

A Markov Regime Switching Approach of Estimating Volatility Using Nigerian Stock Market

Yahaya Haruna Umar, Matthew Adeoye

Statistics Department, Faculty of Science, University of Abuja, Abuja, Nigeria

Email address:

yahagumau@gmail.com (Y. H. Umar), matthewadeoye9@gmail.com (M. Adeoye)

To cite this article:

Yahaya Haruna Umar, Matthew Adeoye. A Markov Regime Switching Approach of Estimating Volatility Using Nigerian Stock Market. *American Journal of Theoretical and Applied Statistics*. Vol. 9, No. 4, 2020, pp. 80-89. doi: 10.11648/j.ajtas.20200904.11

Received: January 26, 2020; **Accepted:** April 7, 2020; **Published:** May 28, 2020

Abstract: Understanding and forecasting the behavior of volatility in stock market has received significant attention among researchers and analysts in the last few decades due to its crucial roles in financial markets. Portfolios managers, option traders, and market makers are all interested in the possibility of forecasting, with a reasonable level of accuracy. This study examined the volatility on the Nigeria stock market by comparing two Markov regime switching Autoregressive (MS-AR) Models estimated at different lagged values using the Nigeria stock exchange monthly All Share Index data from 1988 to 2018 in the Central Bank of Nigeria (CBN) Statistical Bulletin. It was found that factors like financial crisis, information flow, trading volume, economical aspects and investor's behavior are the causes of volatility in the stock market. The results and forecasts obtained from the statistical analysis in this research showed that the stock market will experience a steady growth in 2020 and beyond. Also, the stock market is experiencing fluctuations in the price indices which show that over the years, investors have been exposed to some certain risks in the time past. We therefore recommended that researchers should focus more attention in developing robust statistical model that will reflect and continue to monitor future trends and realities.

Keywords: Markov Regime Switch, Stock Returns, Volatility Clustering, Financial Crisis

1. Introduction

The importance of understanding and forecasting the behavior of volatility in stock market has received significant attention among researchers and analysts in the last few decades due to its crucial roles in financial markets. Portfolios managers, option traders, and market makers are all interested in the possibility of forecasting, with a reasonable level of accuracy. Volatility prediction is a critical task in asset valuation and risk management for investors and financial intermediaries. The price of almost every derivative security is affected by swings in volatility.

A financial market often changes patterns over time. It can also exhibit dramatic changes due to unexpected events such as natural hazards and financial crisis. A widely accepted fact is that financial markets behave quite differently in different economic situations. Traders often adjust their portfolios according to market trend, which is defined as the long term tendency of a financial market to move in a certain direction. A financial market is traditionally classified into 3 categories: bearish, bullish and neutral. The first two terms describe

overall market gain and loss respectively. The term neutral market is used when no strong upward or downward trend is observed. Several benefits can be derived from generating accurate forecast of volatility. Volatility forecast is a key indicator in assessing the performance of the stock market in order for both indigenous and foreign speculators to make accurate speculations and decisions on investments.

The issue with volatility of stock market price refers to the fluctuations that may be observed in stock market prices over time. The major reason for the ups and downs in the stock market may be traced to macroeconomic instability. Since the stock market operate in a macroeconomic environment, it is therefore necessary that the environment must be an enabling one in order to realize its full potentials. The problem with forecasting the stock market price is that the return distribution can change considerably over time. Volatility is an extremely complex thing to forecast because of the inherent instability of the variable (variability of the random). Volatility forecast sometimes may be uncertain, since it is just a mere projection based on some econometric technique in most cases.

The Markov regime switching model is a popular approach, much on behalf of the way it takes the shifts in the time series behavior into account and the way it is able to capture more complex dynamic patterns. What separates the Markov regime switching model from other switching models is that the switching mechanism is controlled by an unobservable variable that follows a hidden Markov chain. By Markov properties, the current value of the variable depends only on its immediate past value. This means that a structure in the series may prevail for a random period of time, before being replaced by another structure when a switching takes place. This way, the Markov regime switching model is able to capture more complex dynamic patterns. [1] stated that a way to model non-linear dynamic patterns in time series such as breaks or asymmetry is through the use of Markov switching models. It has been found that financial time series exhibit some formalized facts which can advantageously be reproduced by a hidden Markov model. This has made the Markov regime switching model one of the most popular nonlinear time series models in the literature [2, 3, 4, 5]. Investors are skeptical especially when it comes to investment, as they require that sense of security that they will attain a return on their investment. Hence, the need of a framework where the potential investors has access to relevant information pertaining to the stock market, so that they are able to conduct their personal research and choose the optimal stocks in the market. According to [6], stock market indexes provide guidance concerning the performance of the overall stock market. Therefore, in this study, volatility assessment and forecast are carried out on All Share Index (ASI) of the Nigeria Stock Exchange (NSE) in order to examine the performance of the Nigeria Stock Market. The All Share Index (ASI) on the Nigeria Stock Exchange (NSE) is used as proxy for stock market prices in order to assess volatility by measuring the trends and thereby examining the performance of the Nigeria Stock Market. This research will examine this issue from the perspective of the Markov regime switching model. A number of studies have used such an approach in volatility forecasting and have found that particularly for inherent instability of the variable (variability of the random), the regime switching models do provide a good forecasting ability. The broad objective of this study is to assess the volatility performance of the Nigeria Stock Market in order to determine the basis for examining its forecast.

