Journal of Modern Applied Statistical Methods

Volume 12 | Issue 1 Article 21

5-1-2013

Parameter Estimation of a Class of Hidden Markov Model with Diagnostics

E. B. Nkemnole
University of Lagos, Nigeria, Africa

O. Abass
University of Lagos, Nigeria, Africa, olabass@unilag.edu.ng

R. A. Kasumu
University of Lagos, Nigeria, Africa

Part of the Applied Statistics Commons, Social and Behavioral Sciences Commons, and the Statistical Theory Commons

Recommended Citation

Nkemnole, E. B.; Abass, O.; and Kasumu, R. A. (2013) "Parameter Estimation of a Class of Hidden Markov Model with Diagnostics," *Journal of Modern Applied Statistical Methods*: Vol. 12 : Iss. 1 , Article 21.
DOI: 10.22237/jmasm/1367382000

Journal of Modern Applied Statistical Methods
May 2013, Vol. 12, No. 1, 181-197
Copyright © 2013 JMASM, Inc.
1538 – 9472/13/$95.00

Parameter Estimation of a Class of Hidden Markov Model with Diagnostics

E. B. Nkemnole O. Abass R. A. Kasumu
University of Lagos,
Nigeria, Africa

A stochastic volatility (SV) problem is formulated as a state space form of a Hidden Markov model (HMM). The SV model assumes that the distribution of asset returns conditional on the latent volatility is normal. This article analyzes the SV model with the student-t distribution and the generalized error distribution (GED) and compares these distributions with a mixture of normal distributions from Kim and Stoffer (2008). A Sequential Monte Carlo with Expectation Maximization (SMCEM) algorithm technique was used to estimate parameters for the extended volatility model; the Akaike Information Criteria (AIC) and forecast statistics were calculated to compare distribution fit. Distribution performance was assessed using simulation study and real data. Results show that, although comparable to the normal mixture SV model, the Student-t and GED were empirically more successful.

Key words: Hidden Markov Model, sequential Monte Carlo, expectation maximization, Student-t distribution, state-space model, stochastic volatility, likelihood, stock exchange.

Introduction

The Hidden Markov Model (HMM), originally introduced in 1957, (MacDonald & Zucchini, 1997, Cappe, et al., 2005) has many applications in fields such as signal processing, medicine, engineering and management. The HMM is a doubly stochastic process, $(X_t, Y_t)_{t\geq 0}$, with an underlying stochastic process, X_t, that is not directly observable but can be observed through another process, Y_t, that produces a sequence of independent random observations.HMM are equivalently defined via a functional representation known as a state space model. The state space model (Doucet & Johansen, 2009) of a HMM is represented by two equations: state (1) and observation (2) as

$$x_t = f(x_{t-1}) + w_t \tag{1}$$

$$y_t = g(x_t) + v_t \tag{2}$$

where f and g are linear or nonlinear function sand w_t and v_t are white noise processes. Models represented by (1) and (2) are referred to as state space models and include a class of HMMs with non-linear Gaussian state-space model such as the stochastic volatility (SV) model. The SV model (Taylor, 1982), accounts for time-varying and persistent volatility and the leptokurtosis in financial return series. The SV model has become popular for explaining the behavior of financial variables, such as stock prices and exchange rates (Durbin & Koopman, 2000; Doucet & Tadic, 2003) and its popularity has resulted in several different proposed approaches for estimating model parameters. Though theoretically attractive, the SV model is empirically challenging due to the fact that the unobserved volatility process enters the model in a non-linear fashion which leads to the likelihood function depending upon high-dimensional integrals.

Estimation procedures, such as the Generalized Method of Moments (GMM)

E. B. Nkemnole is a Lecturer in the Department of Mathematics. Email her at: enkemnole@unilag.edu.ng. O. Abass is a Professor in the Department of Computer Science. Email him at: olabass@unilag.edu.ng. R. A. Kasumu is a Lecturer in the Department of Mathematics. Email him at: bkasumu@yahoo.com.

(Mellino & Turnbull, 1990) and the Efficient Method of Moments (EMM) (Gallant, et al., 1997) have been proposed for the SV model. Other proposed estimation procedures include the method of moments and the quasi maximum likelihood approach methodology to approximate the SV model to a linear Gaussian model (Harvey, et al, 1994; Ruiz, 1994).

Durbin and Koopman (2000) used the idea of linearization of general state-space models and matched terms in the likelihood of a linearized model to those of a linear Gaussian model. Several studies (Jacquier, et al., 1994; Chib, et al., 2002; Kim, et al., 1998) adopted the Gibbs sampling scheme, and Shephard and Pitt (1997) applied the Metropolis-Hastings scheme for the analysis of the SV.Kim and Stoffer (2008) showed how the incorporation of the EM algorithm and SMC (particle filters and smoothers) forms a basic idea to handle the parameter estimation problem in the SV model. Estimation can be accomplished by applying a filtering algorithm. (Kitagawa& Sato, 2001) combined particle filtering methods and gradient algorithms. This article expands the scope of application of SV models, by extending SMC techniques with the EM algorithm developed by Kim and Stoffer (2008) to estimate SV model parameters with the student-t distribution.

The SV model usually assumes that the distribution of asset returns conditional on the latent volatility is normal. However, financial data often have heavier tails than can be captured by the standard SV model: This has led to the use of non-normal distributions to better-model and to deal with the heavy tails (Shephard, 1996; Kim, et al., 1998; Bai, et al., 2003; Sadorsky, 2005; Kim & Stoffer, 2008). Liesenfeld and Jung (2000) fit a Student-t distribution to the error distribution in the SV model using the simulated maximum likelihood method developed by Danielsson and Richard (1993) and Danielsson (1994). A promising distribution that models both skewness and kurtosis is the Skewed Student-t (Fernandez & Steel, 1998). Hence, it is necessary to determine the best-fitted model out of a potentially huge class of candidates; it has become pertinent to develop efficient model selection criteria. As this background illustrates there is an ever-growing literature on time-varying financial market volatility; it is abound with empirical studies in which competing models are evaluated and compared on the basis of their forecast performance (Andersen, et al., 2005).

Stochastic Volatility (SV) Models

SV models belong to class of Hidden Markov model and account for volatility of data. The SV model can be expressed as an autoregressive (AR) process:

$$x_t = \phi x_{t-1} + w_t \tag{3}$$

$$y_t = \beta \exp\left(\frac{x_t}{2}\right) v_t \tag{4}$$

where $w_t \sim N(0,\tau)$, $x_0 \sim N(\mu_0,\sigma_0^2)$, $v_t \sim N(0,1)$, $\{y_t\}_{t\geq 0}$ is the log-return on day t, and β is the constant scaling factor so that $\{x_t\}_{t\geq 0}$ represents the log of volatility of y_t (Taylor, 1982).To ensure stationarity of y_t it is assumed that $|\phi|<1$. Squaring (4) and taking its logarithm results in the linear equation

$$y_t = \alpha + x_t + z_t . \tag{5}$$

Equations (3) and (5) form a version of the SV model that can be modified in many ways; together they form a linear, non-Gaussian, state-space model for which (5) is the observation equation and (3) is the state equation.

Stochastic Volatility with Heavy-Tailed Distribution

The standard form of the SV modelwas shown in equations (3) and (4); in (4) v_t follows a normal distribution. Various authors have argued that real data may have heavier tails than can be captured by the standard SV model.

The Stochastic Volatility Model with Normal Mixture

The observational noise process (Kim & Stoffer, 2008) is a mixture of two normal's with unknown parameters given as

$$y_k = x_t + z_t \quad (6)$$

with $z_t = I_t z_{t1} + (1-I_t) z_{t0}$, $z_{t0} \sim N(m_0, R_0)$, $z_{t1} \sim N(m_1, R_1)$, $m_0 = \alpha - \mu\pi$, $m_1 = \alpha + (1-\pi)\mu$ and $I_t \sim Bernuolli\,(\pi)$. I_t is an indicator variable, where π is an unknown mixing probability, that is, $p(I_t = 1) = \pi = 1 - p(I_t \sim Bernuolli(\pi)$. The likelihood of $\{x_0, \cdots, x_n, y_1, \cdots, y_n, I_1, \cdots, I_n\}$ is shown in Figure 1 where

$$R_t^* = I_t R_1 + (1-I_t) R_0,\ \mu_t^* = I_t q_1 + (1-I_t) q_0.$$

In the SV-normal mixture model defined by (6), the vector of the model parameter is denoted by $\{q_0, q_1, R_0, R_1, \pi\}$. These parameters are estimated along with the other parameters, $\{\phi, \tau\}$ (see Kim & Stoffer, 2008 for details).

