

Discrete-Time Markov's Chain for a Multivariate Stochastic Volatility

Ngassima Fanny Laure¹, Joseph K. Mung'atu², MbeleBidima Martin³

1 Pan African University Institute of Basic Sciences, Technology and Innovation (Pauisti) Nairobi, Kenya

2 Department of Mathematics (Finance) - Jomo Kenyatta University of Agriculture and Technology (Jkuat) Nairobi, Kenya

Abstract

Forecasting volatility is one of the fundamental areas of research in Financial Mathematics, and thus has been the focus of many researchers; also, financial markets are known to be far from deterministic but stochastic and hence random models tend to perfectly model the markets. This study used appropriate Discrete-time Markov models to predict the multivariate stochastic Autoregressive volatility of an equity portfolio on a stock market. Therefore, the idea of modelling volatility as a stochastic process for an accurate forecast using the Markov chain on the financial data sets are based on the risks that often affect investment opportunities and the risk factors for prices changing that investors are most concerned about making decisions. The results provided more accuracy on forecasting price volatility on stock markets. We used a 3-state Discrete-Time Markov Chain (DTMC) for a portfolio of two stocks for the same sector and we compared the used model (fitted on a portfolio) to the multivariate GARCH models using real data from a stock market. The modified model provided better volatility smiles compared to the Multivariate Generalized Autoregressive Conditional Heteroscedasticity (MGARCH) models.

1. Introduction

Changes in volatility over time can be modelled using the approach based on Discrete time Markov chain (DTMC). The main characteristic of any financial asset is its return, which is typically considered to be a random variable. The asset's volatility that describes the spread of outcomes of this variable, plays the principal role in numerous financial applications. We often use it to estimate the value of market risk and we will use it in this work for portfolio management. Then, the main purpose of our research is to allow financial institutions not only to know the current value of the volatility of the managed assets, but also to be able to estimate their future values. However, the generalizations to multivariate series can be difficult to estimate and interpret. Another approach is to model volatility as an unobserved stochastic process. Although it is not easy to obtain the exact likelihood function for such stochastic volatility models, they tie in closely with developments in finance theory and have certain statistical attractions. A number of papers have documented the advantage of modelling stochastic volatility including Harvey et Al (1994) who used the Quasi Maximum Likelihood (QML) methods. Although there have been already many practical and successful applications of multivariate GARCH models, the theoretical literature on multivariate stochastic volatility (MSV) models has developed significantly over the last few years. Nevertheless, compared to the MGARCH literature, the literature on MSV is much more limited, reflected by much fewer published papers on the topic to date. Yet the MSV models remain more difficult to estimate, although estimation is already an issue for the MGARCH models, it is believed that estimation is more of an issue for MSV models. Moreover, as a result of difficulties with parameter estimation, the computation of model comparison criteria becomes extensive and demanding. Also, compared to the multitude alternative specifications in MGARCH models, only a handful of MSV model specifications have been studied; this may be among the multiple reasons why the MSV models have had fewer empirical applications.

Our interests in Stochastic Volatility models stem from their popularity in analysis of macroeconomic and financial market data. It has been shown by Boscher et Al (2000) and Hol and Koopman (2002) that in some empirical studies Stochastic Volatility models make better forecasts than GARCH models do. In GARCH-type models the conditional variance of returns is assumed to be a deterministic function of past returns, whereas in stochastic volatility (SV) models the volatility process is random. There are both economic and econometric reasons why multivariate volatility models are important. One of the advantages of multivariate stochastic volatility (MSV) models over GARCH models is parsimony. The knowledge of correlation structures is very important in many financial applications, such as asset pricing, optimal portfolio, risk management and asset allocation, so that multivariate volatility models are useful for making financial decisions. Two classes of models ARCH and Stochastic Volatility have emerged as the dominant approaches for modelling financial volatility. Volatility should be modelled as a stochastic process. Several real situations can be modelled by stochastic processes including our case, time series are stochastic processes that illustrate the daily closing values of the stock market. One of the main objectives of the study of time series is therefore, the forecasting of future realizations very often for economic reasons, namely to predict the evolution of a financial market.

We use discrete time Markov chain (DTMC) methods to develop a new class of stochastic volatility models in which volatility has a discrete support. The method developed in our work for estimation of the MSV model is simulated maximum likelihood (SML). The model will be easy to estimate. It implies that volatility follows a low-order Autoregressive (AR) process.

Our model has a low dimension of the state space, and is parameterized such that volatility follows a low-order Autoregressive (AR) process, and does not incorporate any type of underlying component structure; see (Calvet and Fisher, 2004).

In SML estimation, the latent variable is simulated conditional on available information and the simulated value is used to construct an unbiased estimate of the marginal density of the observable variable. The SML method has been used by Danielsson (1994a) and Adriana \$ Kirby (2014) in estimating a univariate SV model and the technique is extended here to allow for estimation of the MSV model. Simulated likelihood (SL) has several advantages in estimation of stochastic volatility models, since it is a likelihood method; the classical theory of maximum likelihood (ML) carries over to the simulated likelihood (SL) case. Likelihood methods are more efficient than approximation techniques like quasi maximum likelihood (QML) (Danielsson, 1994a.)

The data used in this study is daily Equity Group and KCB Group Ltd prices data from 2010-2016.

2. Research Method

2.1. Modelling Volatility

In this paper, we developed a discrete multivariate stochastic autoregressive volatility model. Stochastic modelling is a form of financial modelling that includes one or more random variables. Among the stochastic models, one kind of process called Markov process is a specific type of a mathematical object known as a stochastic or random process, this has been studied by several independent researchers. Two related modelling strategies are typically followed in specifying the dynamics of the volatilities; the volatilities can be assumed to be a non-linear function of past returns, as shown in the ARCH type models also, the volatility process is a function of an exogenous shock as well as past volatilities as shown below. The stochastic volatility model can easily be parsimoniously extended to include multiple assets. The standard SV model is defined as

$$r_t = v_t \varepsilon_t \quad t = 1, \dots, T, \quad \varepsilon_t \sim N(0, 1)$$

where r_t is the return for the interval $t-1$ to t , $v_t > 0$ is the return volatility for period t , and ε_t is a white-noise error that is independent of v_{t-j} for all $j \geq 0$.

