Modeling of Stock Indices with HMM-SV Models

Nkemnole E. B.[1*], Wulu J. T. [2**]
*Department of Mathematics, University of Lagos, Nigeria.
enkemnole@unilag.edu.ng [1],
** Department of Mathematics & Statistics, University of Maryland University College
john.wulu@faculty.umuc.edu [2]

Abstract
The use of volatility models to conduct volatility forecasting is gaining momentum in empirical literature. However, it is known that volatility persistence, as indicated by the estimated parameter ϕ , in Stochastic Volatility (SV) model is typically high. Since future values in SV models are based on estimation of the parameters, this may lead to poor volatility forecasts. Furthermore, this high persistence, according to some research scientists, is due to the structure changes (e.g. shift of volatility levels) in the volatility processes, which SV model cannot capture. Hidden Markov Models (HMMs) allow for periods with different volatility levels characterized by the hidden states. This work deals with the problem by bringing in the SV model based on Hidden Markov Models (HMMs), called HMM-SV model. Via hidden states, HMMs allow for periods with different volatility levels characterized by the hidden states. Within each state, SV model is applied to model conditional volatility. Through Empirical analysis using the proposed HMM-SV models does not only address the structure changes, but also, provides better volatility forecasts and establishes an efficient forecasting structure for volatility modeling.

Keywords: Forecasting, Hidden Markov model, Stochastic volatility, stock exchange

1.0 Introduction

A great deal of attention has been paid in finance, as well as in empirical literature for practically measuring risk to modeling and forecasting the volatility of stock market indices via stochastic volatility (SV) model. No doubt, forecasting the volatility of stock indices is an important aspect of many financial decisions. For instance, investment managers, option traders and the financial managerial bodies are all interested in volatility forecasts in order to either construct less risky portfolios or obtain higher profits (Panait and Slavescu, 2012). Various volatility models have been recommended to describe the statistical features of financial time series. The most common models among these are the SV models. Other volatility models include the autoregressive conditional heteroskedasticity (ARCH) models by Engle (1982) and extended to generalized ARCH (GARCH) by Bollerslev, et al., (1995). Their success lies in their ability to capture some empirical stylized facts of financial time series, such as time-varying volatility and volatility clustering.

SV modeling has been applied to time-varying volatility (Taylor, 1982, 1986). For example, SV models can be used to model the variance as an unobserved component that follows a particular stochastic process. In SV models, it is usual to model volatility as a logarithmic first order autoregressive process. This model, though theoretically attractive, is empirically challenging as the unobserved volatility process enters the model in a nonlinear fashion which leads to the likelihood function depending upon high-dimensional integrals.

In this paper, we propose a solution to the problem of high in SV models by bringing in Hidden Markov Models (HMM) to allow for different volatility states (periods with different volatility levels) in time series. Persistence parameter ϕ estimated as being close to unity, implies a high degree of volatility persistence. By using HMM, within each state, we enable the SV model to model the conditional variance. The ensuing HMM-SV models indeed yields better volatility forecast compared to SV models for both simulated data and real financial data sets.

2. Related Literature

Two recent studies that compare the usefulness of the SV model with GARCH models in applied forecasting situations can be seen in So et al. (1999) and Yu (2002). So et al. (1999) ascertained that in modeling and forecasting foreign exchange rates, the SV model estimated as a state space model does not outperform GARCH model. Yu (2002), on his own part, used the SV model to forecast daily stock market volatility for New Zealand. By means of forecast accuracy tests, he discovered that the SV model surpasses performance of GARCH models. The mixed results from these two papers suggest the need for further research on the relative merits of SV models in applied forecasting situations.

Although standard SV models improve the in-sample fit a lot compared with constant variance models, numerous studies find that SV models give unsatisfactory forecasting performances, (Figlewski, 1997). Xiong-Fei and Lai-Wan (2004) argued that the usually overstated volatility persistence in SV models may be the cause of poor forecasting performances. Lamoureux (1990), shows that this well-known high persistence may originate from the structure changes in the volatility processes, which SV models cannot capture. Lamoureux demonstrated that any shift in the structure of financial time series (e.g. the shift of unconditional variance) is likely to lead to misestimating of the SV parameters in such a way that they entail too high a volatility persistence.

Nelson (1991) and Glosten, et al., (1993) have used the GARCH model to compute the effects of negative and positive shocks on volatility. They find different effects for positive and negative unexpected returns, but both lead to variance increases. Kim, et al. (1998), in recent times, applied Hidden Markov model (HMM), instead of ARCH, to handle the effects of volatility in economic data. Here, the difference between the HMM and ARCH is the unconditional variance (shift in the structure of volatility levels of financial time series} . If there are sequential changes in regime, some researchers advise that some more intuitive approaches need to be considered, and using different regimes may contribute to the return-generating process in the market.
Diebold (1986) and Lamoureux and Lastrapes (1990) contended that the high estimated value for persistence parameter may reflect structural changes that occurred during the sample in the variance process. This is related to Perron (1989) observation that changes in regime may give the spurious impression of unit roots in characterizations of the level of a series.