Student-t as an Observation Noise

Equations (3) and (5) are an extension of the linearized version of the SV model wherein it is assumed that the observational noise process, z_t is a student-t distribution. The model, first presented in Shumway and Stoffer (2006), retains the state equation for the volatility as: $x_t = \phi x_{t-1} + w_t$.

However, the proposed Student-t distribution with degrees of freedom, v, for the observation error term, z_t, effects a change in the observation equation:

$$y_t = \alpha + x_t + z_t$$
$$z_t \sim t_v,\ t = 1, \cdots, n \quad (7)$$

The distribution of the error term for this specification according to Shimada & Tsukuda (2005) takes the form

$$f(y_t|x_t) = \frac{1}{\sqrt{\pi(v-2)}} \frac{\Gamma\left(\frac{v+1}{2}\right)}{\Gamma\left(\frac{v}{2}\right)} e^{-\frac{x_t}{2}} \left(1 + \frac{y_t^2 e^{-x_t}}{v-2}\right)^{-\frac{v+1}{2}} \quad (8)$$

where v represents a parameter of degree of freedom and Γ stands for the Gamma function. The likelihood function of $\{x_0, x_1, \cdots, x_n, y_1, \cdots, y_n\}$ is

$$f(X,Y) = \left(\frac{1}{\sqrt{2\pi\sigma_0^2}} exp\left(-\frac{1}{2}\left(\frac{x_0 - \mu_0}{\sigma_0} \right)^2 \right) \times \prod_{t=1}^{n} \frac{1}{\sqrt{2\pi\tau}} exp\left(-\frac{1}{2}\left(\frac{x_t - \varphi x_{t-1}}{\tau} \right)^2 \right) \right) \times \prod_{t=1}^{n} \frac{1}{\sqrt{\pi(v-2)}} \frac{\Gamma\left(\frac{v+1}{2}\right)}{\Gamma\left(\frac{v}{2}\right)} e^{-\frac{x_t}{2}} \left(1 + \frac{y_t^2 e^{-x_t}}{v-2}\right)^{-\frac{v+1}{2}}$$

Figure 1: Likelihood of $\{x_0, \cdots, x_n, y_1, \cdots, y_n, I_1, \cdots, I_n\}$
where $R_t^* = I_t R_1 + (1-I_t) R_0$, $\mu_t^* = I_t q_1 + (1-I_t) q_0$

$$f(X,Y,I) = \frac{1}{\sqrt{2\pi\sigma_0^2}} exp\left(-\frac{(x_0 - \mu_0)}{2\sigma_0^2} \right) \prod_{t=1}^{n} \frac{1}{\sqrt{2\pi\tau}} exp\left(-\frac{(x_t - \varphi x_{t-1})^2}{2\tau} \right) \prod_{t=1}^{n} \pi^{I_t} (1-\pi)^{1-I_t} \prod_{t=1}^{n} \frac{1}{\sqrt{2\pi R_t^*}} exp\left(-\frac{(y_t - x_t - \mu_t^*)^2}{2R_t^*} \right)$$

Generalized Error Distribution as an Observation Noise

The distribution of the GED according to Bao, et al. (2006) takes the form

$$f(y_t \mid x_t) = \frac{v \exp\left[-\frac{1}{2}\left((y_t - \alpha - v_t)/\psi\right)^v\right]}{\psi \Gamma\left(\frac{1}{v}\right) 2^{\left(1+\frac{1}{v}\right)}}$$

where

$$\psi = \left[2^{-\frac{2}{v}} \frac{\Gamma\left(\frac{1}{v}\right)}{\Gamma\left(\frac{3}{v}\right)}\right]^{\frac{1}{2}}$$

The log-likelihood function for the GEDmodel is:

$$\log f(y_t|x_t) = n \log v - \frac{1}{2}\sum_{t=1}^{n}\left(\frac{y_t - \alpha - v_t}{\psi}\right)^v - n \log \psi - n \log \Gamma\left(\frac{1}{v}\right) - n \log 2^{\left(1+\frac{1}{v}\right)}.$$

The EM Algorithm

The paramount parameter estimation toolto achieve maximum likelihood estimator is the EM algorithm and it has been widely applied to the cases where the data is considered to be incomplete in the sense that it is not fully observable. It is comprised of the two following steps:

E-step: Compute the expected likelihood,

$$Q(\theta \mid \theta^{(k)})$$
$$Q(\theta \mid \theta') = E(\log f(x \mid \theta') \mid \mathrm{y}, \theta)$$

M-step: Choose $\theta^{(k+1)}$ the parameter values that maximize the function, $Q(\theta \mid \theta^{(k)})$ (for details see Baum, et al, 1970; Dempster, et al., 1977; Rabiner, 1989).

The E- and M- steps are repeated until some stopping criteria is met, such as $|\hat{\theta}^{n+1} - \hat{\theta}^{n}| < Q$, for some specified Q, obtaining suitable initial parameters inclusive. An online EM algorithm recently proposed for discrete HMM can be extended to more general settings, including non-linear non-Gaussian state-space models that necessitate the use of Sequential Monte Carlo (SMC) filtering approximations.

Sequential Monte Carlo Methods (SMC)

After its introduction in the 1960's, SMC has become an emerging methodology for the nonlinear or non-Gaussian state-space models. The chief initiative is to represent the interested density function $p(x_{0:k-1} \mid y_{0:k-1})$ at time $k-1$ by a set of random samples with associated weights, $\{x_{0:k-1}^{(i)}, w_{0:k-1}^{(i)} i = 1, \cdots, N\}$ and compute estimates based on these samples and associated weights. As the number of samples becomes very large, this Monte Carlo characterization develops into an equivalent representation to the functional description of the probability density function (Arulampalam, et al., 2002).

If $\{x_{0:k-1}^{(i)}, w_{0:k-1}^{(i)} i = 1, \cdots, N\}$ are samples and associated weights approximating the density function $p(x_{0:k-1} \mid y_{0:k-1})$, with $\sum_{i=1:N} w_{k-1}^{(i)} = 1$, then the density function is approximated by

$$p(x_{0:k-1} \mid y_{0:k-1}) \approx \sum_{i=1}^{N} w_{k-1}^{(i)} \delta(x_{k-1} - x_{k-1}^{(i)})$$

where $\delta(x)$ signifies the Dirac delta role. The particle approximation $\{w_k^{(i)}, x_k^{(i)}\}_{i=1}^{N}$ is transformed into an equally weighted random sample from $p(x_{0:k-1} \mid y_{0:k-1})$ by sampling, with replacement from the discrete distribution $\{w_k^{(i)}, x_k^{(i)}\}_{i=1}^{N}$. This procedure, also called resampling, produces a new sample with uniformly distributed weights so that $w_k^{(i)} = N^{-1}$.

Particle filters and smoothers are SMC methods grounded in particle representations and are considered generalizations of Kalman filters and smoothers for general state-space models. The fundamental approach used to obtain particles from the desired density is based on sequential importance sampling (SIS) and resampling. SIS, a Monte Carlo method, forms the basis for most particle filtering methods. To approximate the conditional density of x_t given previous states, x_{t-1}, and past and present data, $y_t, p(x_t \mid x_{t-1}, y_t)$, SIS introduces a importance sampling density, $q(x_t \mid x_{t-1}, y_t)$ where it is easier to sample from $\pi(x_t \mid x_{t-1}, y_t)$ than $p(x_t \mid x_{t-1}, y_t)$ (Doucet, et al., 2001).

Particle Filter Algorithm

If at time t weighted particles $\{f_t^{(i)}, w_t^{(i)}\}$ drawn from $f(x_t \mid y_t)$, $f_t^{(i)}$ is a set of particle filter with associated weight $w_t^{(i)}$, then this is considered an empirical approximation for the density comprised of point masses, $f(x_t \mid y_t) \approx \sum_{i=1}^{M} w_t^{(i)} \delta(x_t - f_t^{(i)})$.