A discrete time stochastic process is a family of random variables $\{v_t, t \in N\}$ defined on a given probability space and indexed by the parameter $t \in N = \{0, 1, 2, \dots\}$. We can assume that v_{t+1} follows a first-order Markov process, i.e.

$$\begin{aligned} & \Pr(v_{t+1} = \sigma_k / v_1 = \sigma_h, \dots, v_{t-1} = \sigma_i, v_t = \sigma_j) \\ &= \Pr(v_{t+1} = \sigma_k / v_t = \sigma_j) \quad \forall t \end{aligned}$$

where σ is the standard deviation, and also that the transition probabilities for this process are time invariant, that is:

$$\begin{aligned} & \Pr(v_{t+n+1} = \sigma_k / v_{t+n} = \sigma_i) \\ &= \Pr(v_{t+1} = \sigma_k / v_t = \sigma_i) \quad \forall n, t \end{aligned}$$

That means we model volatility as a time-homogeneous, first-order Markov Chain with $m=3$ states. Volatility is here, associated with the sample standard deviation of returns over some period of time. We can

compute it using the following formula: $\hat{\sigma} = \sqrt{\frac{1}{T-1} \sum_{t=1}^T (r_t - \mu)^2}$. As mentioned before, r_t is the return

of an asset over period t and μ is an average return over T periods. We could also use the variance, σ^2 , as a measure of volatility but variance and standard deviation are already connected by a simple relationship.

From that, suppose $S = (S_1, \dots, S_k)'$ denotes a vector of log-prices for k financial assets, and

$r_t = (r_{1t}, \dots, r_{kt})'$ denotes a vector of the observed log-returns for k financial assets at time t for $t = 1, \dots, T$. We assume that the conditional mean of r is zero for expositional purposes. Note that the conventional first-order Markov Chain model for k financial data sets of m states has m^k states.

Let $\epsilon_t = (\epsilon_{1t}, \dots, \epsilon_{kt})$, $V_t = (v_{1t}, \dots, v_{kt})$. A Multivariate model of returns is then defined as:

$$r_t = V_t \epsilon_t \quad [2.1]$$

where $V_t' V_t = \Omega_t$ is the $k \times k$ volatility matrix of r_t , and ϵ_t is a $k \times 1$ vector of White-noise error which are independent of V_{t-j} for all $j \geq 0$, and k the number of assets. The model of returns has to focus both the distribution of the shocks ϵ_t and the functional form of the volatilities Ω_t . In order to illustrate the key elements of our strategy, we assume that the dynamics of volatility is governed by a first-order Markov chain properly parameterized.

Let's assume that Ω_{t+1} follows a first-order Markov process, i.e.

$$\begin{aligned} & P(V_{t+1} = \sigma_k / V_1 = \sigma_h, \dots, V_{t-1} = \sigma_i, V_t = \sigma_j) \\ &= P(V_{t+1} = \sigma_k / V_t = \sigma_j), \quad \forall t \end{aligned}$$

In fact, we assume here that the volatilities that will be effective in a period depends only on the current volatilities for a period.

Commonly the return shocks $\boldsymbol{\varepsilon}_t$ are assumed to be normal and \boldsymbol{r}_t is conditionally normal, while the unconditional distribution of \boldsymbol{r} is non-normal, and can exhibit the expected stylized facts about returns, such as fat tails and volatility clusters. The variance of the shock returns is not constant over time or the volatility is clustering.

In the SV case, the volatilities are a dynamic latent variable and estimation is nontrivial since the volatilities have to be integrated out of the joint density for returns and volatilities. The stochastic volatility specification has several advantages over the GARCH class of models e.g. they are much more closely integrated with microeconomic theory (Anderson, 1994).

The expression of the volatility is:

$$\mathbf{V}_{t+1} = \boldsymbol{\sigma}' \mathbf{x}_{t+1} \quad [2.2]$$

and

$$\mathbf{x}_{t+1} = \mathbf{P} \mathbf{x}_t + \mathbf{e}_{t+1} \quad [2.3]$$

where \mathbf{x}_t is the state of today and $\boldsymbol{\sigma}'$ an $k \times 1$ vector of σ_m that specifies the volatilities mass points and where each $\sigma_m = (\sigma_1, \sigma_2, \dots, \sigma_M)'$.

We can represent an M-states Markov chain in terms of a $M \times 1$ vector \mathbf{x}_t whose each j^{th} element equals 1 if the process is in state $j \in \{1, 2, \dots, M\}$ at time t and 0 otherwise; see Hamilton (1994). We have the state-transitions described by a VAR (1) process as below:

$$\mathbf{x}_{t+1} = \mathbf{P} \mathbf{x}_t + \mathbf{e}_{t+1}$$

where \mathbf{P} is a $M \times M$ transition matrix with $\mathbf{P}_{\sigma_k \sigma_j} = P(\mathbf{V}_{t+1} = \sigma_k / \mathbf{V}_t = \sigma_j) \quad \forall k, j \in M$ and \mathbf{e}_{t+1} is a vector martingale difference sequence, i.e., $E(\mathbf{e}_{t+1} / \mathbf{x}_1, \dots, \mathbf{x}_{t-1}, \mathbf{x}_t) = 0$.

Our model allows for leverage effects and time varying correlation. Hence, it is more flexible than those others models, the estimated model should be the same as the Multivariate model of Danielson (1998) and Harvey et Al (1994).

Assume that in the equation of the multivariate model of returns, the return shocks $\boldsymbol{\varepsilon}_t$ are multivariate normal. From the definition of the $k \times k$ matrix of volatilities $\boldsymbol{\Omega}_t$, where the covariance matrix, $\boldsymbol{\phi}_t$ is defined by:

$$\boldsymbol{\phi}_t = \boldsymbol{\Omega}_t' \boldsymbol{\Gamma} \boldsymbol{\Omega}_t$$

where $\boldsymbol{\Gamma}$ is the matrix of correlation coefficient defined by $E[\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t']$ and:

- $-1 < \Gamma_{ij} < 1 \quad \forall i \neq j$
- $\Gamma_{ij} = 1 \quad \forall i = j$

The covariance matrix will always be positive definite since $\boldsymbol{\Gamma}$ is positive definite. In the MSV model, variances and correlations are instantaneous stochastic variables.

