

# **A Class of Nonlinear Stochastic Volatility Models and Its Implications for Pricing Currency Options**

**Jun Yu<sup>1</sup>, Zhenlin Yang**

School of Economics and Social Sciences,  
Singapore Management University, Singapore 259756

**Xibin Zhang**

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

## **Abstract**

This paper proposes a class of stochastic volatility (SV) models by applying the Box-Cox transformation to the volatility equation. This class of nonlinear SV models encompasses all standard SV models that have appeared in the literature, including the well known lognormal model. It allows us 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. We develop a likelihood-based technique for analyzing the model. Daily dollar/pound exchange rate data suggest some evidence against lognormal 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.

## **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 to explain the well documented time varying volatility. Empirical successes of the lognormal SV model relative to GARCH-type models are documented in Kim, Shephard and Chib (1998) in terms of in-sample fitting and in Yu (2002) in terms of out-of-sample forecasting.

---

<sup>1</sup>Author for correspondence. Email: yujun@smu.edu.sg. Tel: +65 68280858; fax: +65-68280833.

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). 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 particular, Heston's model has received a great deal of attention in the option price literature as it provides a closed form expression to 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 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), by Jones (2003) and by Meddahi (2001).

Despite all these alternative specifications, there is a lack of simple procedure for selecting an appropriate functional form of SV.<sup>2</sup> 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.

---

<sup>2</sup>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, Duan (1996) and Hentschel (1995). Such a test, however, is by no mean a direct test of stochastic volatility specifications.

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. It also overlaps with the classes of Jones (2003) and Meddahi (2001). Different from these alternative classes which preclude a simple comparison of competing SV models, an advantage of our proposed class is the ease with which coexisting specifications on stochastic volatility can be tested. In fact, the specification test is based on a single parameter. This parameter also provides a measure of degree of departure from classical SV models. 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 several daily exchange rate data. Our empirical test of all standard SV models is, to the best of our knowledge, the first in the literature. When daily dollar/pound data are used, for example, our empirical test suggests some evidence against all standard SV models and favors a nonlinear SV specification. Economic importance of this nonlinearity is 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 LN distribution. Most importantly, an application of our nonlinear 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. However, when other exchange rate series are used, we find that only LN model but not the other classical SV models is suitable. This finding provides a justification why the LN model enjoys the most popularity in the empirical finance literature (see Table 1 for a partial list of studies on the LN model).

The paper is organized as follows. Section 2 presents this class of nonlinear SV 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 economic importance of the proposed models in terms of their implications for pricing

currency options. Finally in Section 6 we present conclusions and possible extensions.

## 2 A Class of Nonlinear 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), Chesney and Scott (1989), and Scott (1991) specify the following model for the asset price  $P(t)$  and the corresponding volatility  $\sigma^2(t)$ ,

$$dP(t)/P(t) = \alpha dt + \sigma(t)dB_1(t), \quad (1)$$

$$d \ln \sigma^2(t) = \lambda(\xi - \ln \sigma^2(t))dt + \gamma dB_2(t), \quad (2)$$

where  $B_1(t)$  and  $B_2(t)$  are two Brownian motions and  $\text{corr}(dB_1(t), dB_2(t)) = \rho$  with  $\rho$  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 *LN* SV model given by

$$X_t = \sigma_t e_t, \quad (3)$$

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

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  $\text{corr}(e_t, v_{t+1}) = \rho$ . The above model is equivalently represented, in the majority of empirical literature, by

$$X_t = \exp\left(\frac{1}{2}h_t\right)e_t, \quad (5)$$

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

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

The LN 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,

$$X_t = \sigma_t e_t, \quad (7)$$

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

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

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

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 LN SV model and the other popular SV models as special cases, ii) it adds great flexibility to the functional form, iii) it provides a 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  $H_0 : \delta = 0$ , 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

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

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

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

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

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 in the nonlinear ARCH (NARCH) model proposed in Higgins and Bera (1992). Obviously, our model provides a 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.