Kitagawa & Sato (2001) and Kitagawa (1996) provide an algorithm for filtering in general state space model. This is a Monte Carlo filtering for general state-space models:

1. For $i = 1, \cdots, N$, generate a random number $f_0^{(i)} \sim p(x_0)$

2. Repeat the following steps for $t = 1, \cdots, T$.
   a. For $i = 1, \cdots, N$, generate a random number $w_t^{(i)} \sim q(w)$.
   b. For $i = 1, \cdots, N$, compute $p_t^{(i)} = F(f_{t-1}^{(i)}, w_t^{(i)})$
   c. For $i = 1, \cdots, N$, compute $w_t^{(i)} = p(y_t \mid p_t^{(i)})$
   d. Generate $f_t^{(i)}, i = 1, \cdots N$ by resampling $p_t^{(i)}, \cdots, p_t^{(N)}$

The Monte Carlo filter returns $\{ f_t^{(i)}, i = 1, \cdots, N, t = 1, \cdots, m \}$ so that $\sum_{i=1}^{N} \frac{1}{N} \delta(x_t - f_t^{(i)}) \approx f(x_t \mid Y_t)$.

Particle Smoothing Algorithm

If $\{s_t^{(i)}, w_t^{(i)}\}_{i=1}^{M}$ is a set of particle smoothers and associated weights approximating the density function $f(x_t \mid Y_n)$, then the density function is approximated by:

$$f(x_t \mid Y_n) \approx \sum_{j=1}^{M} w_t^{(i)} \delta(x_t - s_t^{(i)}).$$

The problem with smoothed estimates is degeneracy. Godsill, et al. (2004) suggested a new smoothing method (particle smoother using backwards simulation). The method assumes that filtering has already been performed, thus, the particles and associated weights, $\{f_t^{(i)}\}_{i=1}^{M}$, $\{w_t^{(i)}\}_{i=1}^{M}$ can approximate the filtering density,

$$f(x_t \mid Y_t), \text{ by } = \frac{\sum w_t^{(i)} \delta(x_t - f_t^{(i)})}{\sum_{i=1}^{N} w_t^{(i)}}.$$

The algorithm from Godsill, et al. (2004) supposes that weighted particles $\{f_t^{(i)}, w_t^{(i)}; i = 1, 2, \cdots, M\}$ are available for $t = 1, 2, \cdots, n$. The algorithm for $i = 1, 2, \cdots, M$ is:

1. Choose $s_n^{(i)} = f_n^{(j)}$ with probability $w_n^{(j)}$

2. For $n-1$ to 1
   a. Calculate $w_{t|t+1}^{(j)} \propto w_t^{(j)} f(s_{t+1}^{(i)} \mid f_t^{(j)})$ for each j.
   b. Choose $s_t^{(i)} = f_t^{(j)}$ with probability $w_{t|t+1}^{(j)}$.

3. $s_{1:n}^{(i)} = (s_1^{(i)}, \cdots, s_n^{(i)})$ is an approximate realization from $p(X_n \mid Y_n)$.

Sequential Monte Carlo Expectation Maximization (SMCEM) Algorithm Analysis
Parameter Estimation

The SMCEM estimation procedure consists of three main steps: filtering, smoothing and estimation. Parameter estimation for the Student-t and GED SV model were considered. A basic approach for the Student-t SV model, equation (7), is to apply the EM algorithm; with the output of filtering and a smoothing step an approximate expected likelihood is calculated.

Filtering Step

The algorithm for the filtering and smoothing steps shows a slight modification of Godsill, et al. (2004) and Kim and Stoffer (2008). M samples from $f(x_t, \mid Y_t)$ for each t were obtained as:

1. Generate $f_0^{(i)} \sim N(\mu_0, \sigma_0^2)$
2. For $t = 1, \cdots, n$
   a. Generate a random number $w_t^{(i)} \sim N(0, \tau), \quad j = 1, \cdots, M$
   b. Compute $p_t^{(i)} = \phi f_{t-1}^{(i)} + w_t^{(i)}$
   c. Compute
   $$w_t^{(i)} = p(y_t \mid p_t^{(i)},) \propto e^{-\frac{x_t}{2}} \left(1 + \frac{y_t^2 e^{-x_t}}{v-2}\right)^{-\frac{v+1}{2}}$$
   d. Generate $f_t^{(i)}$ by resampling with weights, $w_t^{(j)}$

Smoothing Step

In the smoothing step, particle smoothers that are needed to acquire the expected likelihood in the expectation step of the EM algorithm were obtained. Suppose that equally weighted particles $\{f_t^{(i)}\}, i = 1, \cdots, M$ from $f(x_t, \mid Y_t)$ are available for $t = 1, \cdots, n$, from the filtering step.

1. Choose $[s_n^{(i)}] = [f_n^{(j)}]$ with probability $\frac{1}{M}$
2. For $n-1$ *to* 0 calculate
$$w_{t|t+1}^{(i)} \propto f(s_{t+1}^{(i)} | f_t^{(j)}) \propto exp\left(-\frac{(s_{t+1}^{(i)} - \varphi f_t^{(j)})^2}{2\tau}\right)$$
$$\frac{1}{\sqrt{\pi(v-2)}} \frac{\Gamma\left[\frac{v+1}{2}\right]}{\Gamma\frac{v}{2}} exp^{-\frac{\tilde{s}_{t+1}}{2}} \left(1 + \frac{y_t^2 e^{-\tilde{s}_{t+1}^{(j)}}}{v-2}\right)^{-\frac{v+1}{2}}$$
   for each j
   a. Choose $[s_t^{(i)}] = [f_t^{(j)}]$ with probability $w_{t|t+1}^{j}$.
3. $(s_{0:n}^{(i)}) = \{(s_0^{(i)}, \cdots, s_n^{(i)})\}$ is the random sample from $f(x_0, \cdots, x_n \mid Y_n)$
4. Repeat 1-3, for $i = 1, \cdots, M$ and calculate
$$\hat{x}_t^n = \frac{\sum_{i=1}^{M} s_t^{(i)}}{M}, \quad \hat{p}_t^n = \frac{\sum_{i=1}^{M} (s_t^{(i)} - \hat{x}_t^n)^2}{M-1},$$
$$\hat{p}_{t,t-1}^n = \frac{\sum_{i=1}^{M} (s_t^{(i)} - \hat{x}_t^n)(s_{t-1}^{(i)} - \hat{x}_{t-1}^n)}{M},$$
$$E\left[1 + \frac{y_t^2 e^{x_t}}{v-2}\right]^{-\frac{v+1}{2}} = \frac{n(v-2)}{(v+1)\sum_{t=1}^{n} y_t^2 e^{-y_t + v_t} \left[1 + \frac{y_t^2 e^{x_t}}{v-2}\right]^{-1}}$$

Estimation Step

This step consists of obtaining parameter estimates by setting the derivative of the expected likelihood, of the complete data $\{x_0, \cdots, x_n, y_1, \cdots, y_n\}$ given $\{x_0, \cdots, x_n\}$, with respect to each parameter to zero and solving for $\hat{\phi}$, $\hat{\tau}$, and $\hat{\alpha}$. The complete likelihood of $\{x_0, x_1, \cdots, x_n, y_1, \cdots, y_n\}$ is

$$\log f(X,Y) =$$

$$\log \frac{1}{\sqrt{2\pi}} \frac{1}{\sigma_0} + \log \exp\left(-\frac{(x_0 - \mu_0)^2}{2\sigma_0^2} \right)$$

$$+ \log \prod_{t=1}^{n} \frac{1}{\sqrt{2\pi\tau}} \exp\left(\frac{-(x_t - \varphi x_{t-1})^2}{2\tau} \right)$$

$$+ \left(\begin{array}{l} \log \prod_{t=1}^{n} \frac{1}{\sqrt{\pi(v-2)}} \frac{\Gamma\left(\frac{v+1}{2}\right)}{\Gamma\left(\frac{v}{2}\right)} e^{-\frac{(y_t - \alpha - v_t)}{2}} \\ \left(1 + \frac{y_t^2 e^{-(y_t - \alpha - v_t}}{v-2} \right)^{-\frac{v+1}{2}} \end{array} \right)$$

This method results in the estimates:

$$\hat{\phi} = \frac{S_{10}}{S_{00}}, \quad \hat{\tau} = \frac{1}{n}\left[S_{11} - \frac{S_{10}^2}{S_{00}} \right],$$

$$\hat{\alpha} = \log \frac{n(v-2)}{(v+1)\sum_{t=1}^{n} y_t^2 e^{-y_t + v_t} \left[1 + \frac{y_t^2 e^{x_t}}{v-2} \right]^{-1}}$$

$$\hat{\alpha} = \left[n \sum_{t=1}^{n} (y_t - v_t)^{v-1} \right]^{\frac{1}{v-1}}$$

where

$$S_{00} = \sum_{t=1}^{n} (x_{t-1}^n)^2 + p_{t-1}^n,$$

$$S_{11} = \sum_{t=1}^{n} (x_t^n)^2 - p_t^n, \; S_{10} = \sum_{t=1}^{n} x_t^n x_{t-1}^n + p_{t,t-1}^n$$

When z_t follows the GED, it is not possible to represent it as equation (3.2). Hence, for the SV-GED model, the parameter v, as well as the other parameters, and x_t were sampled from their full conditional distributions using SMCEM techniques.