In the discrete state-space framework, the autocorrelation function of volatility is determined by how we parameterize the transition matrix of the Markov chain. The most transparent way to obtain a multivariate first order autoregressive model is to specify P such that it is immediately apparent. We have:

$$\Omega_{t+1} = \zeta + \varphi \Omega_t + \eta_{t+1} \quad [2.4]$$

This is an AR (1) model of volatility, where φ is the first order autocorrelation coefficient of volatility and,

$\eta_{t+1} = \sigma' e_{t+1}$ is white noise. We can set

$$P = \varphi I_M + (1 - \varphi) 1_M \pi' \quad [2.5]$$

where $\varphi \in [0, 1)$, I_M denotes an $M \times M$ identity matrix, 1_M is the $M \times 1$ vector of 1's and π is the $M \times 1$ vector of ergodic probability with $\pi = E(X_t)$, with X_t denoting the state.

By substitution and simplification, we obtain the volatility process in equation [2.4] with $\omega = (1 - \varphi) \sigma' \pi$. This is a straightforward approach for formulating multivariate first order discrete stochastic volatility model.

Therefore, we will refer equations [2.2], [2.3] and [2.5], as a DMSARV (1, M) model, this designator conveys its two most important features. Volatility follows a discrete AR (1) process, and allows for M different realizations of volatility. We then assume that the distribution of volatility has discrete support for achieving the computational tractability of our approach.

The choice of N and M controls the degree of approximation error, if we wanted to approximate a continuous stochastic AR volatility process in which the marginal distribution of volatility is log normal, we could parameterize π and σ by specifying the mean and variance of the log normal distribution. Formulating a multivariate higher-order of the previous DMSARV model will require some modifications to the methods developed above and this is an interesting avenue for future research.

2.2. Parameterization Strategies

Since M is small (i.e. ≤ 3), the multivariate stochastic volatility specifications developed below are slightly parameterized because each π_k and σ_k have M elements for $k = 1, 2$; we will impose two additional restrictions on the parameter space for more specifications. Let parameterize π_k and σ_k as:

$$\sigma_{jk} = \delta_k + \gamma_k j, \quad k = 1, 2 \text{ and } j = 1, 2, \dots, M \quad [2.6]$$

where each $\gamma > 0$ and $\delta > -\gamma$

Volatility mass points are evenly spaced along a line. Now, π_k can be parameterized as

$$\pi_{jk} = \frac{(M-1)!}{(j-1)!(M-j)!} \omega_k^{j-1} (1 - \omega_k)^{M-j} \quad [2.7]$$

$j = 1, \dots, M$ with $M = 3$ and $k = 1, 2$, where $\omega \in (0, 1)$.

By imposing these two restrictions, we obtain a class of models that have only 8 parameters regardless the size of the state space.

By extending this approach to a log linear specification, this yields

$$\log \sigma_{jk} = \delta_k + \gamma_k j \quad j=1, \dots, M \quad [2.8]$$

where $\gamma > 0$ and the value of δ is unrestricted, one in which the mass points of the log of volatility are evenly spaced along a line. Changing how to parameterize σ has no effect on the basic time series of volatility, Ω_t still follow a discrete AR (1) process.

Therefore, knowing that parameterization strategies of σ is an important factor for the approximation of the marginal distribution of volatility, however, the evidence from the realized-volatility literature suggested that the marginal distribution of volatility is much closer to log normal than to normal. Then, we use the parameterization which is more in line with log-normality and expect it to better fit the data.

Another parameterization of σ that offers greater flexibility in the positioning of the volatility mass points is done by replacing the linear functions in the previous equations with polynomials of any order less than M , That is

$$\log \sigma_{jk} = \delta_k + \gamma_k j + \beta_k j^2 \quad [2.9]$$

where $\gamma > 0$ and the values of δ and β are unrestricted, allowing the mass points of log-volatility to take on a quadratic configuration, provided that $M = 3$.

The normal distribution has the fourth moment equal to 3, although some papers have shown that the distribution of market returns have sample fourth moments larger than 3. Also, prices movements are negatively correlated with volatility, this means that the volatility of shock tends to increase when the stocks prices fall, decrease when the stock prices rise and null when the stock prices are stable, since the stock market prices are highly fluctuating.

The simplest way to estimate volatility is taking daily squared returns, Unfortunately, this method gives an inaccurate estimation of volatility (Taylor, 1986); we then calculate daily returns using the closing price of each asset in the end of a trading session.

2.3. Models with Asymmetric Volatility

There is an asymmetric relation between stock prices changes and the volatility of future stock returns as shown previously. Therefore, the source of this asymmetry has been explained in the literature; the common explanations are known as the leverage hypothesis and the volatility feedback hypothesis.

It is known that the leverage hypothesis asserts that a fall in the stock market price leads to an increase in financial leverage, which makes the stock a riskier investment and that can create a decreasing need to invest, and causes its volatility to increase, while the volatility feedback hypothesis asserts that the risk premium demanded by investors increases whenever they expect volatility to increase, and this increase in the risk premium immediately causes a decrease of the stock prices.

In that case, in order to capture these effects, we may allow the transition probability for the volatility process to be variant over the time.

Let consider the following model

$$\Omega_{t+1} = \sigma' x_{t+1}$$

$$x_{t+1} = P_t' x_t + e_{t+1}$$

with P_t a time varying transition matrix, denoted by

$$P_t = \phi I_M + (1 - \phi) 1_M \pi_t'$$

Here π_t is no more $E(x_t)$ but $f(r_{it}, r_{it-1}, \dots, r_{i1})$ with $i = 1, \dots, k$. The time-varying transition probability still requires some parameterization decisions and since it's well-known that the volatility follows a discrete AR (1) process, and that P_t depends on exogenous variables and predetermined. The transition probabilities for Ω_{t+1} are a function of only the lagged returns which are those predetermined variables.