Hamilton and Susmel (1994) apprehended that the long run variance could obey regime shift; they suggested an ARCH process that will allow the parameters of an ARCH process to come from one of several different regimes, with transitions between regimes governed by an unobserved Markov chain. The effect will vanish if they use weekly data, because sparse time point makes the dependence weaker. In using HMM, Chu et al., (1996) chose a two-stage process to represent the return behavior in the stock market. They first considered the return behavior in stock market as a Markov process. Then, the different return regimes derived from the first stage were utilized to estimate the volatility. Lastly, they found that the negative deviations in returns can have larger increases in volatility than the positive ones. Accordingly, they think the return and volatility are not linear but asymmetrical.

2. HMM-SV model

2.1 Hidden Markov Model

Hidden Markov Model (HMM), originally introduced in 1957 and early 1970's, (see MacDonald & Zucchini, 1997, Cappe et al. 2005) has found many applications in most contemporary fields like the signal processing, medicine, engineering, and management applications. Thus, the contemporary reputation of statistical methods of HMM is not in question. A HMM is a bivariate discrete-time process $\{X_k, Y_k\}_{k\geq 0}$ where $\{X_k\}_{k\geq 0}$ is a homogeneous Markov chain which is not directly observed but can only be observed through $\{Y_k\}_{k\geq 0}$ that produce the sequence of observation. $\{Y_k\}_{k\geq 0}$ is a sequence of independent random variables such that the conditional distribution of Y_k only depends on X_k .The underlying Markov chain $\{X_k\}_{k\geq 0}$ is called the state sequence.

HMM are also defined through a functional representation known as state space model. The state space model (Doucet and Johansen, 2009) of a HMM is represented by the following two equations:

(State equation) $$x_t = F_t(x_{t-1}, w_t) \quad (1)$$

(Observation equation) $$y_t = H_t(x_t, v_t) \quad (2)$$

where f and g are either linear or nonlinear functions, while w_t and v_t are error terms. Models represented by (1) - (2) comprises a class of HMMs which includes non-linear Gaussian state-space models, such as the stochastic volatility (SV) models to be defined in the next section

2.2 Stochastic volatility model

Stochastic volatility models (see Shephard (1996) for a review) are a variant of non-linear Gaussian state-space model which take variation in the volatility of the observed data into account. The SV model due to Taylor (1982) can be expressed as an autoregressive (AR) process:

$$x_t = \phi x_{t-1} + w_t \quad (3)$$

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

where $w_t \sim N(0,\tau), x_0 \sim N(\mu_0, \sigma_0^2), v_t \sim N(0,1), \{r_t\}_{t\geq 0}$ is the log-returns on day t, ,we call β the constant scaling factor, so that $\{x_t\}_{t\geq 0}$ represents the log of volatility of the data, $\log(\sigma_t^2)$ where $\sigma_t^2 = \text{var}(r_t)$ .In order to ensure stationarity of r_t , it is assumed that $|\phi|<1$. Taking the logarithm of the square of equation (4), results in a linear equation,

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

where

$y_t = \log(r_t^2), \alpha = \log(\beta^2) + E\ \log(v_t^2)$, $z_t = \log(v_t^2) - E(\log v_t^2)$. $v_t^2 \sim \chi_1^2$ so that z_t has a centered $\log \chi_1^2$ distribution.

Equations (3) & (5) form the version of the SV model which 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.

2.2.1 Stochastic Volatility with heavy –tailed distribution

The standard form of the SV model is given in equations (3) & (4). In equation (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.

A modification of the linearized version of the SV model (see equation (3) and (5), wherein it is assumed that the observational noise process, z_t is a student-t distribution is considered. The model, first presented in Shumway and Stoffer (2006), retains the state equation for the volatility as:

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

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

$$y_t = \alpha + x_t + z_t \qquad z_t \sim t_v,\ t = 1, \cdots, n, \tag{6}$$

For the parameter estimates of the proposed SV model with student-t, the likelihood functions have been maximized by using the Sequential Monte Carlo Expectation Maximization algorithm (Nkemnole et al., 2015) in the MATLAB optimization routines.

2.3 HMM with stochastic volatility Model

Our model is a blend of the original SV model and HMMs. To start with, we use HMMs to divide the entire time series into regimes with different volatility levels. The return of the time series is assumed to be modeled by a mixture of probability densities and each density function corresponds to a hidden state with its mean and variance. In the HMMs, Sequential Monte Carlo Expectation Maximization algorithm (SMCEM) algorithm is employed in finding the state sequence in the time series (Nkemnole, 2014) which consists of three main steps: filtering, smoothing, and estimation. Subsequently we get the subsets of original time series corresponding to different states (volatility levels). Afterwards, within each regimes, we allow SV model with different parameter sets to model the conditional variance as:

$$x_t = \phi^i\ x_{t-1} + w_t,\quad w_t \sim N(0, \tau)$$

$$y_t = \alpha^i + x_t + z_t$$

where i denotes the state of the time series at time t. ϕ^i, τ^i, and α^i are the parameter sets of the SV model related to state i.