2. Data

The data used for this research was sourced from the Central Bank of Nigeria statistical bulletin [7].

3. Method

This section introduces and explains which statistical tool was deployed in analyzing data and upon what frame work was the conclusion based on. Statistical tools like time series will be deployed in analyzing the data set with the aid of

statistical package E-Views also the graphical aspect.

3.1. Nonlinear Time Series

Traditional time series analysis is based on assumptions of linearity and stationarity. However, there has been a growing interest in studying nonlinear and non-stationary time series model in many practical problems. The first and the simplest reason for this is that many real world problems do not satisfy the assumptions of linearity and/or stationarity. For example, the financial markets are one of the areas where there is a greater need to explain behaviors that are far from being even approximately linear. Therefore, the need for the further development of the theory and application for nonlinear models is essential. In general time series analysis, it is known that there are a large number of nonlinear features such as cycles, asymmetries, bursts, jumps, chaos, thresholds, heteroscedasticity and mixtures of these have to be taken into account. A problem arises directly from a suitable definition of nonlinear model because not every model is linear. This class clearly encompasses a large number of possible choices. Furthermore, forecasting the future values of an observed time series is an important phenomenon for many real world problems. It provides a good basis for production planning and technical decisions. Forecasting means extrapolating observations available up to time t to predict observations at future times. Forecasting methods are mainly classified into quantitative techniques, which are based on unscientific and mathematical and/or statistical models respectively. The quantitative techniques are more important than qualitative techniques for future planning.

3.2. Test of Nonlinearity

Given the wide range of nonlinear time series models available and the inherent flexibility of these models, the possibility of getting a spuriously good fit to any time series data set is very high. Therefore it is usually recommended to perform a test of linearity against nonlinearity before building a possibly complex nonlinear model.

A popular test for nonlinearity, the BDS test, which has been found to have power against a wide range of nonlinear time series models, will be used to determine whether a nonlinear model is suitable for the data. There are many other types of nonlinearity tests that are developed to test against specific nonlinear models.

BDS Test for Nonlinearity

The BDS test developed by [8] was originally designed to test for the null hypothesis of independent and identical distribution (iid) for the purpose of detecting non-random chaotic dynamics. However, many studies have shown that BDS test has power against a wide range of linear and nonlinear alternatives [9].

Linear vs. nonlinearity choice can be made partly on statistical grounds deciding whether a linear specification is sufficient to describe all of the most important features of the data at hand.

Thus, the null and alternative hypothesis of the BDS test for detecting nonlinearity is as follows;

H_0 : The series are linearly dependent

H_1 : The series are not linearly dependent.

The BDS test is based on an integral correlation of the series and is defined as follows;

$$BDS_{m,M}(r) = \sqrt{M} \frac{c_m(r) - c_1^r(r)}{\sigma_{m,M}(r)} \quad (1)$$

Where M is the surrounded points of the space with m dimension, r denotes the radius of the sphere centered on the X_i , C is the constant and $\sigma_{m,M}(r)$ is the standard deviation of $\sqrt{M}C_m(r) - C_1^r(r)$.

3.3. Selection Criteria

It is not all that possible to identify the parameter p of the Markov Switching Autoregressive (MS-AR) process by visualization. The most common model selection criteria are the AIC, BIC. Using the AIC and BIC criterion, we select the Markov Switching Autoregressive model with the lowest value of the AIC or BIC.

3.3.1. The Akaike Information Criteria (AIC)

The Akaike information criterion (AIC) was the first selection criterion to gain widespread acceptance. AIC was introduced in 1973 by [10] as an extension to the maximum likelihood principle. Akaike's seminal idea was to combine estimation and structural determination into a single procedure. Given a family of candidate models of various structures, each model is fit to the data through maximum likelihood. An AIC is computed based on each model fit. The fitted candidate model corresponding to the minimum value of AIC is then selected; one is hoping to identify the fitted model that is closest to the generating model. The AIC is defined by

$$AIC = -2 \log(\text{likelihood}) + 2(N) \quad (2)$$

3.3.2. Bayesian Information Criterion (BIC)

The Bayesian information criterion (BIC) also known as the Schwarz information criterion (SIC) is related to the Bayes factor and useful for model comparison in its own right.

The BIC of a model is defined as

$$BIC = -2 \ln(\text{likelihood}) + (k + k \ln(N)) \quad (3)$$

3.4. Markov Regime Switching Model

A basic time series model that will be used in describing the Markov Regime Switching model is the Autoregressive Process (AR) defined below:

Definition: Autoregressive Processes AR (p)

The idea behind the autoregressive models is to explain the present value of the series, X_t , by a function of the past values, $X_{t-1}, X_{t-2}, \dots, X_{t-p}$.

An autoregressive process of order p is written as

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots + \phi_p X_{t-p} + \varepsilon_t \quad (4)$$

where $\{\varepsilon_t\}$ is white noise, i.e., $\{\varepsilon_t\} \sim WN(0, \sigma^2)$ and ε_t is uncorrelated with

X_s for each $s < t$.