Let,

$$\Omega_{t+1} = \omega_t + \phi \Omega_t + \eta_{t+1}$$

This model is described by a discrete AR (1) process with T-varying intercept, where $\omega_t = (1 - \phi) \sigma' \pi_t$. This process can capture asymmetric volatility effects because it allows the expected value of Ω_{t+1} knowing Ω_t to be correlated with r_t and $r_{t-1}, r_{t-2}, \dots, r_1$. The correlation between returns and volatility like that implied by the leverage and volatility-feedback hypothesis, is generated by having negative returns in periods t and earlier to be associated with changes in π_t that increase the value of ω_t . Another parameterization of π_t is

$$\pi_{jt}^k = \frac{(M-1)!}{(j-1)!(M-j)!} (\omega_t^k)^{j-1} (1 - \omega_t^k)^{M-j} \quad j = 1, \dots, M$$

Let's specify a binomial inspired parameterization for π_t

$$\pi_{jt} = \frac{(M-1)!}{(j-1)!(M-j)!} \omega_t^{j-1} (1 - \omega_t)^{M-j} \quad j = 1, \dots, M$$

where the time varying parameter ω_t is:

$$\omega_t = \frac{\exp(\eta + \psi r_t + \psi \rho r_{t-1} + \psi \rho^2 r_{t-2} + \dots + \psi \rho^{t-1} r_1)}{1 - \exp(\eta + \psi r_t + \psi \rho r_{t-1} + \psi \rho^2 r_{t-2} + \dots + \psi \rho^{t-1} r_1)}$$

with $\rho \in [0, 1)$

In this case, the sign of ψ controls the strength and direction of the asymmetric volatility effect. If we set $\psi < 0$, that means, it gives us a model in which negative returns are associated with increases in expected future volatility. On the other side, ρ controls the rate at which this asymmetric volatility response diminishes with time. If at time t , a negative return tells us that the volatility is expected to increase in the

future, then this expected increase could be entirely transitory if $\rho = 0$, if $\rho = 0.5$ then it could be moderately persistent and highly persistent of and only if it is close to one i.e. $\rho = 0.9$

2.4. Model with time varying volatility persistence

We would like to formulate a model that displays time-varying volatility persistence, allowing for time-varying transition probabilities, we therefore allow the transition probabilities matrix for the volatility process to vary over the time by assuming it is selected in a stochastic manner for each t . If we suppose that $\{Y_t\}_{t=1}^{\infty}$ is a stochastic process with discrete support such that $Y_t \in \{1, 2\}$ for all t and $y_t = (y_{1t}, y_{2t})'$ a $k \times 1$ vector. If we let Y_t be generated by a time-homogenous ergodic and irreducible 3-state Markov chain, then we can express the transition probabilities for Y_{t+1} as

$$\begin{aligned} \Pr(Y_{t+1} = j / Y_t) \\ = (1 - \phi) \left((1 - \bar{\omega})(2 - j)(3 - j) + \bar{\omega}(j - 1)(j - 2) \right) + \phi 1_{[b_{t=j}]} \end{aligned}$$

where $\phi \in [0, 1)$ and $\bar{\omega} \in (0, 1)$

Let $X_t^{(y)}$ denotes a 3×1 vector whose j^{th} element equals one if $Y_t = j$ and zero otherwise.

To obtain a general multivariate stochastic volatility model for volatility that displays time-varying volatility persistence, we assume that the joint transition probabilities of Ω_{t+1} and Y_{t+1} are given by

$$\begin{aligned} \Pr(\Omega_{t+1} = \sigma_k, Y_{t+1} = j / \Omega_t, Y_t) \\ = \left((1 - \phi_j) \pi_k + \phi_j 1_{[\Omega_t = \sigma_k]} \right) \Pr(Y_{t+1} = j / Y_t) \end{aligned}$$

with $\phi_j \in [0, 1]$ for $j \in \{1, 2, 3\}$, and $\phi_1 > \phi_2 > \phi_3$.

2.5. Estimation Method

We used the Simulated Maximum Likelihood method introduced by Danielson and Richard (1993) which depends on Monte Carlo integration, in order to evaluate the likelihood. The likelihood function of multivariate stochastic volatility models involves high-dimensional integration, which is difficult to calculate numerically. Nevertheless, estimation of the parameters can be based on evaluating high-dimensional integrals with simulation methods and then maximizing the likelihood function, resulting in the so-called SML estimators. There are several ways to perform SML estimation for multivariate stochastic volatility models, the most usual approach to SML is the importance sampling method. The basic idea of this method is to approximate first the integrand by a multivariate normal distribution using the so-called Laplace approximation and then draw samples from this multivariate normal distribution.

Let $f(\mathbf{R}_{t+1} / \Omega_{t+1}, I_t; \boldsymbol{\theta})$ be the joint probability density function of \mathbf{R}_{t+1} conditional on observing both Ω_{t+1} and $I_t = \{\mathbf{R}_t, \mathbf{R}_{t-1}, \dots, \mathbf{R}_1\}$, with $\boldsymbol{\theta}$ a vector of unknown parameters which is estimated by maximum likelihood. In order to fit our model in equation [3.2], [3.3], and [3.5], let's assume that

$$\mathbf{R}_{t+1} / \Omega_{t+1}, I_t \sim N(0, \sigma_{t+1}^2) \text{ where } \Omega_{t+1} = \mathbf{V}_{t+1}' \mathbf{V}_{t+1}$$

$$\text{And } \mathbf{V}_{t+1} = \boldsymbol{\sigma}' \mathbf{x}_{t+1}$$

$$\mathbf{x}_{t+1} = (\phi \mathbf{I}_N + (1 - \phi) \mathbf{1}_N \boldsymbol{\pi}')' \mathbf{x}_t + \mathbf{e}_{t+1}$$

$\boldsymbol{\theta} = (\delta, \gamma)'$ since the normal distribution is determined by its mean and variance, in the case where $\boldsymbol{\sigma}$ is parameterized as $\sigma_j = \delta + \gamma j$, $j = 1, \dots, M$.