Then, for the volatility forecast σ_t^2 , $(\{x_t\}_{t\geq 0}$ represents the log of volatility of the data, $\log(\sigma_t^2)$ where $\sigma_t^2 = \text{var}(r_t))$ of the global model, there is need for us to predict the state i of time series at time $t+1$ (next state).

After the next state i at time $t+1$ has been determined, we choose the corresponding SV model with parameter sets ϕ^i , τ^i, and α^i to make volatility forecast.
Criteria for assessing the accuracy of the models to predict which includes mean absolute error (MAE), mean square error (MSE), mean absolute percentage error (MAPE) are listed on section 4. SPSS and MATLAB were used to analyse the data to produce figures and results of the models.

2.4 Sequential Monte Carlo Expectation Maximization (SMCEM) Algorithm Analysis

Estimation procedures

The entire estimation procedure consists of three main steps: filtering, smoothing, and estimation. With the output of filtering and smoothing step an approximate expected likelihood is calculated. $\{\phi, \tau, \alpha\}$ are estimated to model the changing volatility.

2.4.1 Filtering Step:

The algorithm for the filtering and smoothing steps below are sane of the work of Nkemnole, et al. (2015) From here M samples from $f(X,|Y)$ for each t were obtained as follows.

i) Generate $f_0^{(i)} \sim N(\mu_0, \sigma_0^2)$

For $t = 1, \cdots, n$

ii) Generate a random number $w_t^{(i)} \sim N(0, \tau), \quad j = 1, \cdots, M$

iii) Compute $p_t^{(i)} = \phi\, f_{t-1}^{(i)} + w_t^{(i)}$

a.Compute $\omega_t^{(i)} = p(y_t \mid p_t^{(i)},) \propto e^{-\frac{x_t}{2}} \left(1 + \frac{y_t^2 e^{-x_t}}{v-2}\right)^{-\frac{v+1}{2}}$

b. Generate $f_t^{(i)}$ by resampling with weights, $\omega_t^{(i)}$

To save computing time, it is essential to start with good initial parameters. The initial value of τ and ϕ are obtained based on the method of moments, Anderson et al. (1969).

2.4.2 Smoothing step

In the smoothing step, particle smoothers that are needed to get the expected likelihood in the expectation step of the EM algorithm were gotten:

Suppose that equally weighted particles $\{f_t^{(i)}\}, i = 1, \cdots, M$ from $f(x_t,| Y_t)$ 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 $t = n-1$ to 0

a)Calculate

$$\omega_{t|t+1}^{(i)} \propto f(s_{t+1}^{(i)} | f_t^{(j)}) \propto \exp\left(-\frac{(s_{t+1}^{(i)} - \phi f_t^{(j)})^2}{2\tau}\right) \frac{1}{\sqrt{\pi(\nu-2)}} \frac{\Gamma\left[\frac{\nu+1}{2}\right]}{\Gamma \frac{\nu}{2}} \exp^{-\frac{\tilde{s}_{t+1}}{2}} \left(1 + \frac{y_t^2 e^{-\tilde{s}_{t+1}^{(j)}}}{\nu-2}\right)^{-\frac{\nu+1}{2}}$$

for each j

b) Choose $[s_t^{(i)}] = [f_t^{(j)}]$ with probability $\omega_{t|t+1}^{j}$.

3) $(s_{0:n}^{(i)}) = \{(s_0^{(i)}, \cdots, s_n^{(i)})\}$ is the random sample from $f(x_0, \cdots, x_n | Y_n)$

4) Repeat 1-3, for $i = 1, \cdots, M$ and calculate

$$\hat{x}_t^n = \frac{\sum_{i=1}^{M} s_t^{(i)}}{M}, \hat{p}_t^n = \frac{\sum_{i=1}^{M} (s_t^{(i)} - \hat{x}_t^n)^2}{M-1}, \hat{p}_{t,t-1}^n = \frac{\sum_{i=1}^{M} (s_t^{(i)} - \hat{x}_t^n)(s_{t-1}^{(i)} - \hat{x}_{t-1}^n)}{M},$$

$$E\left[1 + \frac{y_t^2 e^{x_t}}{\nu-2}\right]^{-\frac{\nu+1}{2}} = \frac{n(\nu-2)}{(\nu+1)\sum_{t=1}^{n} y_t^2 e^{-y_t+v_t}\left[1 + \frac{y_t^2 e^{\hat{x}_t}}{\nu-2}\right]^{-1}}$$