Compared with the NARCH model, however, our proposed class is more closely related to the option literature as the nested models have been used for pricing options.

It can be seen that as  $\delta \rightarrow 0$ ,  $(1 + \delta h_t)^{1/(2\delta)} \rightarrow \exp(0.5h_t)$  and  $((\sigma_t^2)^\delta - 1)/\delta \rightarrow \ln \sigma_t^2$ . Hence the proposed N-SV model includes the LN SV model as a special case. If  $\delta = 1$ , the variance equation (8) becomes

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

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 (8) becomes

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

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  $\delta, \mu$  and  $\phi$ . In fact, all the models except our model require  $\delta$  to be 0, 0.5, or 1.<sup>3</sup> For a general  $\delta$ , our model is different from any of them and  $\delta$  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. Perhaps the most relevant applications to our work may be that proposed by Higgins and Bera (1992) for reasons mentioned above. Another relevant application is Hentschel (1995) who introduces a family of GARCH models by applying the Box-Cox transformation to the conditional

---

<sup>3</sup>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  $\sigma_t^2$  which is identical to assuming  $\sigma_t$  follows a particular OU process.

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.

Table 1

	Models	$\delta$	$\mu$	$\phi$
Taylor (1982), Wiggins (1987), Scott (1987) Chesney and Scott (1989) Jacquier et al (1994), Harvey et al (1998) Kim, Shephard and Chib (1998)	$\ln \sigma_t^2 = \mu + \phi(\ln \sigma_{t-1}^2 - \mu) + \sigma v_t$	0		
Scott (1987), Andersen (1994) 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) 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	$\frac{(\sigma_t^2)^\delta - 1}{\delta} = \mu + \phi[\frac{(\sigma_{t-1}^2)^\delta - 1}{\delta} - \mu] + \sigma v_t$			

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

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

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  $\sigma_t^2$  is nonnegative, the exact normality of  $v_t$  is incompatible unless  $\delta = 0$  or  $1/\delta$  is an even integer.<sup>4</sup> 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.<sup>5</sup> Unfortunately, even in

<sup>4</sup>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)  $\delta$  is small, or ii)  $\mu$  is large, or iii)  $\sigma_h$  is small. See Yang (1999) for a discussion on this.

<sup>5</sup>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.”

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.<sup>6</sup>

To conclude this section, we attempt to offer a heuristic interpretation of  $\delta$  from a finance perspective.<sup>7</sup> For ease of interpretation, we restrict ourselves to the range of positive  $\delta$ . Define  $m = 1/\delta$  and re-write the inverse Box-Cox transformation as

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

where  $\{h_{it}\}$  can be understood as a sequence of intra-day volatility movements. From a market microstructure perspective, intra-day volatility movement 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  $m \rightarrow \infty$  and  $\sigma_t^2 \rightarrow \exp(h_t)$ , new information arrives at the market very frequently. In the N-SV model with a positive, finite value of  $\delta$ , say  $\delta = 0.25$ , on average new information arrives at the market 4 times per day.

### 3 Estimation and Inference Using MCMC

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

Since volatility in SV models is latent, the computation of likelihood requires integrating out the latent variables, which in turn makes the direct likelihood-based analysis numerically difficult. Let  $\mathbf{X} = (X_1, X_2, \dots, X_T)$  be the vector of observations and  $f(\mathbf{X}|\boldsymbol{\theta})$  the likelihood

---

<sup>6</sup>As an alternative, Yu and Yang (2002) approximate the normality of  $u_t = \sigma_t^2 = (1 + \delta h_t)^{1/\delta}$  by a generalized LN distribution. This alternative specification gives rise to the analytic expression for model moments and can be thought of nesting the standard SV models in approximation.