Now, let $\mathbf{x}_{t+1/t} = E(\mathbf{x}_{t+1} / I_t)$ denotes the expectation of the $M \times 1$ vector \mathbf{x}_{t+1} given the period t information set. Hamilton (1989) show that $\mathbf{x}_{t+1/t}$ is given by

$$\mathbf{x}_{t+1/t} = \mathbf{P}' \left(\frac{\mathbf{x}_{t/t-1} \square \boldsymbol{\eta}_t}{\mathbf{1}'_N (\mathbf{x}_{t/t-1} \square \boldsymbol{\eta}_t)} \right)$$

where $\boldsymbol{\eta}_t = (\eta_{1t}, \dots, \eta_{Mt})'$ is a $M \times 1$ vector with j^{th} element,

$$\eta_{jt} = f(\mathbf{R}_t / \Omega_t = \boldsymbol{\sigma}_j, I_{t-1}; \boldsymbol{\theta})$$

Then we can write the log likelihood function as

$$L(\boldsymbol{\theta}) = \sum_{t=1}^T \log \mathbf{1}'_N (\mathbf{x}_{t/t-1} \square \boldsymbol{\eta}_t)$$

where $\boldsymbol{\theta}$ contains both parameters that determine the transition probabilities and those contained in $\boldsymbol{\theta}$, with \mathbf{P} parameterized as in equation [3.5] and $\boldsymbol{\pi}$ is as in equation [3.7]. We then use a quasi-Newton method to find the value of $\boldsymbol{\theta}$ that maximizes $L(\boldsymbol{\theta})$ and we compute standard errors using the second-derivative estimate of the information matrix. In order to select the model that fit better the data, we measure the performance of the out-of-sample variance forecasts produced by various models, and we require a proxy for the unobserved variance of daily returns.

For example, if we want to evaluate one-step ahead forecasts, we might fit a regression of the following form

$$\mathbf{R}\Omega_{t+1} = a + b\hat{\sigma}_{t+1/t}^2 + \varepsilon_{t+1}$$

where $\mathbf{R}\Omega_{t+1}$ is the realized variance joint variance for period $t+1$ and $\hat{\sigma}_{t+1/t}^2$ is constructed using maximum likelihood estimates of the model parameters, and models are rank using the regression R-squared. To conduct formal comparisons of the various models under study, we used either AIC or BIC tests.

3. Data Analysis and Results

In this section, we assess the empirical performance of the proposed DMSARV models by fitting a number of different specifications to a portfolio of daily data on stock indices. We begin with a description of the data, where the general statistical features of the NSE data are investigated and the rest of the sections discuss the application of benchmark models together with DMSARV models in real life data.

3.1. The data

The data sequences are generated by the same source. Daily closing prices of NSE Equity and KCB shares data over a period of 7 years extending from 01/01/2010 to 31/12/2016 with 1756 observations were used. The Equity and KCB shares are the most traded and most profitable companies trading in NSE market. They track the daily performance of the most capitalized companies in the sector of Banking among the eight (08) segments listed on the NSE. In order to make forecasts, the full sample was divided into two parts, in sample and out-of-sample observations.

3.1.1. Assets returns

Most financial studies involve returns instead of prices of assets to forecast volatility. This is because the return of an asset is a complete and scale-free summary of the investment opportunity for average and aware investors, and returns series are easier to handle than price series because return series have more attractive statistical properties. (Giot and Laurent, 2001). We used the daily percentage returns for the stock indices namely Equity and KCB stocks, in order to fit the discrete MSARV models.

Let P_t and P_{t-1} denote the closing asset prices of NSE assets at the current (t) and previous (t-1) day respectively. The rate of returns on an asset price is defined as

$$r_t = \log \left(\frac{P_t}{P_{t-1}} \right)$$

3.1.2. Summary statistics of NSE returns series data

In order to describe the behaviour of NSE return series, we drawn descriptive statistics table for the returns. The data are in log-difference form. The skewness, kurtosis, Kolmogorov test for normality, and correlation coefficients are used as the diagnostic tools under this study. They are defined as in table 1.

This is implemented by using the estimated mean, μ and the standard deviation, σ . The null hypothesis of normality is rejected if the p-valued of the Kolmogorov statistic is less than the significance level.

The summary of the descriptive statistics for the NSE returns series are shown in table 2. As it is expected for a time series of returns the mean is close to zero. The return series are both negatively skewed, an indication that the NSE data used have symmetric returns. The kurtosis is greater than three for the normal distribution, this indicates that the underlying distribution of the returns are leptokurtic or heavy tailed. The series fail the Kolmogorov normality test statistic which rejects normality at the 1% confidence level in both cases; that means they have positive excess kurtosis which confirms that the returns are effectively leptokurtic or heavy tailed.

3.2. Benchmark Models

We used two multivariate GARCH models which have been revealed to fit better the data. The first is the CCC model introduced for the first time by Bollerslev, the conditional correlation matrix in this class of models is time invariant. We then choose a GARCH-type model for each conditional variance and we model the conditional correlation matrix, based on the conditional variances.

Since the conditional correlation matrix is time invariant, the conditional covariances are therefore proportional to the product of the corresponding conditional standard deviations. Hence,

Definition 3.1

The CCC (p, q) process is a martingale difference sequence X_t , relative to a given filtration F_t , whose conditional covariance matrix $H_t = \text{cov}(X_t / F_{t-1})$ satisfy

$$H_t = D_t R D_t = \rho_{ij} (\sigma_{iit} \sigma_{jjt}) \quad [3.1]$$

$$\text{where } D_t = \text{diag}(\sigma_{11t}, \dots, \sigma_{kkt}) \quad [3.2]$$

$$\text{and } R = (\rho_{ij}) \quad [3.3]$$

is a symmetric positive definite matrix with $\rho_{ii} = 1$, $\forall i$ then off diagonal elements of the conditional covariance matrix are defined as $[H_t]_{ij} = \sigma_{iit} \sigma_{jjt} \rho_{ij}$ for $i \neq j$, $1 \leq i, j \leq k$. σ_{iit}^2 is defined as univariate GARCH (p, q) model

$$\sigma_t^2 = \omega + \sum_{i=1}^p A_i X_{t-i}^2 + \sum_{i=1}^q B_i \sigma_{t-i}^2 \quad [3.4]$$

where ω is $k \times 1$ vector, A_i and B_i are diagonal $k \times k$ matrices. See Francq and Zakoian (2010) for more details. The DCC Model is a generalization of the CCC model was proposed by Engle (2002), the so-called DCC is a new class of multivariate models which conditional correlation matrix is time-dependent. These models are flexible like the previous univariate GARCH and parsimonious parametric models for the correlations.