Methodology

The proposed method to compare the fit of the distributions is illustrated using three simulated data sets and daily exchange rates of the Nigerian Naira, Ghana Cedi, British Pound and Euro compared to the U. S. Dollar, from March 3, 2009 to March 3, 2011. Figures 1-3 show the plots and histograms of data generated from the normal mixture, Student-t and GED SV model respectively and Tables 1-4 show the results of the estimation for the models.

Simulation 1

Data were generated from the normal mixture SV model $x_t = 0.7x_{t-1} + w_t$, $y_t = -2.75 + x_t + v_t$ where $w_t \sim N(0, 0.96)$, $v_t \sim I_t N(-2, 6) + (1 - I_t) N(-3.5, 4)$ and $I_t \sim Bernoulli(0.5)$ with true parameter set $(\phi, \tau, q_0, q_1, R_0, R_1, \pi)$ = (0.7, 0.96, –3.5, –2, 4, 6, 0.5). The technique based on mixture and Student-t SV was applied to this data to examine the performance of the proposed model. To make the process stationary, 11,000 samples were generated and the first 10,000 values were discarded. Figures 2a and 2b show the plot and histogram for Simulation 1.

Simulation 2

Data were generated from the Student-t SV model with true parameter set (ϕ, τ, α, v) = (0.81, 1.45,–3.01, 8). The technique based on the mixture and Student-t SV models was applied to this data to examine the merit of the Student-t idea; the length of the data, $\{y_t\}$, was 1,000. Figures 3a and 3b show the plot and histogram for Simulation 2. The second data set was used to observe the behavior of the estimation procedure when a departure from the normal mixture observational error assumption exists.

Figure2a: Representation of SMCEM SequenceSimulated from the Normal Mixture SV Model

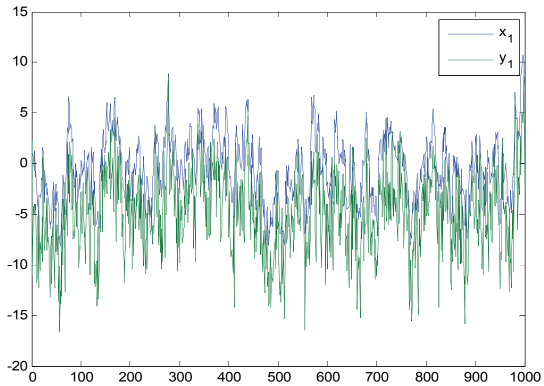


Figure 2b: Histogram of Final Values of Parameters of the Normal Mixture SV Model

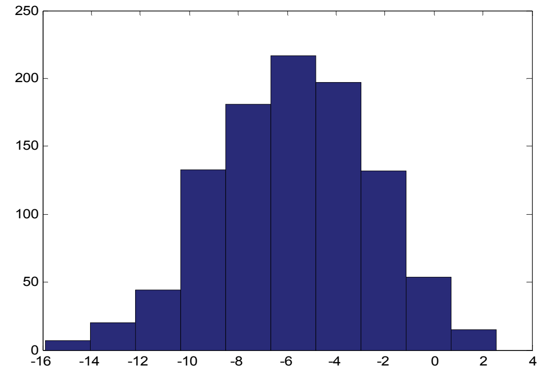


Figure3a: Representation of SMCEM Sequence Simulated from the Student-t SV Model

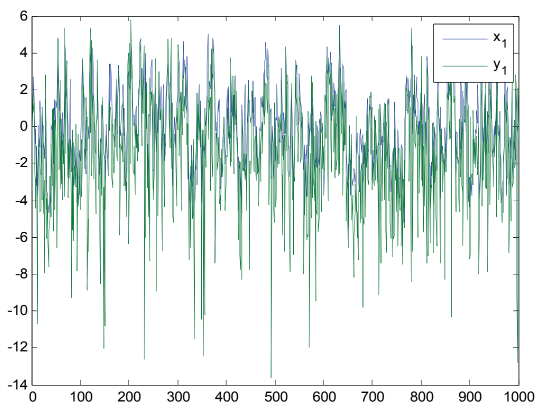


Figure 3b: Histogram of Final Values of the Parameters of the Student-t SV Model

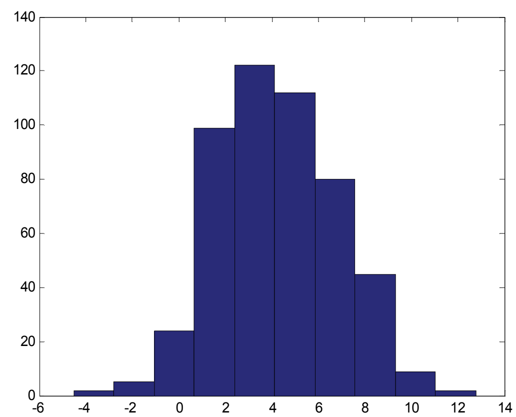


Figure4a: Representation of SMCEM Sequence Simulated from the Normal Mixture SV Model

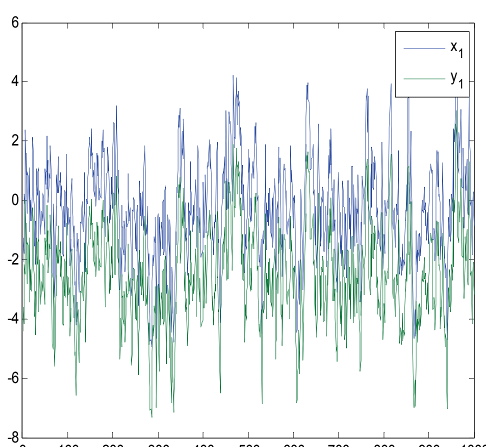
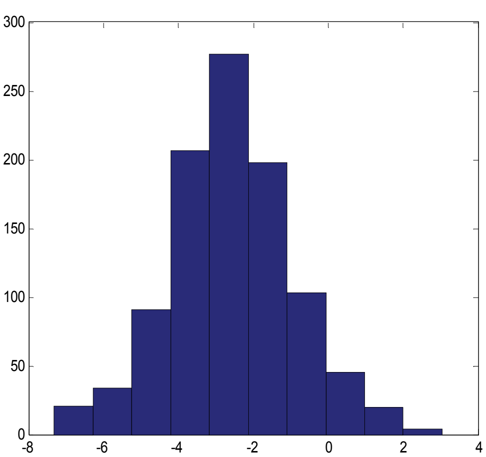


Figure 4b: Histogram of Final Values of Parameters of the Normal Mixture SV Model



Simulation 3

Data were generated from the GED SV model with true parameter set (ϕ, τ, α, v) = (0.9, 1.6, 0.7). Techniques based on mixture and the GED SV model were applied to this data to examine the merit of the GED idea; the length of the data, $\{y_t\}$, is 1,000. Figures 4a and 4b show the plot and histogram for Simulation 3.