<sup>7</sup>Our treatment here is analogous to the introduction of continuously compounded returns. We are grateful to Steve Satchell for pointing this out to us.



function. To circumvent the problem caused by the latent process, a common practice is to augment the parameter vector to  $(\theta, \mathbf{h})$ , where  $\mathbf{h} = (h_1, h_2, \dots, h_T)$ . 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 joint posterior. When the 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 Nonlinear SV Model

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

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

where  $p$ ,  $S_\sigma$ ,  $\omega$ ,  $\gamma$ ,  $\mu_\delta$  and  $\sigma_\delta^2$  are hyperparameters to be defined by users. When  $\sigma^2$  is integrated out of (16), we can obtain the logarithm of the marginal posterior of  $(\phi, \delta, \mu, \mathbf{h})$ ,

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

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

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

Thus,  $\mu$  can be sampled directly from  $N(\hat{\mu}^*, \hat{\sigma}_\mu^2)$ .<sup>8</sup> Third, we sample  $\sigma^2$  directly from its conditional posterior,

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

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 what follows.

1. Initialize  $\theta$  and  $\mathbf{h}$ ;
2. Sample  $\phi$  and  $\delta$  based on (17), given all the other parameters and  $\mathbf{h}$ ;
3. Sample each component of  $\mathbf{h}$  sequentially according to (16), given  $\theta$ ;
4. Sample  $\sigma^2$  from (19), given all the other parameters and  $\mathbf{h}$ ;
5. Sample  $\mu$  from its conditional posterior, given  $\sigma^2$ ,  $\phi$  and  $\mathbf{h}$ ;
6. Goto 2 and iterate for  $N_0 + N$  times;

---

<sup>8</sup>As  $\mu$  can be sampled directly, its prior has no effect on sampling the other parameters and latent volatilities. That is why we have not inserted a prior of  $\mu$  into the joint posterior (16). When the prior of  $\mu$  is required for further inferences, it can be assumed to be a uniform over a known interval.

where  $N_0$  is the number of iterations in the burn-in period, and  $N$  is the number of draws after the burn-in period.

Two important points should be noted. First,  $\phi$  and  $\delta$  are sampled simultaneously according to the Metropolis-Hastings rule, rather than a single-move procedure.<sup>9</sup> Second, when we calculate the acceptance probability to update a component of  $\mathbf{h}$ , say  $h_t$ , we only compute related items that change with the update of  $h_t$ . Thus, the conditional posterior of  $h_t$ , given parameters and the other components of  $\mathbf{h}$ , is (ignoring end conditions to save space)

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

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

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

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 and 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.

---

<sup>9</sup>See Chib and Greenberg (1995) for detailed discussion on the Metropolis-Hastings algorithms. When updating  $\phi$  and  $\delta$ , the random numbers are generated from the proposal Gaussian density on an elliptical contour. This strategy may increase the sampling efficiency.

### 3.2 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; see, for example, 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 this paper, following Berg et al (2004), we employ Kitagawa's filtering algorithm using 50,000 particle points.

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 (Spiegelhalter et al. (2002)) regardless of properties of the prior (Berg et al. (2004)). We focus on the likelihood ratio test in this paper.

### 3.3 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 simulated 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  $\delta = 0.2$ . This parameter setting is selected to be empirically realistic for daily exchange rates.

In both the simulation and empirical studies, we estimate the N-SV model using the proposed MCMC algorithm. For comparison purposes, we also estimate the LN 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 LN SV model and arbitrarily initialized for the N-SV model. The same prior distributions are used for the common parameters in both models.<sup>10</sup> 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

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

Table 2 summarizes 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 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.

<sup>10</sup>The only exception is for  $\mu$ . In the LN SV model we choose an informative but reasonably flat prior distribution for  $\mu$  (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 precisely estimate all the parameters in the N-SV model, including the key parameter,  $\delta$ . Second, the 90% Bayesian credible interval of  $\delta$  includes the true value and excludes 0. Although not reported, we find that even 99% Bayesian credible interval of  $\delta$  does not include 0.5 or 1. The likelihood ratio statistic favors the true specification and suggests some evidence against the LN model once again. 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) and Mahieu and Schotman (1998)). In this section we estimate 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 mean-corrected 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 LN SV model is the most widely used one in the empirical finance literature, we also estimate it for comparison.