Definition 3.2

The DCC process is a martingale difference sequence X_t , relative to a given filtration F_t , whose conditional covariance matrix $H_t = \text{cov}(X_t / F_{t-1})$ satisfy

$$H_t = D_t R_t D_t \quad [3.5]$$

where

$$D_t = \text{diag}(\sigma_{1t}, \dots, \sigma_{kt}) \quad [3.6]$$

and R_t is $k \times k$ time varying correlation matrix of X_t , σ_{iit}^2 is defined as univariate GARCH (p, q) model.

$$\sigma_{it}^2 = \omega_i + \sum_{j=1}^{p_i} \theta_{ij} X_{t-j}^2 + \sum_{j=1}^{q_i} \phi_{ij} \sigma_{t-j}^2$$

where ω_i , θ_{ij} , and ϕ_{ij} are non-negative parameters for $i = 1, \dots, k$, with the usual GARCH restriction for

non-negativity and stationary being imposed, such as non-negativity of variances and $\sum_{j=1}^{p_i} \theta_{ij} + \sum_{j=1}^{q_i} \phi_{ij} < 1$.

In a bivariate case, the number of parameters to be estimated equals $(k+1)(k+4)/2$. Note that H_t , being a covariance matrix has to be positive definite, D_t is positive definite since all the diagonal elements are positive, this ensure R_t to be positive definite. Also, all the elements in the correlation matrix R_t have to be equal or less than one by definition; See Engle (2002) for more details.

3.3. Empirical Results

We used the DCC and CCC models as benchmarks Multivariate models which have been revealed to fit better the data. To address the issue of model selection, we measured the performance of the in-sample and the out of sample variance forecasts produced by the various models. We required a proxy for the unobserved variance of daily returns. Fitting Mincer and Zarnowitz regressions is a common strategy for evaluating the forecasting performance of volatility models (Calvet and Fisher, 2004) and (Fleming and Kirby, 2013). If we are evaluating one-step-ahead forecasts for example, we might fit a regression of the form

$$RV_{t+1} = \beta_0 + \beta_1 \hat{\sigma}_{t+1}^2 + e_{t+1} \quad [3.7]$$

where RV_{t+1} is the realized variance for period $t+1$ variance based on the period t information set.

Unbiased forecasts correspond to the hypothesis $\beta_0 = 0$ and $\beta_1 = 1$. And $\hat{\sigma}_{t+1}^2$ is constructed using maximum likelihood estimates of the model parameters and models are ranked using the regression R-squared. We used the Diebold and Mariano test of equal predictive accuracy to conduct formal comparisons.

For example to compare model i and j under a specified loss function $L(\sigma_{t+1}^2, \hat{\sigma}_{t+1}^2)$. Our null

hypothesis is $E(e_{t+1}^{(ij)}) = 0$. Where,

$$e_{t+1}^{(ij)} = L(\sigma_{t+1}^2, \hat{\sigma}_{it+1}^2) - L(\sigma_{t+1}^2, \hat{\sigma}_{jt+1}^2) \quad [3.8]$$

denotes the loss differential for period $t+1$. To implement the test, we used the MSE loss function, we fitted to daily percentage returns, the regressions are estimated via OLS and the forecasts are for one-day horizon.

Let's start with the linear discrete SARV model. Results from estimation of DSARV (1, M) model of the data are presented in table 3. We fitted first-order discrete SARV models to daily percentage returns for the two stocks first individually with $T=1756$. All the specifications employ the linear parameterization of σ given by $\sigma_j = \delta + \gamma j$ with $j = 1, 2, \dots, M$. Table 3.a and 3.b report maximum likelihood parameter estimates

for $M=3$ and $M=10$. And table 3.c reports the in-sample and out-of-sample model selection criteria for all values of M from 3 to 10.

The former is the BIC obtained by fitting the model and the latter is the R-squared for a regression of daily realized variances on the variance forecasts produced by the fitted model. The parameter estimates for the $M=3$ display the expected characteristics, the estimates of δ and γ provide clear evidence of time-varying volatility for both stocks while the estimate of ω the ergodic probability of the high volatility state are all below 0.5. This implies that the process spends more time in the low-volatility state, and the estimates of ϕ is high for both stocks, this indicates a strong persistence in volatility. Increasing the volatility mass point to $M=3$ changes all the estimates and the BIC decreases monotonically with M in each case; this suggests that it is suitable to work with $M \geq 3$.

The question is whether first-order discrete MSARV models capture the dynamics of volatility.

From table 4, we see how changing the parameterization of σ affects the performance of the model. We fitted the linear parameterization in equation [2.6], the log-linear parameterization in equation [2.8] and the log quadratic in [2.9], then we notice that from $M \geq 3$ the BIC values are decreasing and the parameterization for log volatility values lower than those for the volatility itself for every $M \geq 3$, this is because there is gain in moving to log parameterization. BIC values match with higher R-squared values, and the R-squared values are increasing with M . Moreover, we notice that almost all the values of R-squared are closed to one, this suggests that the model fits the data well; for example with $M \geq 3$, the R-squared improves from linear to log quadratic parameterization. It is worth noting that, since the BIC values diminish quickly as the value of M increases, this means there is benefit of increasing the number of states. We can conclude that the parameterization σ impacts on the performance of the discrete multivariate SARV and shows the advantage of working with a portfolio of stocks instead of a single stock.

Finally, the model selection criteria suggest that there is more benefit fitting the log quadratic parameterization for the bivariate model, since for every number of state chosen from $M=3$ the BIC decreases.

To conduct suitable pairwise comparison for the selected multivariate models, we used Diebold and Mariano to evaluate the forecast accuracy of the models by comparing the out-of-sample forecasting performance of selected first-order discrete multivariate model to that of the two benchmarks models above. We considered the discrete MSARV (1, 10) model, our benchmarks are a DCC and CCC models which are nonlinear combinations of univariate GARCH models.