Results

Using the procedures described [0.9500, 1.0729, −0.6794, −3.6794, 4.000, 4.000, 0.5000] were selected for the initial parameters for $(\phi, \tau, q_0, q_1, R_0, R_1, \pi)$. Table 1 shows final estimates with their standard error (in parenthesis) for Simulation 1. The final estimates, along with their standard deviations (in parentheses), were: $\hat{\phi}$ = 0.7568 (0.027826), $\hat{\tau}$ = 0.3466 (0.00931), $\hat{q}_0$ = −1.9486 (0.10989), $\hat{q}_1$ = −3.7620 (0.08690), $\hat{R}_0$ = 2.3169 (0.20936), $\hat{R}_1$ = 7.5643 (0.67241), $\hat{\pi}$ = 0.3854 (0.02635) where the true parameters are (0.7, 1.06, −3.5, −2, 4, 6, 0.5). In this approach, $\hat{\alpha} = \hat{\pi}\hat{q}_1 + (1-\hat{\pi})\hat{q}_0$ = −2.6475; $(\hat{\phi}, \hat{\tau}, \hat{\alpha})$ = (0.7568, 0.3466, −2.6475); based on results, the estimation procedure based on the normal mixture model works well because that the estimates are close to the true parameters.

Based on the Student-t technique, (0.9500, 1.0729, −2.1496) were used as the initial values for parameters (ϕ, τ, α); the process was stopped when the value of relative likelihood was less than 0.001. The final estimates, along with their standard deviations (in parentheses) were: $\hat{\phi}$ = 0.6913 (0.037981), $\hat{\tau}$ = 1.0336 (0.14839), $\hat{\alpha}$ =−2.9009 (0.024501). These results show that the model provides good estimates despite the fact that the true observation noise is not a normal mixture distribution.A similar simulation study was performed using the data from simulation 2 (see Table 2).

The initial parameter set [0.8214, 1.3359, −2.7823, −5.7823, 4.000, 4.000, 0.5000] was selected for parameters $(\phi, \tau, q_0, q_1, R_0, R_1, \pi)$. Table 2 shows the results of the parameter estimation procedure based on the normal mixture. The final estimates, along with their standard deviations (in parentheses) were: $\hat{\phi}$ = 0.6547 (0.005272), $\hat{\tau}$ = 1.2930 (0.002473), $\hat{q}_0$ = −3.0180 (0.0.035241), $\hat{q}_1$ = −5.8536 (0.012445), $\hat{R}_0$ = 3.3275 (0.15), $\hat{R}_1$ = 5.2663 (0.32564), $\hat{\pi}$ = 0.4806 (0.004338) where the true parameters are (0.81, 1.45, −3.01) for the parameters, $(\hat{\varphi}, \hat{\tau}, \hat{\alpha})$; where $\hat{\alpha} = \hat{\pi}\hat{q}_1 + (1-\hat{\pi})\hat{q}_0$ = −4.3808.

When the data from simulation 2 was fitted with the techniques based on the Student-t (0.8214, 1.3359, −2.2823) were used as initial parameters of (ϕ, τ, α). At the 11th iteration the relative likelihood was less than 0.001 and the process was considered converged. The final estimates, along with their standard deviations were: $\hat{\phi}$ = 0.8383 (0.008552), $\hat{\tau}$ = 1.5357 (0.12403), $\hat{\alpha}$ =−3.0912 (0.005302). These estimates are similar to the true parameters (0.81, 1.45, −3.01), while Normal mixture returns (0.6547, 1.2930,−4.3808) as $(\hat{\phi}, \hat{\tau}, \hat{\alpha})$. The method based on the Student-t SV model worked well in both cases. When the estimation procedure based on the normal mixture SV model was applied, the estimates were distant to the true parameter. Conversely, the application of the technique based on Student-t model indicated a better proximity to the true parameters; therefore, extension of the SV model by adopting Student-t is meaningful.

Table 3 shows the results of the parameter estimation procedure on technique based on the normal mixture SV and GED on data generated from the normal mixture model; [0.8699, 3.6899, −4.8897, −7.8897, 4.000, 4.000, 0.5000] were selected for the initial parameter for the parameters $(\phi, \tau, q_0, q_1, R_0, R_1, \pi)$. The final estimates, along with their standard deviations (in parentheses) were: $\hat{\phi}$ = 0.9869 (0.0013341), $\hat{\tau}$ = 4.1936 (0.51468), $\hat{q}_0$ = −4.3900

(0.021891), $\hat{q}_1 = -6.4554$ (0.026309), $\hat{R}_0 = 3.7342$ (0.069579), $\hat{R}_1 = 4.3517$ (0.16268), $\hat{\pi} = 0.4895$ (0.0032833) where the true parameters are (0.8, 3.5, −5, −8, 3, 4.1, 0.5). In this approach, $\hat{\alpha} = \hat{\pi}\,\hat{q}_1 + (1-\hat{\pi})\,\hat{q}_0 = -5.4010$; $(\hat{\phi}, \hat{\tau}, \hat{\alpha})$ = (0.9869, 4.1936, −5.4010). Results show that the estimation procedure based on the normal mixture model worked well in the sense that the estimates are close to the true parameters.

For the GED technique, (0.8699, 3.6899, −4.4365,) were used as the initial values for parameters (ϕ, τ, α). Table 3shows the results of the estimation procedure. The process was stopped when the value of relative likelihood was less than 0.001. The final estimates, along with their standard deviations (in parentheses) were: $\hat{\phi}$ = 0.8127 (0.0027854), $\hat{\tau}$ = 4.2368 (0.020652), $\hat{\alpha}$ =–4.7144 (0.0083309). Results show that the GED model gives good estimates even though the true observation noise is not a normal mixture distribution.

Table 4 shows results of the parameter estimation procedure on technique based on the normal mixture SV and GED on data generated from the GED SV model. The method based on the GED model works well in both cases. When the estimation procedure based on the normal mixture SV model was applied, the estimates were far from the true parameters. By contrast, the application of the technique based on GED model indicated a better proximity to the true parameters. (0.9500, 1.3288, 0.6309) were used as initial parameters (ϕ, τ, α). The final estimates, along with their standard deviations were: $\hat{\phi}$=0.9749 (0.0026845), $\hat{\tau}$=2.3496 (0.2678), $\hat{\alpha}$=0.6821 (0.014247). These estimates are similar to the true parameters (0.9, 1.6, 0.7) while the normal mixture returns (0.7515, 2.3496, 1.3259732) as (ϕ, τ, α). Thus, the method based on the GED works well in both cases.

Table 1: Parameter Estimates and Standard Errors (in parenthesis) on Technique Based on Normal Mixture and Student-t on Data Generated from Normal Mixture Model

	True Parameter	Normal Mixture SV	Student-t SV	Normal Mixture SV	Student-t SV	Normal Mixture SV	Student-t SV
		$M = 500$	$\varepsilon = 0.1$	$M = 1000$	$\varepsilon = 0.01$	$M = 1000$	$\varepsilon = 0.001$
$\phi^{(i)}$	0.7	0.7368 (0.028438)	0.6976 (0.01475)	0.8677 (0.048671)	0.7308 (0.049383)	0.7568 (0.027826)	0.6913 (0.037981)
$\tau^{(i)}$	1.06	0.9408 (0.014796)	1.3654 (0.0029492)	0.1186 (0.0059242)	0.7625 (0.20677)	0.3466 (0.00931)	1.0336 (0.14839)
${q_0}^{(i)}$	−3.5	-0.9826 (0.085676)		−0.8746 (0.084237)		−1.0486 (0.10989)	
${q_1}^{(i)}$	-2	-3.5081 (0.037916)		−3.4744 (0.00879)		−3.7620 (0.08690)	
${R_0}^{(i)}$	4	2.2144 (0.38686)		1.8883 (0.27135)		2.3169 (0.20936)	
${R_1}^{(i)}$	6	7.5758 (1.0062)		7.9490 (0.59001)		7.5643 (0.67241)	
$\pi^{(i)}$	0.5	0.4330 (0.017748)		0.4267 (0.014455)		0.3854 (0.02635)	
$\alpha^{(i)}$	−2.75	−2.0761	−2.1465 (0.0038743)	−1.983935	−1.6703 (0.10626)	−2.094344	−2.9009 (0.024501)
Rel. Lik		0.0865	0.0250	0.0025	−0.0021	−0.0004	−0.000

Table 2: Parameter Estimation on Technique Based on the Normal Mixture and Student-ton Data Generated from the Student-t Model