Table 3

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

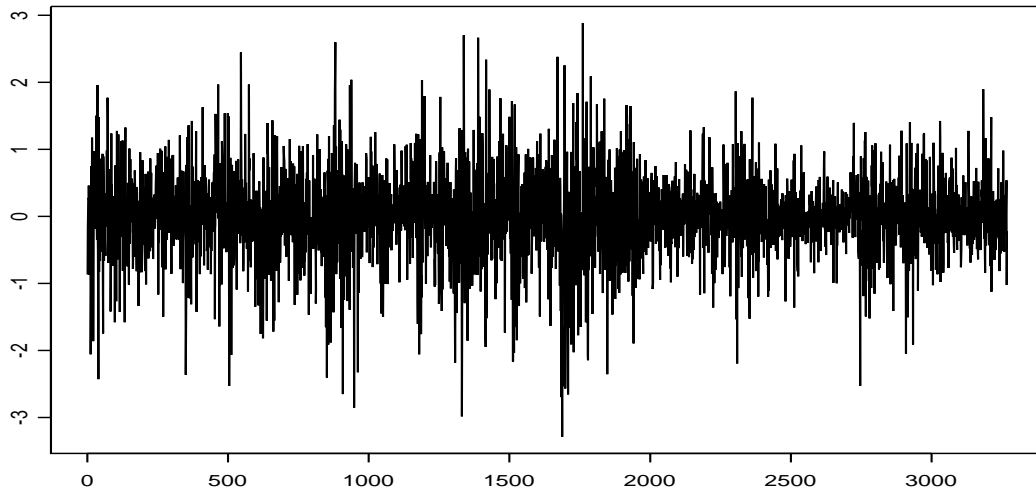


Figure 1: Time series plots for dollar/pound exchange rate return.

Figure 1 displays the adjusted return series. 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 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  $\delta$  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  $\delta$  does not include 0.5 or 1. As a consequence, we find some, albeit weak, evidence against the LN SV model and strong 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  $\delta$  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  $\phi$  (0.9676) is close to 1 in

the LN 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 LN 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 LN model. Fifth, compared with other parameters,  $\delta$  appears more difficult to estimate and has the largest value of standard deviation. Finally, according to our interpretation of  $\delta$ , 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 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., 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 the LN SV model in this case.

As argued in Section 2, a byproduct 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 Figure 4, 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. 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 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