2.4.3 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 $\{x_0, x_1, \cdots, x_n, y_1, \cdots, y_n\}$

$$\log f(X,Y) = \log \frac{1}{\sqrt{2\pi}} \frac{1}{\sigma_0} + \log \exp\left(-\frac{(x_0 - \mu_0)^2}{2\sigma_0^2}\right) + \log \prod_{t=1}^{n} \frac{1}{\sqrt{2\pi\tau}} \exp\left(\frac{-(x_t - \phi x_{t-1})^2}{2\tau}\right)$$

$$+ \log \prod_{t=1}^{n} \frac{1}{\sqrt{\pi(\nu-2)}} \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)} e^{-\frac{(y_t - \alpha - v_t)}{2}} \left(1 + \frac{y_t^2 e^{-(y_t - \alpha - v_t}}{\nu-2}\right)^{-\frac{\nu+1}{2}} \quad (7)$$

By the above method, we got the following estimates

$$\hat{\phi} = \frac{S_{10}}{S_{00}}, \quad \hat{\tau} = \frac{1}{n}\left[S_{11} - \frac{S_{10}^2}{S_{00}}\right] \quad (8)$$

$$, \quad \hat{\alpha} = \log \frac{n(\nu-2)}{(\nu+1)\sum_{t=1}^{n} y_t^2 e^{-y_t+v_t}\left[1 + \frac{y_t^2 e^{x_t}}{\nu-2}\right]^{-1}} \quad (9)$$

where

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

$$S_{11} = \sum_{t=1}^{n} (x_{t}^{n})^2 - p_{t}^{n} \ ,$$

$$S_{10} = \sum_{t=1}^{n} x_{t}^{n} x_{t-1}^{n} + p_{t,t-1}^{n}$$

3 Volatility Forecast Evaluation and Comparison

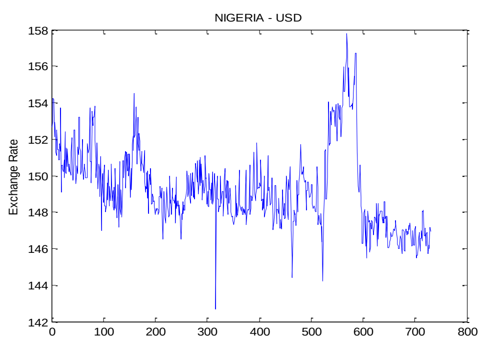
3.1 Data and Methodology

Both simulated data sets and real financial data sets were utilized in the volatility forecast experiments. Also, the in-sample and the out-of-sample forecasting performances were considered. To start with, we used simulated data set to verify if the proposed model solves the problems of excessive persistence in SV model; we generated more than 11000 observations and discarded the initial 10000 samples.

Then, we employed the use of real financial data sets in our experiments to establish the viability of the proposed model. The real financial data sets consist of the daily exchange rate series of the Nigerian Naira, Ghana Cedi, British Pound and Euro, all against the U. S. Dollars (from January 2, 2010 to December 31, 2014).

3.1.1 Jarque-Bera Statistics

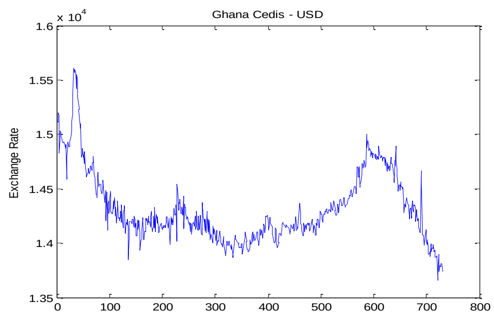
Jarque-Bera statistics is applied to examine the non-normality of the exchange rate series



Statistics	Naira/Dollar rate
Mean	36.28380
Std. Dev.	64.67169
Skewness	1.220108
Kurtosis	3.489442
Jarque-Bera	194.4878
Probability	0.000000

Figure 1. Naira/dollar exchange rate index summary statistics

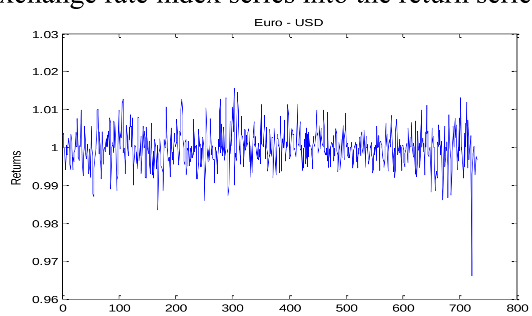
Figure 1 shows a positive skewness, 1.220108, and a high positive kurtosis, 3.489442. With reference to the Jarque-Bera statistics, Naira/dollar exchange rate index is non-normal at the confidence interval of 99%, since probability is 0.000000 which is less than 0.01. Consequently, there is need to convert the Naira/dollar exchange rate index series into the return series.