Suppose that the typical historical behavior could be described with a first-order autoregression,

$$y_t = c_1 + \phi y_{t-1} + \varepsilon_t, \quad (5)$$

with $\varepsilon_t \sim N(0, \sigma^2)$, which seemed to adequately describe the observed data for $t = 1, 2, \dots, t_0$.

Suppose that at date t_0 there was a significant change in the average level of the series, so that we would instead wish to describe the data according

$$y_t = c_2 + \phi y_{t-1} + \varepsilon_t, \quad (6)$$

For $t = t_0 + 1, t_0 + 2, \dots$

This fix of changing the value of the intercept from c_1 to c_2 might help the model to get back on track with better forecasts, but it is rather unsatisfactory as a probability law that could have generated the data. We surely would not want to maintain that the change from c_1 to c_2 at date t_0 was a deterministic event that anyone would have been able to predict with certainty looking ahead from date $t = 1$. Instead there must have been some imperfectly predictable forces that produced the change. Hence, rather than claim that expression (5) governed the data up to date t_0 and (6) after that date, what we must have in mind is that there is some larger model encompassing them both,

$$y_t = c_{s_t} + \phi y_{t-1} + \varepsilon_t, \quad (7)$$

where s_t is a random variable that, as a result of institutional changes, happened in our sample to assume the value $s_t = 1$ for $t = 1, 2, \dots, t_0$ and $s_t = 2$ for $t = t_0 + 1, t_0 + 2, \dots$. A complete description of the probability law governing the observed data would then require a probabilistic model of what caused the change from $s_t = 1$ to $s_t = 2$. The simplest such specification is that s_t is the realization of a two-state Markov chain with

$$\Pr(s_t = j | s_{t-1} = i, s_{t-2} = k, \dots, y_{t-1}, y_{t-2}, \dots) = \Pr(s_t = j | s_{t-1} = i) = p_{ij}. \quad (8)$$

Assuming that we do not observe s_t directly, but only infer its operation through the observed behavior of y_t , the parameters necessary to fully describe the probability law governing y_t are then the variance of the Gaussian innovation σ^2 , the autoregressive coefficient ϕ , the two intercepts c_1 and c_2 , and the two state transition probabilities, p_{11} and p_{22} .

The specification in (8) assumes that the probability of a change in regime depends on the past only through the value of the most recent regime, though, as noted below, nothing in the approach described below precludes looking at more general probabilistic specifications. But the simple time-invariant Markov chain (8) seems the natural starting point and is clearly preferable to acting as if the shift from c_1 to c_2 was a deterministic event. Permanence of the shift would be represented by $p_{22} = 1$, though the Markov formulation invites the more general possibility that $p_{22} < 1$. Certainly in the case of business cycles or financial crises, we know that the situation, though dramatic, is not permanent.

Furthermore, if the regime change reflects a fundamental change in monetary or fiscal policy, the prudent assumption would seem to be to allow the possibility for it to change back again, suggesting that $p_{11} < 1$ is often a more natural formulation for thinking about changes in regime than $p_{22} = 1$.

The probability of switching is captured in the matrix P known as a transition matrix.

$$P = \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix} \quad (9)$$

Where $p_{11} + p_{12} = 1$ and $p_{21} + p_{22} = 1$

Suppose that we observe y_t directly but can only make an inference about the value of s_t based on what we see happening with y_t . This inference will take the form of two probabilities

$$\xi_{jt} = \Pr(s_t = j | \Omega_t; \theta) \quad (10)$$

for $j = 1, 2$,

where these two probabilities sum to unity by construction. Here $\Omega_t = \{y_t, y_{t-1}, y_1, y_0\}$ denotes the set of observations obtained as of date t , and θ is a vector of population parameters, which for the above example would be $\theta = (\sigma, \phi, c_1, c_2, p_{11}, p_{22})'$, and which for now we presume to be known with certainty. The inference is performed iteratively for $t = 1, 2, \dots, T$ with step t accepting as input the values

$$\xi_{i,t-1} = \Pr(s_{t-1} = i | \Omega_{t-1}; \theta) \quad (11)$$

for $i = 1, 2$ and producing as output (10). The key magnitudes one needs in order to perform this iteration are the densities under the two regimes,

$$\eta_{jt} = f(y_t | s_t = j, \Omega_{t-1}; \theta) = \frac{1}{\sqrt{2\pi}\sigma} \exp \left[-\frac{(y_t - c_j - \phi y_{t-1})^2}{2\sigma^2} \right], \quad (12)$$

for $j = 1, 2$.

Specifically, given the input (11) we can calculate the conditional density of the t th observation from

$$f(y_t | \Omega_{t-1}; \theta) = \sum_{i=1}^2 \sum_{j=1}^2 p_{ij} \xi_{i,t-1} \eta_{jt} \quad (13)$$

and the desired output is then

$$\xi_{jt} = \frac{\sum_{i=1}^2 p_{ij} \xi_{i,t-1} \eta_{jt}}{f(y_t | \Omega_{t-1}; \theta)} \quad (14)$$

As a result of executing this iteration, we will have succeeded in evaluating the sample conditional log likelihood of the observed data

$$\log f(y_1, y_2, \dots, y_T | y_0; \theta) = \sum_{t=1}^T \log f(y_t | \Omega_{t-1}; \theta) \quad (15)$$

for the specified value of θ . An estimate of the value of θ can then be obtained by maximizing (15) by numerical optimization.