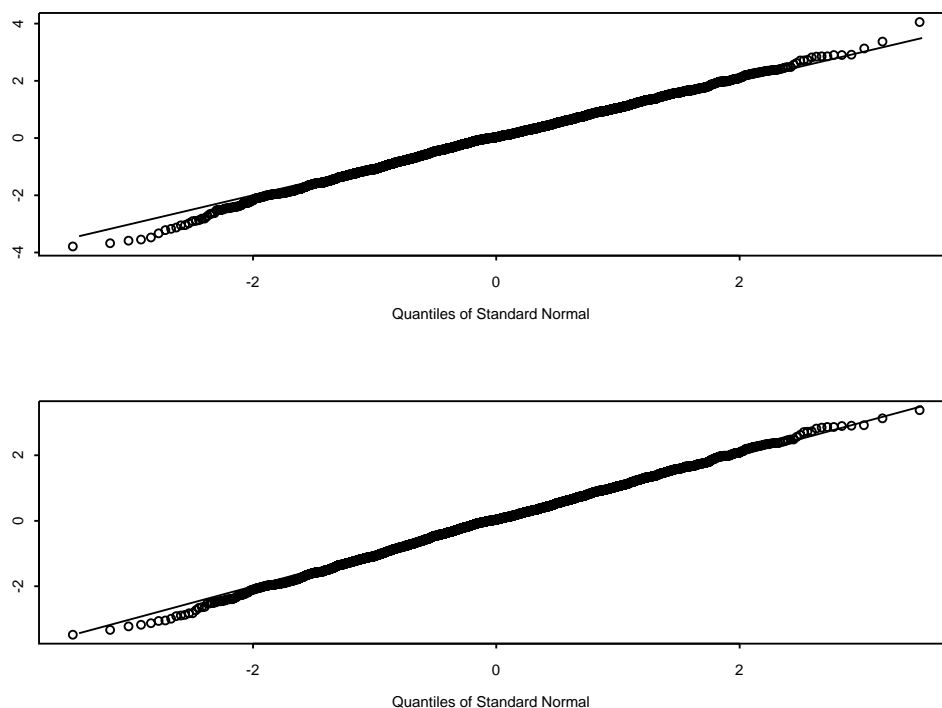


Figure 2: Diagnostic checks of two SV models for dollar/pound exchange rate returns. The first panel is the QQ-plot of the normalized residuals from the LN SV model; the second panel is the QQ-plot of the normalized residuals from the N-SV model.

is not high. When volatility is high, the differences become large. Moreover, we find that the two filter volatilities 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 for 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 economic importance of our N-SV models on option pricing, we price options using both the LN and N-SV models provided 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  $\tau$  (measured in number of days), strike price  $X$ , current volatility  $\sigma_0^2$ , 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

$$C = e^{-\tau r_d} \int_0^\infty \text{BS}(w_\tau) \text{pdf}(w_\tau | h_0) dw_\tau, \quad (20)$$

where  $w_\tau^2$  is given by

$$w_\tau^2 = \int_0^\tau g(h_s, \delta) ds, \quad (21)$$

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

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

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

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

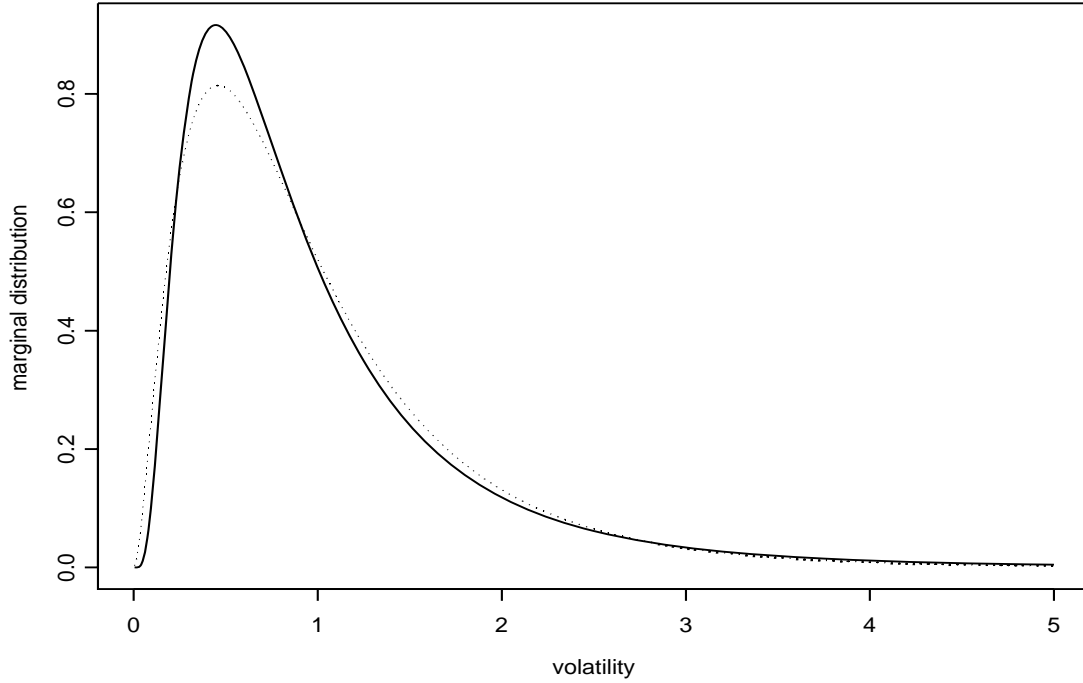


Figure 3: Marginal densities of dollar/pound exchange rate volatility implies from the LN SV model and the N-SV model. The solid line is for the LN SV model; the point line is for the N-SV model.