Statistics	Cedi/Dollar
Mean	0.345512
Std. Dev.	0.615862
Skewness	1.220410
Kurtosis	3.490659
Jarque-Bera	194.5413
Probability	0.000000

Figure 2. Cedi/dollar exchange rate index summary statistics

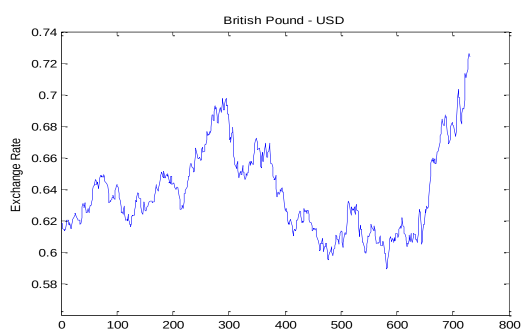
Figure 2 shows a positive skewness, 1.220410 as well as a positive kurtosis, 3.490659. As indicated by Jarque-Bera statistics, the Cedi/dollar exchange rate index is non-normal at the confidence interval of 99%, since probability is 0.0000 which is less than 0.01. So the need also arises to convert the Cedi/dollar exchange rate index series into the return series.



Statistics	Euro/Dollar
Mean	0.180556
Std. Dev.	0.322022
Skewness	1.224487
Kurtosis	3.506883
Jarque-Bera	195.2801
Probability	0.000000

Figure 3. Euro/dollar exchange rate index summary statistics

Figure 3 shows a positive skewness, 1.224487, and a positive kurtosis, 3.506883. As indicated by the Jarque-Bera statistics, Euro/dollar exchange rate index is non-normal at the confidence interval of 99%, since probability is 0.0000 which is less than 0.01; hence the need to convert the Euro/dollar exchange rate index series into the return series.



Statistics	Pound/Dollar
Mean	0.152918
Std. Dev.	0.272588
Skewness	1.220855
Kurtosis	2.492430
Jarque-Bera	194.6209
Probability	0.000000

Figure 4. Pound/dollar exchange rate index summary statistics

Figure 4 shows a positive skewness, 1.220855, and a positive kurtosis, 2.492430. As indicated by the Jarque-Bera statistics, Euro/dollar exchange rate index is non-normal at the confidence interval of 99%, since probability is 0.0000 which is less than 0.01; hence the need to convert the Euro/dollar exchange rate index series into the return series.

3.1.2 Transformation of the exchange rate index series of the Nigerian Naira,Ghana Cedi, British Pound and Euro

On the whole, the movements of the stock indices series are non-stationary, and therefore, not suitable for the study purpose. The stock indices series are transformed into their returns so that we get stationary series. The transformation is:

$$r_t = 100 \ln \frac{p_t}{p_{t-1}} \qquad (11)$$

where r_t , p_t is the exchange rate at time index t, p_{t-1} the exchange rate just prior to the time t.

3.1.3 Augmented Dickey-Fuller (ADF) Test and Phillips-Perron (PP) Test on Naira/Dollar,Cedi/Dollar, Pound/Dollar and Euro/Dollar exchange rates index Returns Series

Both the ADF and PP tests are used to obtain verification regarding whether Naira/Dollar, Cedi/Dollar, Pound/Dollar and Euro/Dollar exchange rates return series is stationary or not.

Table 1. ADF test on Naira/Dollar, Cedi/Dollar, Pound/Dollar and Euro/Dollar exchange rate returns

		t-Statistic			
		Naira/Dollar index	Cedi/Dollar index	Pound/Dollar index	Euro/Dollar index
ADF test statistic		-43.12567	-45.56412	-47.34789	-46.78622
	1% level	-3.331562	-3.33253	-3.331562	-3.33253
	5% level	-2.751341	-2.751341	-2.751341	-2.751341
Test critical values	10% level	-2.456200	-2.456200	-2.456200	-2.456200
Prob.		0.0001	0.0001	0.0001	0.0001

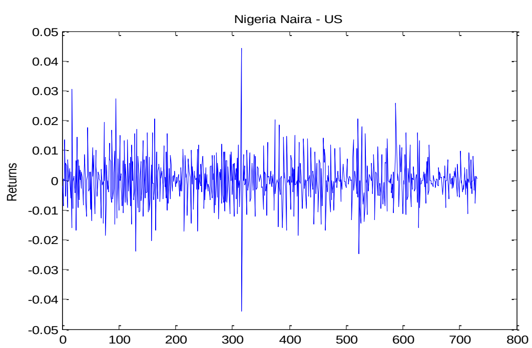
Table 1 shows that the values of ADF test statistic, -43.12567, is less than its test critical value, -2.751341, at 5%, level of significance which implies that the Naira/Dollar exchange rates return series is stationary. The result of ADF test also demonstrates that the Cedi/Dollar, Pound/Dollar and Euro/Dollar return series are stationary, as the values of ADF test statistic is less than its test critical value.

Table 2. PP test on Naira/Dollar, Cedi/Dollar, Pound/Dollar and Euro/Dollar exchange rates returns

		t-Statistic			
		Naira/Dollar index	Cedi/Dollar index	Pound/Dollar index	Euro/Dollar index
PP test statistic		-43.32035	-45.80403	-47.34789	-46.78622
	1% level	-3.331562	-3.33253	-3.331562	-3.33253
	5% level	-2.751341	-2.751341	-2.751341	-2.751341
Test critical values	10% level	-2.456200	-2.456200	-2.456200	-2.456200
Prob.		0.0001	0.0001	0.0001	0.0001

Table 2 illustrates the results of the PP test and proves that the Naira/Dollar index returns series is stationary, as the values of PP test statistic, -43.32035, is less than its test critical value, -2.751341, at the level of significance of 5%. The outcome of the PP test equally shows that the Cedi/Dollar, Pound/Dollar and Euro/Dollar exchange rates returns series are stationary, since the values of PP test statistic is less than its test critical value.