Several options are available for the value ξ_{i0} to use to start these iterations. If the

Markov chain is presumed to be ergodic, one can use the unconditional probabilities

$$\xi_{i0} = \Pr(s_0 = i) = \frac{1 - p_{jj}}{2 - p_{ii} - p_{jj}}. \quad (16)$$

Other alternatives are simply to set $\xi_{i0} = 1/2$ or estimate ξ_{i0} itself by maximum likelihood.

The calculations do not increase in complexity if we consider an $(r \times 1)$ vector of observations y_t whose density depends on N separate regimes. Let $\Omega_t = \{y_t, y_{t-1}, \dots, y_1\}$ be the observations through date t , P be an $(N \times N)$ matrix whose row j , column i element is the transition probability p_{ij} , η_t be an $(N \times 1)$ vector whose j th element $f(y_t | s_t = j, \Omega_{t-1}; \theta)$ is the density in regime j , and $\hat{\xi}_{t|t}$ an $(N \times 1)$ vector whose j th element is $\Pr(s_t = j | \Omega_t, \theta)$. Then (13) and (14) generalize to

$$f(y_t | \Omega_{t-1}; \theta) = 1' (P \hat{\xi}_{t-t|t-1} \odot \eta_t) \quad (17)$$

$$\hat{\xi}_{t|t} = \frac{P \hat{\xi}_{t-t|t-1} \odot \eta_t}{f(y_t | \Omega_{t-1}; \theta)} \quad (18)$$

where 1 denotes an $(N \times 1)$ vector all of whose elements are unity and \odot denotes element-by-element multiplication. There further is no requirement that the elements of η_t be Gaussian densities or even from the same family of densities. One is also often interested in forming an inference about what regime the economy was in at date t based on observations obtained through a later date T , denoted $\hat{\xi}_{t|T}$. These are referred to as “smoothed” probabilities. The calculations in (17) and (18) remain valid when the probabilities in P depend on lagged values of y_t or strictly exogenous explanatory variables. However, often there are relatively few transitions among regimes, making it difficult to estimate such parameters accurately, and most applications have assumed a time-invariant Markov chain. For the same reason, most applications assume only $N = 2$ or 3 different regimes.

Nturns out also to be a natural device for handling transition probabilities that are functions of observable variables. It is natural to want to test the null hypothesis that there are N regimes against the alternative of $N + 1$, for example, when $N = 1$, to test whether there are any changes in regime at all. Unfortunately, the likelihood ratio test of this hypothesis fails to satisfy the usual regularity conditions, because under the null hypothesis, some of the parameters of the model would be unidentified. For example, if there is really only one regime, the maximum likelihood estimate \hat{p}_{11} does not converge to a well-defined population magnitude, meaning that the likelihood ratio test does not have the usual χ^2 limiting distribution. To interpret a likelihood ratio statistic, one needs to rely on generic tests of the hypothesis that an N -regime model accurately describes the data, though these tests are not designed for optimal power against the specific alternative hypothesis of $N + 1$ regimes. Other alternatives are to use Bayesian methods to calculate the value of N implying the largest value for the marginal likelihood or to compare models on the basis of their ability to forecast.

A specification where the density depends on a finite number of previous regimes, $f(y_t | s_t, s_{t-1}, \dots, s_{t-m}, \Omega_{t-1}; \theta)$ can be recast in the above form by a suitable redefinition of regime. For example, if s_t follows a 2-state Markov chain

with transition probabilities $\Pr(s_t = j | s_{t-1} = i)$ and $m = 1$, one can define a new regime variable s_t^* such that $f(y_t | s_t^*, \Omega_{t-1}; \theta) = f(y_t | s_t, s_{t-1}, \dots, s_{t-m}, \Omega_{t-1}; \theta)$ as follows:

$$s_t^* = \begin{cases} 1 & \text{when } s_t = 1 \text{ and } s_{t-1} = 1 \\ 2 & \text{when } s_t = 2 \text{ and } s_{t-1} = 1 \\ 3 & \text{when } s_t = 1 \text{ and } s_{t-1} = 2 \\ 4 & \text{when } s_t = 2 \text{ and } s_{t-1} = 2 \end{cases} \quad (19)$$

Then s_t^* itself follows a 4-state Markov chain with transition matrix

$$P^* = \begin{bmatrix} p_{11} & 0 & p_{11} & 0 \\ p_{12} & 0 & p_{12} & 0 \\ 0 & p_{21} & 0 & p_{21} \\ 0 & p_{22} & 0 & p_{22} \end{bmatrix} \quad (20)$$

More problematic are cases in which the order of dependence m grows with the date of the observation t . Such a situation often arises in models whose recursive structure causes the density of y_t given Ω_{t-1} to depend on the entire history $y_{t-1}, y_{t-2}, \dots, y_1$ as is the case in ARMA, GARCH, or state-space models.

3.5. Transition Probabilities

A finite homogeneous Markov Chain with m -states has transition probabilities:

$$P_{ij} = \Pr\{X_n = j | X_{n-1} = i\} \quad (21)$$

where $(i, j = 1, \dots, m)$.