Therefore, table 5 reports the results of t-statistics for pairwise tests of equal predictive accuracy under MSE loss. We considered forecast horizons one day and one week; this is 5 trading days and the loss differentials for day $t+1$ are computed as follow:

$$e_{t+1,H}^{(ij)} = \left(\mathbf{R}\mathbf{V}_{t+1,H} - \hat{\sigma}_{it+1,H}^2 \right)^2 - \left(\mathbf{R}\mathbf{V}_{t+1,H} - \hat{\sigma}_{jt+1,H}^2 \right)^2$$

With $H \in \{1, 5\}$, the forecast horizon, $i = (1, 2)$ and $j \in \{1, 2\}$ indicates the benchmark models. The null hypothesis for the test is $E(e_{t+1,H}^{ij}) = 0$. The t-statistics are based on robust standard errors that are

constructed using Newey and West (1987) weight. The lag length for the weights is 10 for the one-day horizon and 20 for the 5-day horizon, and a negative or positive t-statistic indicates that the model produces a lower or higher loss on average than the benchmark model.

Test Statistics	
Skewness	$= \frac{1}{T} \sum_{t=1}^T \left(\frac{r_t - \mu}{\sigma^3} \right)^3$
Kurtosis	$= \frac{1}{T} \sum_{t=1}^T \left(\frac{r_t - \mu}{\sigma^4} \right)^4$

4. Conclusion

The empirical analysis highlights the promise of our approach. Ultimately, volatilities and correlations among market returns are widely used in asset pricing. Although researchers have built many multivariate models, the multivariate stochastic volatility models remain the least mentioned in the literature. The multivariate volatility of NSE returns has been modelled and forecasted for a period of 1/01/2010 to 31/12/2016 using different GARCH-type model and by building on well-established techniques for constructing Markov chains with a specified autocorrelation function, we developed a multivariate stochastic volatility model in which volatility follows a low-order autoregressive process; the model specifications assume that volatility has discrete support, and we used SML estimation to estimate the models. One of the main findings is that volatility forecasts produced by first order discrete MSARV models outperform those produced by multivariate GARCH-type models; these findings hold for the stock indices on the NSE. Therefore, we can conclude that, there is a number of interesting directions in which our analysis could be extended. One possibility is to investigate the performance of higher-order discrete MSARV models in assets returns context and this should be relatively straightforward.

Corr. Coefficient	$= \frac{\text{cov}(r_{it}, r_{jt})}{\sigma_{r_{it}} \sigma_{r_{jt}}}$
Kolmogorov (Reject H_0)	$= \max_{1 \leq t \leq T} \left(F(r_t) - \frac{t-1}{N}, \frac{t}{N} - F(r_t) \right)$ F is a cumulative distribution

Appendix A.

Table 1. Test Statistics

Table 2. Summary statistics of NSE return series

Statistics	Equity Group	KCB Group Ltd
Observations	1757	1757
Max	0.0946	0.0878
Min	-0.1022	-0.1121
Mean	0.00043	0.00022
Variance	0.00037	0.00031
SD	0.0192	0.0176
Skewness	-0.093	-0.402
Kurtosis	7.3388	7.541
P-normal	< 5%	< 5%
correlation coef	0.2409	
corr coef for squared returns	0.2803	

Table 3. Estimation results for linear discrete SARV (1, M) models

Table 3.a: *DSARV (1, 3) parameters estimates*

Parameters	Equity		KCB	
	Est.	Std. E	Est.	Std. E
ω	0.22	0.046	0.24	0.03
δ	-0.131	0.041	-0.392	0.066
γ	1.63	0.051	1.78	0.069
ϕ	0.91	0.014	0.922	0.015

Table 3.b: *DSARV (1, 10) parameters estimates*

Parameters	Equity		KCB	
	Est.	Std. E	Est.	Std. E
ω	0.12	0.012	0.23	0.015
δ	0.51	0.023	0.121	0.028
γ	0.56	0.015	0.53	0.018
ϕ	0.93	0.009	0.938	0.009

Table 3.c: BIC and R^2 of volatility forecasts

Model	Equity		KCB	
	BIC	R^2	BIC	R^2
DSARV (1, 3)	4921.67	0.442	5090.51	0.4331
DSARV (1, 4)	5002.58	0.5512	5025.16	0.5606
DSARV (1, 5)	4928.4	0.6713	4977.65	0.658
DSARV (1, 6)	5034.72	0.7374	5055.28	0.743
DSARV (1, 7)	4987.67	0.802	4989.12	0.8149
DSARV (1, 8)	4957.72	0.844	4956.42	0.8403
DSARV (1, 9)	4859.71	0.8487	5021.37	0.87
DSARV (1, 10)	4985.99	0.8695	5000.37	0.8907

Table 4. Model selection criteria for linear, log linear and log-quadratic discrete MSARV (1, M) models

Model	Linear		Log-linear		Log quadratic	
	BIC	R^2	BIC	R^2	BIC	R^2
DMSARV(1, 3)	10073.74	0.4078	10054.82	0.897	10016.83	0.971
DMSARV(1, 4)	9984.37	0.5818	9957.53	0.941	9915.89	0.985
DMSARV(1, 5)	9977.49	0.6806	9947	0.961	9897.388	0.989
DMSARV(1, 6)	9971.65	0.748	9935.94	0.973	9897.115	0.993
DMSARV(1, 7)	9969.22	0.8	9924.11	0.98	9897.065	0.995
DMSARV(1, 8)	9967.643	0.841	9907.74	0.9838	9850.042	0.997
DMSARV(1, 9)	9925.37	0.8685	9889.73	0.987	9824	0.997
DMSARV(1, 10)	9916.64	0.8933	9842.88	0.989	9816.065	0.997

Table 5. Pairwise Comparison-Diebold Mariano test for stock indices

Model	DCC		CCC	
	1-day	5-day	1-day	5-day
DMSARV(1,3)	-0.84	-0.38	-0.95	-0.58
DMSARV(1,10)	-1.80	-1.61	-1.90	-1.70