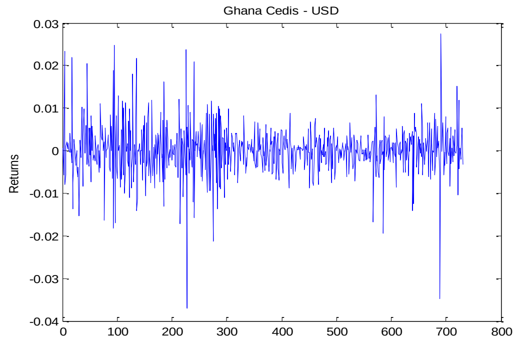
3.2 Summary Statistics of the Naira/Dollar, Cedi/Dollar, Pound/Dollar and Euro/Dollar exchange rates returns



Statistics	Naira/Dollar rate
Mean	-0.001385
Std. Dev.	0.708650
Skewness	-0.074139
Kurtosis	8.805879
Jarque-Bera	735.0376
Probability	0.000000

Figure 5. Naira/dollar exchange rate index returns summary statistics

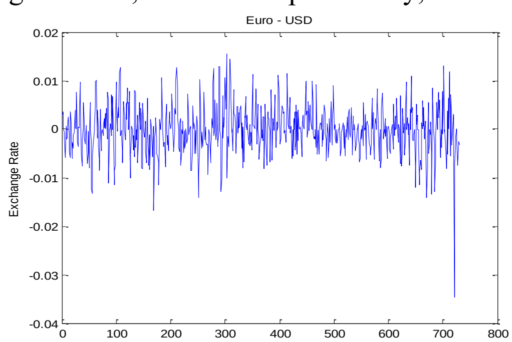
Figure 5 reveals a negative skewness, -0.074139, and a positive kurtosis, 8.805879. As indicated by the Jarque-Bera statistics, the Naira/dollar exchange rate index returns series is non-normal at 95% confidence level, since probability is 0.0000 which is less than 0.05.



Statistics	Cedi/Dollar
Mean	0.006258
Std. Dev.	0.536507
Skewness	-0.096923
Kurtosis	13.11769
Jarque-Bera	2312.255
Probability	0.000000

Figure 6. Cedi/dollar exchange rate index returns summary statistics

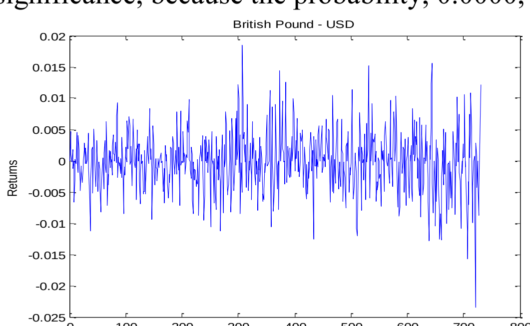
Figure 6 also reveals a negative skewness, -0.096923, and a positive kurtosis, 13.11769. Based on the Jarque-Bera statistics, the Cedi/dollar exchange rate index returns series is non-normal at 5% level of significance, because the probability, 0.0000, is less than 0.05.



Statistics	Euro/Dollar
Mean	0.001676
Std. Dev.	0.488392
Skewness	-0.434943
Kurtosis	7.993814
Jarque-Bera	559.9343
Probability	0.000000

Figure 7. Euro/dollar exchange rate index returns summary statistics

Figure 7 also reveals a negative skewness, -0.434943, and a positive kurtosis, 7.993814. Based on the Jarque-Bera statistics, the Euro/dollar exchange rate index returns series is non-normal at 5% level of significance, because the probability, 0.0000, is less than 0.05.



Statistics	Pound/Dollar
Mean	0.020803
Std. Dev.	0.506541
Skewness	-0.022958
Kurtosis	4.262290
Jarque-Bera	34.76827
Probability	0.000000

Figure 8. Pound/dollar exchange rate index summary statistics

Figure 8 also reveals a negative skewness, -0.022958, and a positive kurtosis, 4.262290. Based on the Jarque-Bera statistics, the Pound/dollar exchange rate index returns series is non-normal at 5% level of significance, because the probability, 0.0000, is less than 0.05.