	True Parameter	Normal Mixture SV	Student-t SV	Normal Mixture SV	Student-t SV	Normal Mixture SV	Student-t SV
		$M = 500$	$\varepsilon = 0.1$	$M = 1000$	$\varepsilon = 0.01$	$M = 1000$	$\varepsilon = 0.001$
$\phi^{(i)}$	0.81	0.6388 (0.021863)	0.8439 (0.005971)	0.5601 (0.052847)	0.8693 (0.002036)	0.6547 (0.005272)	0.8383 (0.008552)
$\tau^{(i)}$	1.45	1.2585 (0.029249)	1.2696 (0.0029492)	1.5805 (0.092768)	1.4500 (0.037812)	1.2930 (0.002473)	1.5357 (0.12403)
${q_0}^{(i)}$		−2.8256 (0.035245)		−2.8979 (0.02815)		−3.0180 (0.035241)	
${q_1}^{(i)}$		−5.7964 (0.011406)		−5.7544 (0.002826)		−5.8536 (0.012445)	
${R_0}^{(i)}$		3.8225 (0.075907)		3.8111 (0.057768)		3.3275 (0.15)	
${R_1}^{(i)}$		4.6134 (0.14959)		4.2066 (0.060831)		5.2663 (0.32564)	
$\pi^{(i)}$		0.4911 (0.0025628)		0.4911 (0.001607)		0.4806 (0.004338)	
$\alpha^{(i)}$	−3.01	−4.28456	−3.1243 (0.000596)	−4.300727	−3.1645 (0.005662)	−4.3808	−3.0912 (0.005302)
Rel. Lik		0.0643	0.0045	0.0057	0.0042	0.0009	−0.0010

Table 3: Parameter Estimation on Technique Based On the Normal Mixture SV and GED on Data Generated from the Normal Mixture Model

	True Parameter	Normal Mixture SV	GED SV	Normal Mixture SV	GED SV	Normal Mixture SV	GED SV
		$M = 500$	$\varepsilon = 0.1$	$M = 1000$		$M = 1000$	
$\phi^{(i)}$	0.8	0.7875 (0.0045855)	0.8485 (0.0059856)	0.8425 (0.0046651)	0.8788 (0.0037027)	0.9869 (0.0013341)	0.8127 (0.0027854)
$\tau^{(i)}$	3.5	3.4771 (0.17068)	4.0273 (0.13077)	2.9591 (0.17631)	3.0502 (0.052579)	4.1936 (0.51468)	4.2368 (0.020652)
${q_0}^{(i)}$	−5	−5.5322 (0.0059081)		−4.8080 (0.026411)		−4.3900 (0.021891)	
${q_1}^{(i)}$	−8	−8.5269 (0.00932)		v7.9658 (0.032108)		−6.4554 (0.026309)	
${R_0}^{(i)}$	3	3.6937 (0.10267)		3.7054 (0.081339)		3.7342 (0.069579)	
${R_1}^{(i)}$	4.1	4.4847 (0.16612)		4.1121 (0.057281)		4.3517 (0.16268)	
$\pi^{(i)}$	0.5	0.4895 (0.003056)		0.4930 (0.0025453)		0.4805 (0.0032833)	
$\alpha^{(i)}$	−5.55	−6.9980	−4.3205 (0.035379)	−6.3648	−4.2576 (0.003807)	−5.3824	−4.7144 (0.0083309)
Rel. Lik		0.0248	0.0853	−0.0148	0.0064	−0.0684	0.0007

Table 4: Parameter Estimation on Technique Based On the Normal Mixture SV and GED on Data Generated from the GED Model

	True Parameter	Normal Mixture SV	GED SV	Normal Mixture SV	GED SV	Normal Mixture SV	GED SV
		$M=500$	$\varepsilon=0.1$	$M=1000$		$M=1000$	
$\phi^{(i)}$	0.9	0.9050 (0.015127)	0.9770 (0.0021463)	0.8871 (0.026021)	0.9754 (0.0025267)	0.7515 (0.015215)	0.9749 (0.0026845)
$\tau^{(i)}$	1.6	0.3136 (0.01557)	2.1311 (0.2265)	0.4605 (0.025703)	2.3108 (0.26119)	1.2777 (0.069801)	2.3496 (0.2678)
$q_0^{(i)}$		0.1491 (0.08715)		−0.3568 (0.11521)		−0.3193 (0.088286)	
$q_1^{(i)}$		−2.4081 (0.1269)		−2.6609 (0.13263)		−3.1988 (0.15449)	
$R_0^{(i)}$		2.9473 (0.2282)		2.5108 (0.28846)		1.6365 (0.47054)	
$R_1^{(i)}$		7.8467 (0.84311)		9.4387 (1.1097)		8.7981 (1.1994)	
$\pi^{(i)}$		0.4448 (0.015076)		0.3916 (0.028448)		0.3496 (0.041351)	
$\alpha^{(i)}$	0.7	−0.9883	0.7654 (0.018455)	1.2590	0.7627 (0.021224)	−1.3259732	0.6821 (0.014247)
Rel. Lik		0.0052	0.0248	−0.0125	0.0057	0.0003	0.0007

Application to Real Life Financial Data

The normal mixture, Student-t and GED SV model were applied to analyze daily rates on the Naira/Dollar, Cedi/Dollar, Pound/Dollar and Euro/Dollar exchange rates from March 3, 2009 to March 3, 2011. Figures 5-8 show the plots of the daily exchange rates and log returns of the data.Patterns of behavior are evident in the second plots in Figures 5-8: the data experience a small variance for some periods of time, and for other periods they show a large variance. For this reason, it cannot be assumed that the data have a constant variance.

Table 6 presents the estimation results along with their standard deviations for the Student-t, normal mixture and the GED SV models. These distributions produce comparable maximum likelihood values, indicating an acceptable overall fit. The values (ranging from 0.927 to 0.988) suggest high persistence of the volatility of the series indicatingthat volatility clustering is observed in all the exchange rates return series.

The Akaike values and the evaluation statistics using the all data are shown in Table 7. The AIC and the log-likelihood values highlight the fact that (GED) Student-t distribution better estimates the series than the normal mixture distribution for the SV model. In fact, the log-likelihood function increases, leading to AIC criteria of 2.805, 3.4593, 3.9989 and 9.6632 with the normal mixture versus (2.7814304) 2.776433, (3.4223827) 3.391374, (3.9749741) 3.969968 and (9.6513786) 9.646376 with the non-normal densities, for the Naira/Dollar, Cedi/Dollar, Pound/Dollar and Euro/Dollar rate respectively. The statistics from the volatility forecasts (Sadorsky, 2005) are presented. In terms of MSE, the Student-t performs better than the normal mixture for the Naira/Dollars and the Euro/Dollar exchange rate while the opposite is true for the Cedi/Dollars and Pound/Dollar exchange rate. Generally, the MAE results are not different from the MSE results. In terms of MAPE, the Student-t SV model is preferred in three cases and the GED SV model once.

Table 5: Descriptive Statistics of Daily Returns for the Exchange Rate

Statistics	Naira/Dollar Rate	Cedi/Dollar	Pound/Dollar	Euro/Dollar
Mean	–0.001385	–0.006258	0.020803	0.001676
Standard Deviation	0.708650	0.536507	0.506541	0.488392
Skewness	–0.074139	0.966923	0.022958	0.434943
Kurtosis	8.805879	13.11769	4.262290	7.993814
Jarque-Bera	735.0376	2312.255	34.76827	559.9343