References

- Adrianas, C., & Kirby, C. (2014, June). DISCRETE STOCHASTIC AUTOREGRESSIVE VOLATILITY. *Journal of Banking and Finance*, 43(1), 160-178.
- Akaike, H. (1974). A new look at the Statistical model identification . *IEEE Transactions on Automatic Control*, 19(6), 716-1974.
- AL-Najjar, D. (2016). Modelling and Estimation of Volatility Using ARCH/GARCH Models in Jordan's Stock Market. *Asian Journal o Finance and Accounting*, 8, 152-167.
- Baba, Y., Engle, R. F., Kraft, D., & Kroner, K. (1990). Multivariate Simultaneous Generalized ARCH. *mimeo, UCSD*.
- Bauwens, L., Laurent, S., & Rombouts, J. (2006). Multivariate GARCH models: A survey. *Journal of Applied Econometrics*, 21, 79-109.
- Bollerslev, T. (1986). Generalized Autoregressive Conditional Heteroscedasticity . *Journal of Econometrics*, 31, 307-327.
- Bollerslev, T., Engle, R., & Wooldridge, J. (1988). A capital Asset Pricing model with time-varying covariances. *The Journal of Political Economy*, 96, 116-131.
- CALVET, & FISHER. (2004). How to Forecast Long-Run Volatility: Regime Swithing and the Estimation of Multifractal Processes. *Journal of Financial Econometrics*, 49-83.
- Chang, C.-Y., Xi-Yuan Q., & Sheng-Yuan J. (2011). Bayesian estimation and Comparison of MGARCH and MSV models via WinBUGS. *Journal of Statistical Computation and Simulation*, 82:1, 19-39.
- Comte, F., & Lieberman, O. (2003). Asymptotic theory for multivariate GARCH processes. *Journal of Multivariate Analysis*, 84, 61-84.
- Dickey D. and Fuller, W. (1981). Distribution of the estimators for autoregressive time series with a unit root. *Econometrica*, 49, 1057-1072.
- Dielbold, F., & R. Mariano. (1995). Comparing Predictive Accuracy. *Journal of Business and Economic Statistics*, 13, 253-265.
- Engle, R. F. (2002). Dynamic Conditional Correlation. A Simple Class of Multivariate Generalized Autoregressive conditional Heteroscedasticity models. *Journal of Business and Economic statistics*, 20, 339-350.
- Engle, R. F., & Kroner, K. (1995). Multivariate Simultaneous Generalized ARCH. *Econometric Theory*, 11, 122-150.
- Engle, R. F., & Sheppard, K. (2001). Theoretical and Empirical properties of Dynamic Conditional Correlation Multivariate GARCH. *NBER Working Papers 8554, National Bureau of Economic Research, Inc*.
- FLEMING, & KIRBY. (2013). Component-Driven Regime-Switching Volatility. *Journal of Financial Econometrics*, 263-301.
- Francq, C., & Zakoian, J. (2004). Maximum Likelihood estimation of pure GARCH-ARMA processes. *Bernoulli*, 10(4), 605-637.
- Francq, C., & Zakoian, J.-M. (2010). QML Estimation of a class of multivariate GARCH models without moment conditions on the observed process. *MPRA Paper No. 20779, Unpublished*.
- Hamilton, J. D. (1994). Time series Analysis. *Princeton University Press, Princeton New Jersey*.
- Harvey, Esther Ruiz, & Neil Shephard. (1994). Multivariate Stochastic variance Models. *Oxford University Press, The Review of Economic Studies*, 61, 247-264.
- Koopman, S. J., & Eugenie Hol U. (2002). The Stochastic Volatility in Mean Model. Empirical Evidence from International Stock Markets. *Journal of Applied econometrics*, 667-689.
- Laurent E., C., & ADLAI J., F. (2004). How to forecast Long-Run Volatility: Regime Switching and the Estimation of Multifractal Processes. *Journal of Financial Econometrics*, 49-83.

- Léon, N. K. (2015). Forecasting stock return Volatility: Evidence from the West African Regional stock market. *Journal of Applied Finance and Banking*, 5(6), 25-35.
- Lopes, M. H. (2004). *Modelling and Forecasting Stochastic Volatility*. Departement d'Economia i d'Historia Economica: International Doctorate in Economic analysis.
- Mincer, J., & V. Zarnowitz. (1969). The Evaluation of Economic Forecasts. In J. Mincer, *Economic Forecasts and Expectations: Analysis of Forecasting Behavior and Performance* (pp. 3-46). NBER.
- Nelson, D. B. (1991). Conditional Heteroscedasticity in assets returns: a new approach. *Econometrica*, 59, 347-370.
- Newey, W., & K., W. (1987). A simple, positive semi-definite, heteroscedasticity and autocorrelation consistent covariance matrix. *Econometrica* 55, 703-708.
- Norris, J. (1997). *Markov Chains*. Cambridge University and Probabilistic Mathematics.
- Perrelli, T. R. (2001). Introduction to ARCH and GARCH models . *University of Illinois. Department of Economics. Econ* 472.
- Poon, S.-H. (2005). A practical Guide to Forecasting Financia Market Volatility. *John Wiley and Sons, Ltd*.
- Sheppard, K. (revised 25 Oct. 2003). Multi-step estimation of Multivariate GARCH models. *Proceedings of the International ICSC symposium: Advanced Computing in Financial Markets, University of California at San Diego*.
- Simeyo, O., Otumba, D. E., & Nyabwanga, R. N. (2015, April). Application of Markov Chain to model and Forecast stock market trend. *International Journal of current research*, 7(04), 14712-14721.
- Suliman, Z. S., & Winker, P. (2012). Modelling Stock Market Volatility Using Univariate GARCH models: Evidence from Sudan and Egypt . *International Journal of Economics and Finance*, 4(8), 161-176.
- Taylor, & Francis. (2016). Robust Parametric Tests of Constant Conditional Correlation in MGARCH model. *Econometric Reviews*.
- Taylor, S. J. (2005). Asset Price Dynamics, Volatility and Prediction. *Princeton University Press*.
- Tsay, R. (2010). Analysis of Financial Times Series. *John Wiley and Sons, Inc., New York, United States of America*, 3rd edition.
- TSE, Y. K., & TSUI, K. (2002). A multivariate Generalized autoregressive conditional heteroscedasticity model with time-varying correlation. *Journal of Business and Economic statistics*, 20, 351-362.
- Zhang, D., & Zhang, X. (2009). *Study on Forecasting the Stock Market Trend Based on Stochastic Analysis Method* (Vol. 4). International Journal of Business and Management.