If $P_{ij} > 0$, then state E_i is said to communicate with state E_j such that a two way communication is possible, and if additionally $P_{ji} > 0$, then $E_i \leftrightarrow E_j$. Transition probabilities form an $m \times m$ array that can be combined into a transition matrix T . If $\sum_{j=1}^m p_{ij} = 1$, then T is said to be row stochastic.

For this study, a two state or regime process was assumed and the probabilities of the Markov process were represented in matrix form as follows:

$$\begin{bmatrix} P(S_t = 1) & P(S_t = 2) \\ P(S_{t-1} = 1) & P(S_{t-1} = 2) \end{bmatrix} \begin{pmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{pmatrix} \quad (22)$$

Where $p_{i1} + p_{i2} = 1$, for $i = 1, 2$ and $T = \begin{pmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{pmatrix}$

The specification in equation (22) assumes that the probability of a change in regime depends on the past only through the value of the most recent regime.

3.6. Expected Duration

We can determine the duration of each regime by using the diagonals elements of equation (9). These diagonal elements denote the transition probabilities of being in the same state both in the current and previous period. Using this information, we can find the average length of which a particular regime lasts on average.

Let D_j denote the number of periods the system is in state j . The probability to stay k periods in state j is given as

follows:

$$\Pr(D_j = k) = p_{jj}^{k-1}(1 - p_{jj}) \quad (23)$$

This implies the expected duration of state j is given by:

$$E(D_j) = \sum_{k=0}^{\infty} k \Pr(D_j = k) \quad (24)$$

$$E(D_j) = \frac{1}{1 - p_{jj}} \quad (25)$$

3.7. Forecasting

Forecasts for one period or several periods into the future with the parameters for a tentative model have been selected. The principal objective of developing a time series model for a variable is to generate post sample period forecast for the same variable. The ultimate test for any model is whether it is capable of predicting future events accurately or not. The forecast accuracy for the several Markov Switching Autoregressive models (MS-AR) were chosen based on the values of some error metrics, namely MSE, RMSE, MAE, MAPE.

However, in this study three different error metrics are considered for the evaluation of forecasting models. They are root mean square error (RMSE), mean absolute error (MAE), mean square error (MSE).

a) Root Mean Square Error (RMSE)

Root mean Square Error (RMSE) is square root of average of sum-squared errors and is given by following formula:

$$RMSE = \sqrt{\frac{1}{n} \sum_{j=1}^n (y_j - \hat{y}_j)^2} \quad (26)$$

There is one problem with RMSE and it is that they may be close to 0 if large positive and negative errors cancel out each other. RMSE gives high weight to the large errors and are generally useful where large errors are not of importance. RMSE are more sensitive than other metrics to the infrequent large errors as the squaring process gives large weight to very large errors [11].

b) Mean Absolute Error (MAE)

The problem of RMSE, canceling out of large positive and negative errors can be avoided by using Mean Absolute Errors. In average, MAE weights all the differences equally.

$$MAE = \frac{1}{n} \sum_{j=1}^n |y_j - \hat{y}_j| \quad (27)$$

MAE and RMSE can be used together to study the variation in the errors in a set of forecasts. RMSE are always larger or equal than MAE.

c) Mean Square Error (MSE)

The sum of the squared forecast errors for each of the observations divided by the number of observations. It is an alternative to the mean absolute deviation (MAE), except that more weight is placed on larger errors. While MSE is popular among statisticians, it is unreliable and difficult to interpret. [12] found no empirical support for the use of the MSE or RMSE in forecasting. Fortunately, better measures are available [13].

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (28)$$

4. Statistical Data Analysis and Discussion

This section presents the statistical analysis of our findings.

4.1. Time Series Graph of the Raw Data

Time series plots which display observations on the y-axis against equally spaced lag on the x-axis used to evaluate patterns and behaviors in data over time is displayed in Figure 1 below. The data used for this research was sourced from Central Bank of Nigeria Statistical Bulletin.

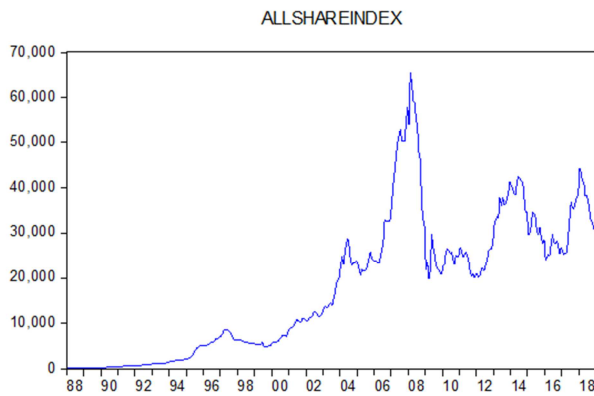


Figure 1. Time series graph of the All Share Index (ASI) on the Nigeria Stock Exchange (NSE): 1988 to 2018.

Volatility in the Nigeria stock exchange has been experiencing instability over the years as presented in Figure 1 above. Clearly, Figure 1 depicts trend in the NSE All Share Index series from 1988 to 2018. The series, the prices were below 5,000 from 1988 to around 1994; it went up and became stable within 5,000 and 10,000 around 1995 to 2001. Again it went beyond 10,000 in 2002 and kept going on an increase at a steady rate until in 2006 when it increased explosively up till the peak around 60,000, and declined explosively to 20,000 in 2008. Again it became stable when the prices were within 20,000 and 30,000 from 2008 until around 2013 when it started going on an increase beyond 30,000 up till 2017 where it went beyond 40,000 and dropped to around 30,000 in 2018.