Figure5: Naira/Dollar Daily Exchange Rate and Log Returns

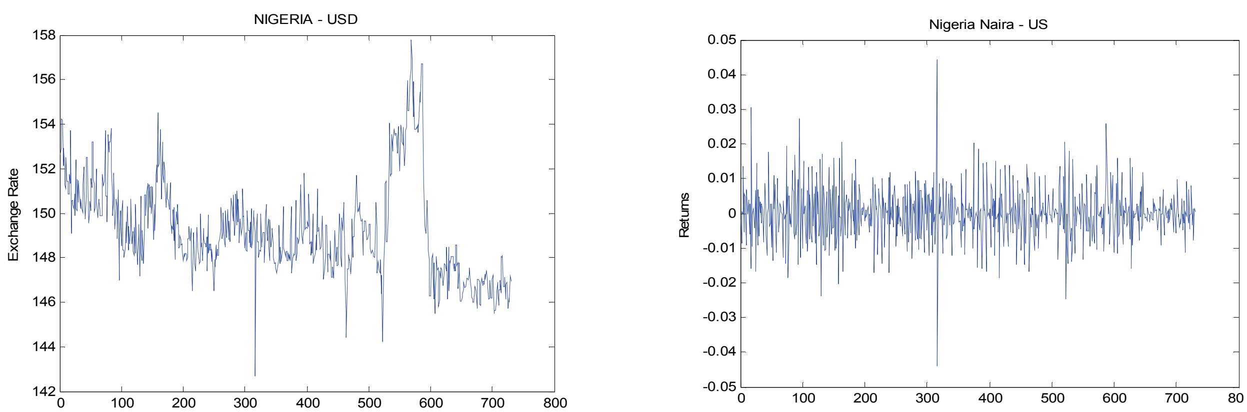


Figure 6: Cedi/Dollar Exchange Rate and Log Returns

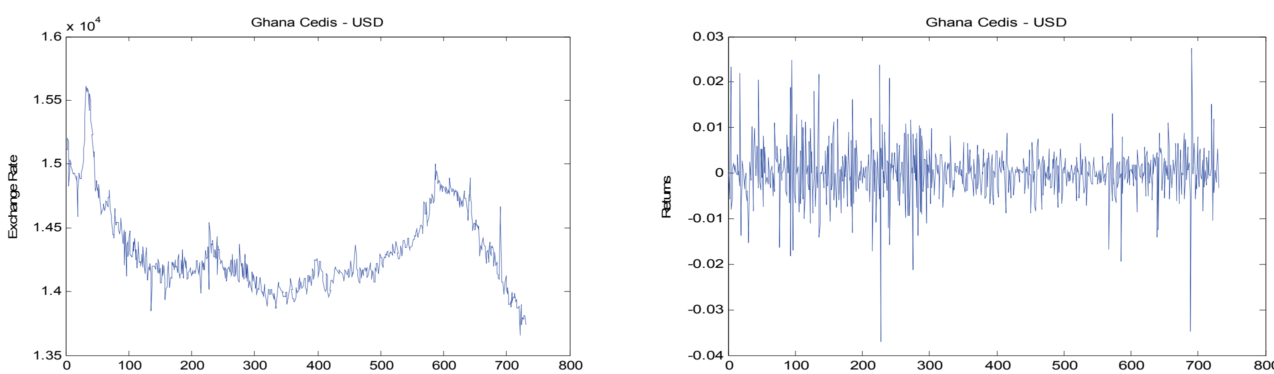


Figure 7: Euro/Dollar Daily Exchange Rate and Log Returns

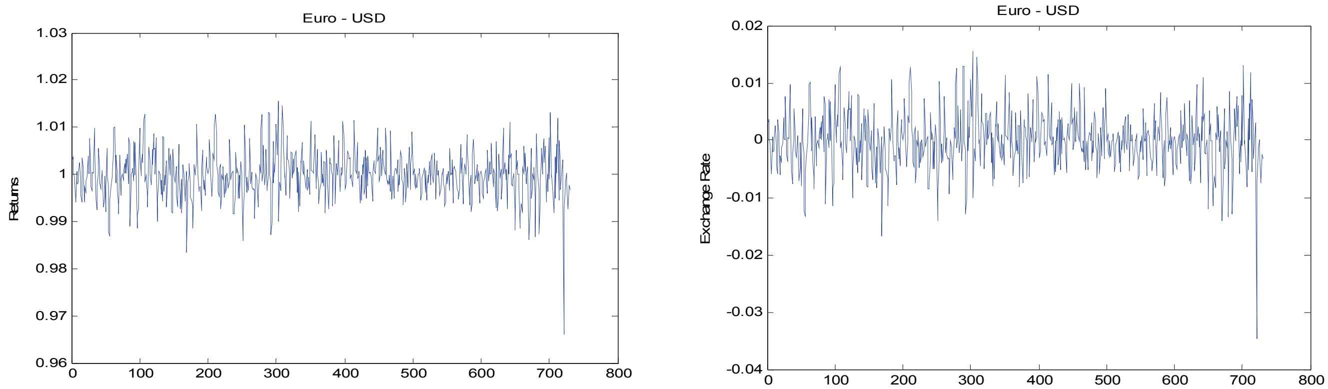
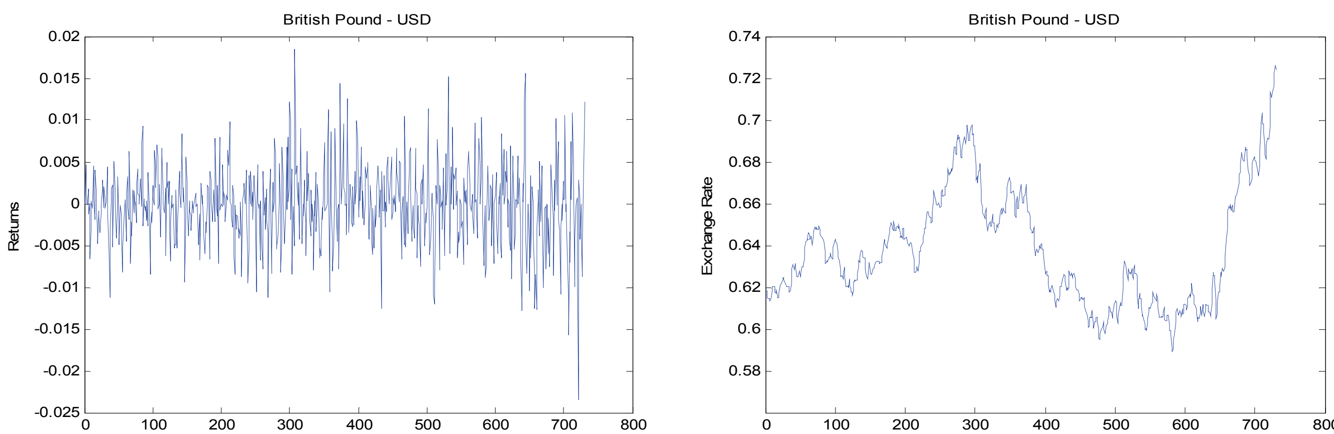


Figure 8: Pound/Dollar Daily Exchange Rate and Log Returns



Conclusion

An extension of the observation error in the SV model from normal mixture to Student-t and GED distributions was presented. A sequential Monte-Carlo expectation maximization experiment was used to estimate parameters for the extended SV model. Functions provided by MATLAB enabled techniques based on the Student-t and GED SV model to be developed along with a strategy for fitting a model that combines the EM algorithm and SMC; this change to the proposed model allowed for a more robust fit, providing a new tool to explore the tail fit. The Student-t and GED SV model was compared with the normal mixture. The EM algorithm makes it possible to obtain maximum likelihood estimators. The estimation Algorithm was completed by applying the Godsill, et al. (2004) particle smoothing algorithm to the SV model with (3) and (5) as the observation and state equations. The outcome of the simulation and real data analyses confirm the viability of the proposed method. Results show that the proposed estimation algorithm yields acceptable results when the normal assumption is violated as well as when it holds, thus widening the range of application of the SV model.Statistics were calculated to compare the fit of distributions. Results, based on data from the Naira/Dollar, Cedi/Dollar, Pound/Dollar and Euro/Dollar exchange rates, reveal that the Student-t is comparable to the normal mixture SV model but is empirically more successful.