4 Empirical Results and Evaluation

As the actual volatility at time t is not observable, there is need for some measures of volatility to assess the forecasting performance. In this paper we apply the standard approach suggested by Pagan and Schwert, (1990). A proxy for the actual volatility $\hat{\sigma}_t^2$ is given by

$$\hat{\sigma}_t^2 = (r_t - \bar{r})^2 \tag{12}$$

where $\bar{r}$ is the mean of the time series over the sample period. The statistical performance measures Mean Squared Error (MSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE), are applied to select the best performing model both in the in-sample and the out-of-sample data set independently in this study:

$$\text{MSE} = \frac{\sum_{t=1}^{n}(\hat{\sigma}_t^2 - \sigma_t^{\,2})^2}{n} \tag{13}$$

$$\text{MAE} = \frac{1}{n}\sum_{t=1}^{n}|\,\hat{\sigma}_t^2 - \sigma_t^{\,2}\,| \tag{14}$$

$$\text{MAPE} = \frac{\sum_{t=1}^{n}|\,\hat{\sigma}_t^2 - \sigma_t^{\,2}\,|/\sigma_t^{\,2}}{n} \tag{15}$$

where $\hat{\sigma}^2$ is the forecasted variance and σ^2 the actual variance time period t and n is the number of forecasts.

4.1 Statistical Performance

The evaluation results are shown in Tables 3 and 4 below. A two-state HMM-SV model was used in our experiments. In both tables, t-v represents true value, HSV stands for HMM-SV model and SV stands for SV model. s_1 and s_2 designate the two states with low and high volatility levels, respectively. MSE_1, MAE_1 and MAPE_1 are the in-sample MSE, MAE and MAPE while MSE_2 MAE_2 and MAPE_2 are the out-of-sample MSE, MAE and MAPE.

Table 3. Statistical performance results for the simulated data set and the true parameter sets compared with those obtained from HMM-SV and SV models

Models	ϕ	τ	α	MSE_1	MAE_1	MAPE_1	MSE_2	MAE_2	MAPE_2
t-v S_1	0.41	0.7417	2.0461						
	0.62	1.1445	2.0854						
S_2									
SV	0.72	0.8416	2.2761	0.0402	0.1001	0.2325	0.1010	0.1776	0.2641
HMMSV S_1	0.57	1.2454	2.0445						
	0.75	1.3534	2.1034				0.0161	0.0562	0.1684
S_2				0.0321	0.0623	0.1224			

Table 4. .Statistical performance results for the stock return data sets and the parameter sets obtained from HMM-SV and SV models

Stock Exchange	Models	ϕ	τ	α	MSE_1	MAE_1	$MAPE_1$	MSE_2	MAE_2	$MAPE_2$
Naira/Dollar	SV	0.8485	4.0273	4.3205	0.3401	0.3211	0.2334	0.2401	0.2211	0.3334
	HMMSV S_1	0.7875	3.4771	6.9980						
	S_2	0.0685	1.2341	3.8746	0.1021	0.1743	0.2534	0.0021	0.0743	0.2424
Cedi/Dollar	SV	0.9869	4.1936	5.3824	0.1370	0.1716	0.2265	0.1360	0.1706	0.2255
	HMMSV S_1	0.8127	4.2368	4.7144						
	S_2	0.0712	3.1134	2.1345	0.0210	0.0595	0.1473	0.0110	0.0495	0.1464
Pound/Dollar	SV	0.9770	2.1311	0.7654	0.1783	0.0928	0.1922	0.1773	0.0718	0.1812
	HMMSV S_1	0.9050	1.3136	0.9883						
	S_2	0.0805	1.2136	0.8564	0.0516	0.0783	0.0452	0.0416	0.0783	0.1112
Euro/Dollar	SV	0.9754	2.3108	0.7627	0.1706	0.2601	0.5784	0.0502	0.1402	0.5138
	HMMSV S_1	0.8871	1.4605	1.2590						
	S_2	0.0762	1.3605	1.1590	0.0956	0.1943	0.3578	0.0144	0.0943	0.4548

The above results are indicative that that HMM-SV model capture the volatility structure changes processes between two different volatility regimes with different volatility persistence ϕ . Nonetheless, the SV model cannot capture such volatility structure changes and always show very high volatility persistence. Consequently, HMM-SV model offers better volatility forecasts as the MSE (MAE) of HMM-SV model is considerably smaller than the SV models for most of the time.

5. Conclusion

The volatility persistence of widely-used SV model is usually too high leading to poor volatility forecasts. The root for this excessive persistence seems to be the structure changes (e.g. shift of volatility levels) in the volatility processes, which the SV model cannot capture.
As we developed our HMM-SV model to allow for both different volatility states in time series and state specific SV model within each state, the empirical results for both artificial data and real financial data not only takes care of the structure changes (hence giving better volatility forecasts), but also helps to establish an proficient forecasting structure for volatility models.

Accordingly, the results for both in-sample and out-of-sample evaluation forecasting performance confirm that our model outperforms widely-used SV model, Hence, the results suggest that it is promising to deepen the study of volatility persistence, the hidden regime-switching mechanisms inclusive. On long run, this will improve volatility forecasts in future research.

References

Bhar R., & Hamori S. (2004) Hidden Markov Models: Application to Financial Economics. Dordrecht, kluwer Academic Publishers.

Bollerslev, T. (1986).Generalized Autoregressive Conditional Heteroskedasticity.Econometrics,31, 307-327.