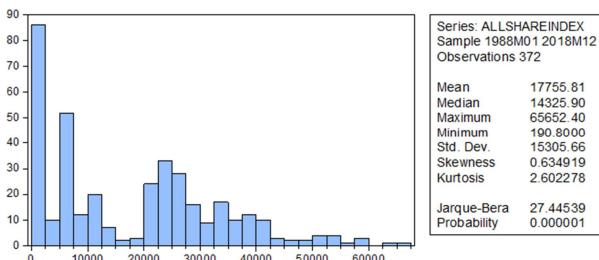


Figure 2. Histogram and Descriptive Statistics for All Share Index (ASI) Monthly Data: 1988 to 2018.

Before the estimation of the Markov switching models, a nonlinearity test might still be necessary to describe the important features of the data at hand. Table 1 below, reports the results of the nonlinearity test developed by [8] on the All Share Index (ASI) data from the Nigeria Stock Exchange (NSE).

4.2. BDS Test for Nonlinearity

Table 1. Parameter Estimates of BDS Test for Nonlinearity.

Dimension	BDS Statistics	Std. Error	P-value
2	0.186569	0.003026	0.0000
3	0.314866	0.004789	0.0000
4	0.401217	0.005676	0.0000
5	0.457852	0.005887	0.0000
6	0.495708	0.005650	0.0000

H_0 : The series are linearly dependent.

H_1 : The series are not linearly dependent.

The BDS test results in Table 1 above indicates that there is nonlinearity effect in the All Share Index of the Nigeria Stock Exchange. Table 1 shows that the probabilities are less than 5%, thus implying a rejection of the null hypothesis that the series is linearly dependent. The results in Table 1 suggest that All Share Index of the Nigeria Stock Exchange are nonlinear and unstable which is an indication of the dynamic behavior of financial time series data hence the data can be estimated using a nonlinear model.

Next, this study proceeds with the estimation of the Markov switching model (MS-AR). However, before obtaining the final form of the model used in this research, two models were estimated at different lagged values of the All Share Index on the Nigeria Stock Exchange.

MS (2)-AR (1)

$$y_t = \begin{cases} c_{s_t} + \phi_1 y_{t-1} + \varepsilon_t s_t = 1 \\ c_{s_t} + \phi_1 y_{t-1} + \varepsilon_t s_t = 2 \end{cases} \quad (29)$$

Where s_t denotes the regime/state

Table 2. Parameter Estimates For MS (2)-AR (1) Model.

Regime 1			
Variable	Coefficient	Std. Error	P-value
C	2594.267	490.0558	0.0000
LOG (SIGMA)	7.651628	0.046812	0.0000
AR (1)	0.988384	0.001973	0.0000
Regime 2			
Variable	Coefficient	Std. Error	P-value
C	2734.858	477.1843	0.0000
LOG (SIGMA)	3.980784	0.094791	0.0000
AR (1)	0.988384	0.001973	0.0000

Akaike's Information Criterion (AIC)= 16.09972

$$\hat{y}_t = \begin{cases} 2594.267 + 0.988384 y_{t-1} + \varepsilon_t & s_t = 1 \\ 2734.858 + 0.988384 y_{t-1} + \varepsilon_t s_t = 2 \end{cases} \quad (30)$$

From the estimated Markov switching model above, it can be seen that all the model parameters are significant at the 5% level since their P-value (0.0000) < 0.05, this is an

indication of a good model.

At the first regime, the estimated mean of the model is 2594.267 with a sigma of 20.7992 which are both significant with P-value of (0.0000) each, while at the second regime, the estimated mean of the model switched to 2734.858 with a sigma of 10.8209 which are also both significant at the 5% level of significance.

In common is the Autoregressive AR (1) parameter, $\hat{\phi}_1 = 0.988384$ with p-value (0.000) which is also significant at the 5% level of significance.

MS (2)-AR (2)

$$y_t = \begin{cases} c_{s_t} + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \varepsilon_t s_t = 1 \\ c_{s_t} + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \varepsilon_t s_t = 2 \end{cases} \quad (31)$$

Where s_t denotes the current regime

Table 3. Parameter Estimates For MS (2)-AR (2) Model.

Regime 1			
Variable	Coefficient	Std. Error	P-value
C	5500.129	395.5156	0.0000
LOG (SIGMA)	7.728381	0.053745	0.0000
AR (1)	1.306792	0.050581	0.0000
AR (2)	-0.299854	0.051198	0.0000
Regime 2			
Variable	Coefficient	Std. Error	P-value
C	165.6161	235.2329	0.4841
LOG (SIGMA)	5.498214	0.057957	0.0000
AR (1)	1.306792	0.050581	0.0000
AR (2)	-0.299854	0.051198	0.0000

Akaike's Information Criterion (AIC)= 16.20645

$$\hat{y}_t = \begin{cases} 5500.129 + 1.306792y_{t-1} - 0.299854y_{t-2} + \varepsilon_t s_t = 1 \\ 165.6161 + 1.306792y_{t-1} - 0.299854y_{t-2} + \varepsilon_t s_t = 2 \end{cases} \quad (32)$$

From the estimated Markov switching model with two lagged values above, it is evident that all the model parameters are significant at the 5% level since their P-value (0.0000) < 0.05 except the mean at the second regime with a P-value greater than 0.05.