$$d_2 = d_1 - w_\tau. \quad (24)$$

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

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

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  $\sigma_0^2$ ;
2. Draw independent standard normal variates  $\nu_i$  for  $1 \leq i \leq n$ ;
3. Generate  $h_i$  according to

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

4. Calculate  $w_\tau^2$  using equation (25);
5. Calculate  $BS(w_\tau)$  using equation (22) and call it  $p_1$ ;
6. Repeat Steps 3-5 using  $\{-\nu_i\}$  and define the value of  $BS(w_\tau)$  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 closely related to the one suggested by Mahieu and Schotman (1998), but there are two differences. The first difference is we use an antithetic method in Step 6 to reduce the variance of simulation errors. Finally, we use a much larger value of  $K$  (10,000 as opposed to 500) to ensure that the approximation errors in calculating equation (20) are very small.

The algorithm is then applied to price a half-year call option based on the LN and N-SV models with the estimated parameter values in Table 2 imposed.<sup>11</sup> In both models, we choose  $n = 126$ ,  $S_0 = 1.5$ ,  $r_d = 0$ ,  $r_f = 0$ ,  $K = 10,000$ ,  $\sigma_0 = 0.006349$ ,<sup>12</sup> and  $S_0/X$

---

<sup>11</sup>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.

<sup>12</sup>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.

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 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.

Table 4

	LN 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

## 6 Conclusions and Extensions

In this paper a class of nonlinear SV models has been proposed. The new class facilitates comparing and testing all standard parametric SV models. Since these alternative para-

metric 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. 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 LN models tend to overprice out-of-the-money options. The deeper the out-of-the-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 LN model. As a result, daily exchange volatility is well described by the LN distribution as its marginal distribution, consistent with the results found in recent literature (Andersen et al. (2001)).

Although in this paper we focus on one-factor SV models which are, at the same time, free of jumps, there are some 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 as well as a higher kurtosis than that could be generated from the mixture of distributions, one has to incorporate a leverage effect and a fat tail error distribution into the nonlinear SV model. Also, because 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, Pan and Singleton (2000) and Breidt, Crato and De Lima (1998). Thirdly, although we implement our theory based on the discrete time SV models in the connection with MCMC, 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 nonlinear SV models relative to standard SV models.

## References

- Amin, K.I and V.K. Ng (1993). Jump diffusion option valuation in discrete time. *Journal of Finance* 48, 881-909.
- Andersen, T. (1994). Stochastic autoregressive volatility: A framework for volatility modelling. *Mathematical Finance* 4, 75-102.
- Andersen, T., Bollerslev, T., Diebold, F.X., and Labys, P. (2001). The distribution of realized exchange rate volatility. *Journal of the American Statistical Association*, 96, 42-55.
- Atkinson, A.C. (1986). Monte Carlo tests of separate families of hypothesis. Unpublished Paper, Imperial College, London.
- 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.
- Berg, A., R. Meyer, and J. Yu (2004). Deviance information criterion for comparing stochastic volatility models. *Journal of Business and Economic Statistics*, 22, 107-120.
- Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics* 31, 307-327.
- Broto, C. and E. Ruiz (2002). Estimation methods for stochastic volatility models: A survey. working paper 02-54, Universidad Carlos III de Madrid.
- Box, G.E.P. and D.R. Cox (1964). An analysis of transformations. *Journal of Royal Statistical Society B* 26, 211-243.
- 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.
- 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.