Table 6a: Estimation Results - Distribution Comparison

	Naira/Dollar			Cedi/Dollar		
	Normal Mixture SV (SD)	Student-t SV (SD)	GED SV (SD)	Normal Mixture SV (SD)	Student-t SV (SD)	GED SV (SD)
$\phi^{(i)}$	0.9759 (0.0041766)	0.9769 (0.0029595)	0.9684 (0.015815)	0.983 (0.0049939)	0.9887 (0.00078342)	0.9741 (0.01001)
$\tau^{(i)}$	0.0988 (0.019774)	0.1300 (0.16604)	0.3227 (0.030638)	0.0902 (0.033484)	0.0854 (0.0057923)	0.09831 (0.048961)
$q_0^{(i)}$	3.0673 (0.043267)			3.4686 (0.015039)		
$q_1^{(i)}$	1.6633 (0.65851)			0.9826 (0.22583)		
$R_0^{(i)}$	0.6954 (0.11387)			0.1865 (0.015231)		
$R_1^{(i)}$	1.9713 (0.88689)			2.3618 (0.10685)		
$\pi^{(i)}$	0.0546 (0.017691)			0.0705 (0.006188)		
$\alpha^{(i)}$	−2.0731	−2.1245 (0.5369)	−2.8836 (0.35001)	−4.28356	−3.2243 (0.0019277)	−4.9834 (0.0022972)
	Pound/Dollar			Euro/Dollar		
	Normal Mixture SV (SD)	Student-t SV (SD)	GED SV (SD)	Normal Mixture SV (SD)	Student-t SV (SD)	GED SV (SD)
$\phi^{(i)}$	0.9895 (0.0013354)	0.9754 (0.0025267)	0.9697 (0.0084966)	0.9579 (0.0090241)	0.9113 (0.022502)	0.9763 (0.0021063)
$\tau^{(i)}$	0.7114 (0.36943)	0.7627 (0.26119)	0.9018 (0.183815)	0.4170 (0.24553)	0.4731 (0.45638)	0.7258 (0.361172)
$q_0^{(i)}$	3.3388 (0.022015)			−1.9638 (0.21412)		
$q_1^{(i)}$	0.4351 (0.023938)			−4.3710 (0.54791)		
$R_0^{(i)}$	0.7835 (0.056033)			4.1677 (1.1257)		
$R_1^{(i)}$	4.4129 (0.14603)			8.1096 (0.4471)		
$\pi^{(i)}$	0.4941 (0.0028666)			0.4404 (0.03083)		
$\alpha^{(i)}$	−4.28356	2.3108 (0.021224)	1.9588 0.0082487	−4.28356	−2.2692 (0.17681)	−3.5033 (0.058033)

Table 7: Evaluation Statistics - Distribution Comparison

	AIC	Log-like	MSE	MAE	MAPE
Naira/Dollar					
Normal Mixture SV	2.805	2537.403	0.2601914	0.210278	0.133428
Student-t SV	2.776433	2565.97	0.1111914	0.075379	0.053437
GED SV	2.7814304	2563.16	0.1131914	0.081256	0.064035
Cedi/Dollar					
Normal Mixture SV	3.4593	2652.836	0.037074	0.059524	0.147437
Student-t SV	3.391374	2769.66	0.121086	0.171621	0.126563
GED SV	3.4223827	2760.56	0.16377	0.188346	0.131424
Pound/Dollar					
Normal Mixture SV	3.9989	2801.144	0.151623	0.072823	1.192232
Student-t SV	3.969968	2869.07	0.178377	0.098361	0.045245
GED SV	3.9749741	2857.27	0.184677	0.103932	0.034675
Euro/Dollar					
Normal Mixture SV	9.6632	2874.968	0.170612	0.260176	0.0988
Student-t SV	9.646376	2903.9	0.095612	0.194364	0.0578
GED SV	9.6513786	2897.8	0.101414	0.196355	0.0634

References

Andersen, T. G., Bollerslev, T., & Diebold, F. X. (2005). *Parametric and nonparametric financial econometrics*. Amsterdam: North-Holland.

Arulampalam, M. S., Maskell, S., Gordon, N., & Clapp, T. (2002). A tutorial on particle filter for on-line non-linear/non-Gaussian Bayesian tracking. *IEEE Transaction on Signal Processing*, *50*, 174-188.

Bai, X., Russell, J. R., & Tiao, G. C. (2003). Kurtosis of GARCH and Stochastic Volatility Models with Non-normal Innovations. *Journal of Econometrics*, *114*, 349-360.

Bao, Y., Tae-Hwy, L., & Burak, S. (2006). Comparing density forecast models. *Journal of Forecasting*, *26*(*3*), 203-225.

Baum, L. E., Peterie, T., Souled, G., & Weiss, N. (1970). A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains. *Annals of Mathematical Statistics*, *41*(*1*), 164-171.

Cappe, O., Moulines, E., & Ryden, T. (2005). *Inference in hidden Markov models*. New York, NY: Springer.

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

Danielsson, J., & Richard, J. F. (1993). Accelerated Gaussian importance sampler with application to dynamic latent variable models. *Journal of Applied Econometrics*, *8*, 153-173.

Danielson, J. (1994). Stochastic volatility in asset prices: Estimation with simulated maximum likelihood.*Journal of Econometrics*, *61*, 375-400.

Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society*, *39*(*1*), 1-38.

Doucet, A., De Freitas, J. F. G., & Gordon, N. (2001). *Sequential Monte Carlo methods in practice.* New York, NY: Springer-Verlag.

Doucet, A., & Tadic, B. B. (2003). Parameter estimation in general state-space models using particle methods. *Annals of Institute of Statistical Mathematics*, *55*, 409-422.

Doucet A., & Johansen, A. M. (2009). A tutorial on particle filtering and smoothing: Fifteen years later. In *Oxford handbook of nonlinear giltering*, D. Crisan & B. Rozovsky (Eds.), 231-260. Oxford: Oxford University Press.

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

Fernandez, C., & Steel, M. (1998). On Bayesian modelling of fat tails and skewness. *Journal of the American Statistical Association*, *93*, 359-371.

Gallant, A. R., Hsieh, D., & Tauchen, G. (1995). Estimation of stochastic volatility model with diagnostics. *Journal of Econometrics*, *81*, 159-192.

Godsill, S., Doucet, A., & West, M. (2004). Monte Carlo smoothing for non-linear time series. *Journal of the American Statistical Association*, *199*, 156-168.

Harvey, A. C., Ruiz, E., & Shephard, N. (1994). Multivariate stochastic variance models. *Review of Economic Studies*, *61*, 247-264.

Jacquier, E., Polson, N. G., & Rossi, P. E. (1994). Bayesian analysis of stochastic volatility models. *Journal of Business and Economic Statistics*, *12*(*4*), 371-417.

MacDonald, I. L., & Zucchini, W. (1997). *Hidden Markov and other models for discrete-valued time series, Vol. 70: Monographs on statistics and applied probability*. London: Chapman & Hall.

Kim, S., Shephard, N., & Chib, S. (1998). Stochastic volatility: Likelihood inference and comparison with ARCH models. *The review of economic studies*, *65*(*3*), 361-393.

Kim, J., & Stoffer, D. S. (2008). Fitting stochastic volatility models in the presence of irregular sampling via particle methods and the EM algorithm. *Journal of Time Series Analysis*, *29*(*5*), 811-833.

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

Kitagawa, G., & Sato, S. (2001). Monte Carlo smoothing and self-organising state space model. In *Sequential Monte Carlo methods in practice*, A. Doucet, N. de Freitas & N. Gordon, Eds.,177-195. New York, NY: Springer-Verlag.

Liesenfeld, R., & Jung, R. C. (2000). Stochastic volatility models: conditional normality versus heavy-tailed distributions. *Journal of Applied Econometrics*, *15*, 137-160.

MacDonald, I. L., & Zucchini, W. (1997). *Hidden Markov and other models for discrete-valued time series, Vol. 70: Monographs on statistics and applied probability*. London: Chapman and Hall.

Mellino, A., & Turnbull, S. (1990). Pricing foreign currency options with stochastic volatility. *Journal of Econometrics*, *45*, 7-39.

Rabiner, L. R. (1989). A tutorial on hidden Markov Models and selected applications in speech recognition. *Proceedings of the IEEE*, *77*(*2*), 257-286.

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

Sadorsky, P. (2005). Stochastic volatility forecasting and risk management. *Applied Financial Economics*, *15*, 121-135.

Shephard, N. (1996). Statistical aspects of ARCH and stochastic volatility. In *Time series models in econometrics, finance and other fields*, D. R. Cox, D. V. Hinkley & O. E. Barndorff-Nielsen (Eds.), 1-67. London: Chapman and Hall.

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

Shimada, J., & Tsukuda, Y. (2005). Estimation of stochastic volatility models: Anapproximation to the nonlinear state space representation. *Communications in Statistics - Simulation and Computation*, *34*, 429-450.

Shumway, R. H., & Stoffer, D. S. (2006). *Time series analysis and its applications*. New York, NY: Springer.

Taylor, S. J. (1982). Financial returns modelled by the product of two stochastic processes: A study of daily sugar prices, 1961-1979. In *Time series analysis: Theory and practice*, O. D. Anderson (Ed.), 203-226. New York, NY: Elsevier Science Publishing Co.