At the first regime, the estimated mean of the model is 5500.129 with a sigma of 21.0079 which are both significant with P-value of (0.0000) each, while at the second regime, the estimated mean of the model switched to an insignificant mean of 165.6161 with a sigma of 14.9457 which is significant at the 5% level of significance.

In common is the Autoregressive AR (1) and AR (2) parameter, $\hat{\phi}_1 = 1.306792$, and $\hat{\phi}_2 = -0.299854$ for both regime 1 and regime 2 respectively with p-value (0.000) each, and this implies they are both significant at the 5% level of significance.

Table 4. Parameter Estimates of MS-AR Models and Models selection.

Model [MS-AR]	Parameter Estimate	P-value	S. E	Log Likelihood	AIC
MS (2)-AR (1)	$\hat{\phi}_1 = 0.988384$	0.0000	0.001973	-2987.548	16.09972
MS (2)-AR (2)	$\hat{\phi}_1 = 1.306792$ $\hat{\phi}_2 = -0.299854$	0.0000 0.0000	0.050581 0.051198	-3006.399	16.20645

Table 4 contained the summary results and the parameters estimate of the possible two state Markov switching

Autoregressive models. Comparing the Akaike Information Criteria (AIC) of the models, clearly we prefer the MS (2)-AR (1) model as the best since it has the smallest AIC (16.09972) and the largest log likelihood (-2987.548).

4.3. Inferences on the selected Markov Regime Switching Model

MS (2)-AR (1)

$$y_t = \begin{cases} c_{s_t} + \phi_1 y_{t-1} + \varepsilon_t s_t = 1 \\ c_{s_t} + \phi_1 y_{t-1} + \varepsilon_t s_t = 2 \end{cases} \quad (33)$$

Where s_t denotes the regime/state

Table 5. Parameter Estimates For the Selected MS (2)-AR (1) Model.

Regime 1			
Variable	Coefficient	Std. Error	P-value
C	2594.267	490.0558	0.0000
LOG (SIGMA)	7.651628	0.046812	0.0000
AR (1)	0.988384	0.001973	0.0000
Regime 2			
Variable	Coefficient	Std. Error	P-value
C	2734.858	477.1843	0.0000
LOG (SIGMA)	3.980784	0.094791	0.0000
AR (1)	0.988384	0.001973	0.0000

Akaike's Information Criterion (AIC)= 16.09972

Regime 1 (bull market):

$$\hat{y}_t = 2594.267 + 0.988384y_{t-1} + \varepsilon_t \quad (34)$$

Regime 2 (bear market):

$$\hat{y}_t = 2734.858 + 0.988384y_{t-1} + \varepsilon_t \quad (35)$$

From the estimated Markov switching model above, it is evident that all the model's parameters are significant at the 5% level since their P-value (0.0000) < (0.05).

Table 6 below contains the forecasts obtained for (January-December, 2018) using the MS (2)-AR (1) model in order to examine the forecasting performance of the model.

Table 6. Forecasting Results for MS (2)-AR (1) of All Share Index Data.

Month	Forecast	Actual
January 2018	37,903.41	44,343.65
February 2018	43,944.99	43,330.54
March 2018	42,941.66	41,504.51
April 2018	41,133.26	41,268.01
May 2018	40,899.05	38,104.54
June 2018	37,766.09	38,278.55
July 2018	37,938.43	37,017.78
August 2018	36,689.82	34,848.45
September 2018	34,541.43	32,766.37
October 2018	32,479.44	32,466.27
November 2018	32,182.87	30,874.17
December 2018	30,605.49	31,430.50

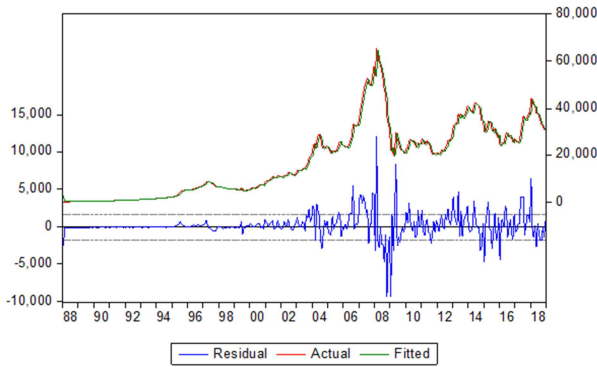


Figure 3. Residual – Actual – Fitted Plot for MS (2)-AR (1) Model: 1988 to 2018.

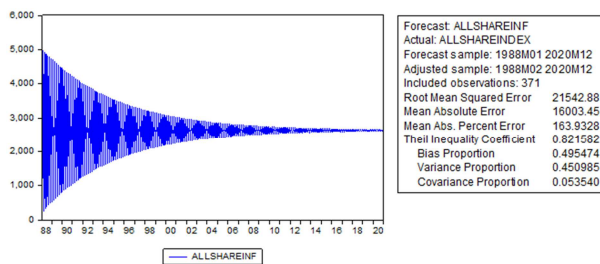


Figure 4. Forecasts Plot for MS (2)-AR (1) Model: 1988 to 2020.