- Chib, S. (1995). Marginal likelihood from the Gibbs output. *The Journal of the American Statistical Association* 90, 1313-1321.
- Chib, S. and E. Greenberg (1995). Understanding the Metropolis-Hastings algorithm. *The American Statistician* 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. *Journal of Econometrics* 79, 97-127.
- Duffie, D., J. Pan and K.J. Singleton (2000). Transform analysis and asset pricing for affine jump-diffusions. *Econometrica* 68, 1343-1376.
- Gallant, A.R. and G. Tauchen (2001). Efficient method of moments. Working paper, Department of Economics, University of North Carolina.
- 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.
- 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.
- Harvey, A.C., E. Ruiz, and N. Shephard (1994). Multivariate stochastic variance models. *Review of Economic Studies* 61, 247-264.
- Hentschel, L. (1995). All in the family: Nesting symmetric and asymmetric GARCH models. *Journal Financial Economics* 39, 71-104.
- 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.
- Higgins, M.L. and A. Bera (1992). A class of nonlinear ARCH models. *International Economic Review* 33, 137-158.



- Hull, J. and A. White (1987). The pricing of options on assets with stochastic volatilities. *Journal of Finance* 42, 281–300.
- 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.
- Johnson, H. and D. Shanno (1987). Option pricing when the variance is changing. *Journal of Financial and Quantitative Analysis* 22, 143–152.
- Jones, C. (2003). The dynamics of stochastic volatility models: evidence from underlying and option markets. *Journal of Econometrics*, 116, 181-224.
- Kim, S., N. Shephard, and S. Chib (1998). Stochastic volatility: Likelihood inference and comparison with ARCH models. *Review of Economic Studies* 65, 361–393.
- Kitagawa, G. (1996). Monte Carlo filter and smoother for Gaussian nonlinear state space models. *Journal of Computational and Graphical Statistics* 5, 1–25.
- Mahieu, R. J. and P.C. Schotman (1998). An empirical application of stochastic volatility models. *Journal of Applied Econometrics* 13, 330–360.
- Meddahi, Nour (2001). An eigenfunction approach for volatility modelling. Manuscript, Department of Economics, University of Montreal.
- Melino, A. and S.M. Turnbull (1990). Pricing foreign currency options with stochastic volatility. *Journal of Econometrics* 45, 239–265.
- Meyer, R. and J. Yu (2000). BUGS for a Bayesian analysis of stochastic volatility models. *Econometrics Journal*, 3, 198–215.
- Nelson, D. (1990). ARCH models as diffusion approximations. *Journal of Econometrics* 45, 7–38.
- Pitt, M. and N. Shephard (1999). Filtering via simulation: Auxiliary particle filter. *The Journal of the American Statistical Association* 94, 590–599.
- 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.

- Shephard, N. and Pitt, M.K. (1997). Likelihood analysis of non-Gaussian measurement time series. *Biometrika* 84, 653–667.
- Spiegelhalter, D. J., Best, N. G., Carlin, B. P., and van der Linde, A. (2002), Bayesian measures of model complexity and fit (with discussion). *Journal of the Royal Statistical Society, Series B*, 64, 583-639.
- Stein, E.M. and J.C. Stein (1991). Stock price distributions with stochastic volatility: An analytical approach. *Review of Financial Studies* 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., Editor, *Time Series Analysis: Theory and Practice*, 1, 203–226. North-Holland, Amsterdam.
- Wiggins, J.B. (1987) Option values under stochastic volatility: Theory and empirical estimate. *Journal of Financial Economics* 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. *Applied Financial Economics* 12, 193–202.
- Yu, J. (2005) On leverage in a stochastic volatility model. *Journal of Econometrics* 128, 165–178.
- Yu, J. and Z. Yang (2002) A class of nonlinear stochastic volatility models. Department of Economics, University of Auckland, Working Paper # 229.