Bollerslev, T., Engle, R. F. and Nelson, D. B. (1995). ARCH models, in R. F. Engle and D. L. Mc-Fadden (eds), Handbook of Econometrics, Vol. 4, North-Holland, New York, chapter 49.

Chu. C.J., Santoni. G.J. and Liu. T. (1996) Stock market volatility and regime shifts in returns. Information Sciences, 77, 179 -190.

Diebold. F., (1986). Modeling the perststence of conditional v'artances: A comment. Econometric Reviews 5, 51 - 56.

Doucet A., and Johansen A. M. (2009). A tutorial on particle filtering and smoothing: Fifteen years later. In Oxford Handbook of Nonlinear Filtering, D. Crisan and B. Rozovsky (eds.). Oxford University Press.

Engle, R., (1982): Autorregressive Conditional Heteroskedasticity with estimates of United Kingdom Inflation, Econometrica 50 (4): 987-1008.

Figlewski, S. (1997). Forecasting Volatility. Financial Markets, Institutions and Instruments, 6, 1-88.

Glosten. L.R., Jagannathan. R.and Runkle. D.E. (1993) On the relation between the expected value of and the volatility of nominal excess return on stocks. Journal of Finance, 53, 1779 -1801.

Hamilton. J.D. and Susmel. R. (1994) Autoregressive conditional heteroscedasticity and changes in regime. Journal of Econometrics, 64, 307 -333.
Jarque, C., and Bera, A. (1987). A Test for Normality of Observations and Regression Residuals, International Statistical Review, 55: 163–172.

Kim. C.J, Nelson. C.R. and Startz. R (1998) Testing for mean reversion in heteroscedastic data based on Gibbs sampling augmented randomisation. Journal of Empirical Finance, 5, 131-157

Kim, J. and 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.

Lawrence R. Rabiner and Fellow (1989) A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition. Proceedings of the IEEE, 77, 257 -286.

Lamoureux, C., (1990). Persistence in Variance, Structural Change, and the GARCH Model. J. Bus Econ. Statist., 8:2, 225-234.

Lamoureux. C. and Lastrapes, W. D. (1990). Persistence in variance, structural change and the GARCH model, Journal of Business and Economic Statisttcs, 8. 225 - 234.

MacKay, R. J. (2003). Hidden Markov Models: Multiple Processes and Model Selection. Ph.D Thesis - Dept. of Statistics, The University of British Columbia

Nkemnole E. B., Abass, O., and Kasumu, R. K (2013): "Parameter Estimation of a Class of Hidden Markov Model with Diagnostics" Journal of Modern Applied Statistical Methods 12 (1): 181-197

Nkemnole E. B., Abass,O.,and Kasumu, R. K (2015):"A t-distribution Based Particle Filter for Univariate and Multivariate stochastic volatility models" Journal of the Nigerian Mathematical Society 34:227-242.

Nelson . D.B. (1991): Conditional heteroskedasticity in asset returns: a new approach. Econometrica, 85, 347 - 370.

Panait, I., & Slavescu, F. O. (2012). Using GARCH-in-mean model to investigate volatility and persistence at different frequencies for Bucharest Stock Exchange during 1997 - 2012. Theoretical and Applied Economics, 19(5), 55-76.

Perron, P. (1989). The great crash, the oil price shock, and the unit root hypothesis, Econometrica 57, 1361-1401.

Rabiner, L. R. (1989). A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition, in Proceedings of the IEEE, 77 (2): 257-286.

Shumway, R. H. and Stoffer, D. S. (2006). Time Series Analysis and its Applications: Springer, New York.
Pagan, A. R., and G.W. Schwert, (1990): "Alternative Models for Conditional Stock Volatility.
J. Econometrics, 45, 267-290.

Shephard, N. (1996). Statistical aspects of ARCH and stochastic volatility. In Time Series Models: In Econometrics, Finance and Other Fields (eds D. R. Cox, D. V. Hinkley and O. E. Barndorff- Nielsen). London: Chapman and Hall, 1–67.

So, M. K. P., Lam, K. and Li, W. K. (1999) Forecasting exchange rate volatility using autoregressive random variance model. Applied Financial Economics, 9, 583–91.

Taylor, S. J. (1986). Modeling Financial Time Series. Wiley, Chichester.

Taylor, S. J. (1982). Financial returns modelled by the product of two stochastic Processes–a study of daily sugar prices, 1961–79. In Time Series Analysis: theory and Practice,1, (ed. O. D. Anderson). New York: Elsevier Science Publishing Co., Amsterdam: North-Holland Publishing Co., pp. 203–26.

Xiong-Fei Zhuang, and Lai-Wan Chan (2004).Volatility Forecasts in Financial Time Series with HMM-GARCH Models. IDEAL, volume 3177 of Lecture Notes in Computer Science, 807-812.

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

Zhang, D, Ning X., Liu, X., and Hongwei M. (2007). Prediction in Hidden Markov Models Using Sequential Monte Carlo Methods, In Proceedings of IEEE International Conference on Grey Systems and Intelligent Services, 718-722.