4.4. Transition Probabilities and Expected Durations

Table 7. Estimates of Transition Probabilities and Expected Durations.

Constant transition probabilities:		
P (i, k) = P (s (t) = k s (t-1) = i)		
(row = i / column = j)		
	1	2
1	0.933390	0.066610
2	0.029268	0.970732
Constant expected durations:		
	1	2
	15.01287	34.16748

The probability transition matrix which is a tool for describing the behavior of Markov chains represents the probability of moving from state i to state j in a Markov process and is denoted as p_{ij} .

From the matrix above, given that we are in regime/state 1 (bull market), $p_{11} = 0.93$ represents the probability similar to a 93% chance of remaining in state/regime 1 and $p_{12} = 0.07$ represents the probability similar to a 7% chance of switching to regime/state 2. Also, given that we are in regime/state 2 (bear market) $p_{22} = 0.97$ represents the probability similar to a 97% chance of remaining in state/regime 2 (bear market) and $p_{21} = 0.03$ represents the probability similar to a 3% chance of switching to regime/state 1.

Furthermore, the estimated transition probabilities indicated that none of the regime is permanent since all the estimated transition probabilities are less than one.

We can also determine the duration of each regime by using the diagonal elements of the probability transition matrix as described in equation (25).

From the table of constant expected duration above, the

expected duration of being in regime 1 (bull market) in the stock market is approximately 15 months, while the expected duration of being in regime 2 (bear market) is approximately 34 months. This is because we are dealing with monthly All Share Index data from the Nigeria Stock Exchange.

4.5. Filtered and Smoothed Probabilities

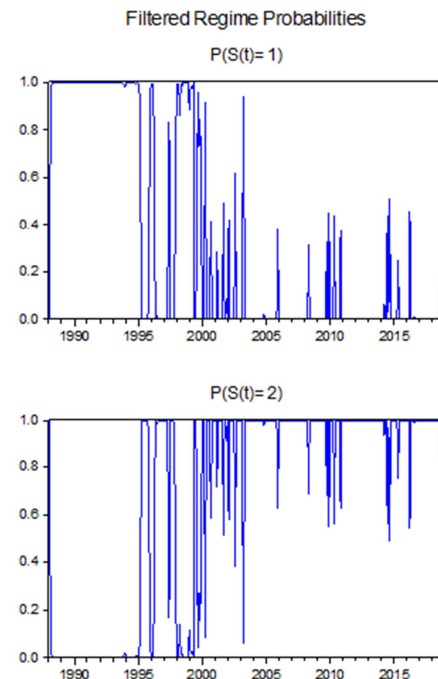


Figure 5. Filtered Regime Probabilities for All Share Index (ASI) Monthly Data.

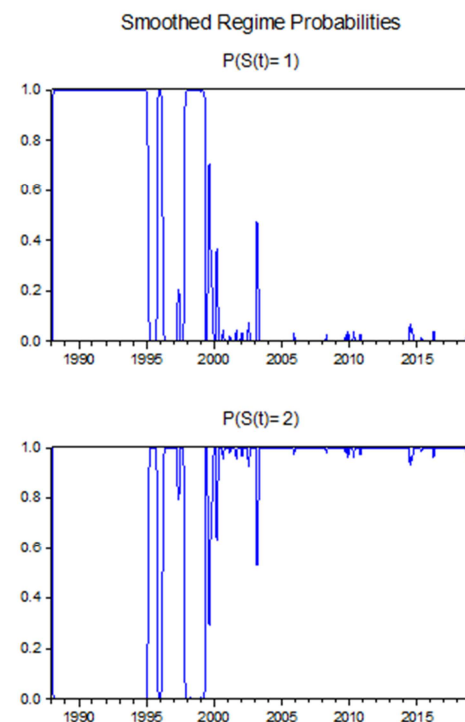


Figure 6. Smoothed Regime Probabilities for All Share Index (ASI) Monthly Data: 1988 to 2018.

The filtered and smoothed probabilities respectively, shows that the model accurately captures the two different regimes, because the probability of regime/state 1 is high when $s_t = 1$, and low when $s_t = 2$.

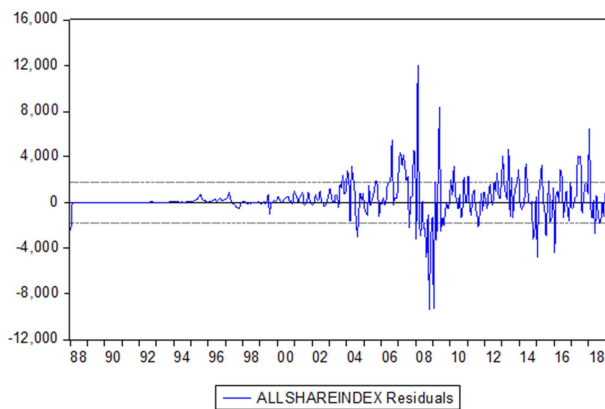


Figure 7. Residuals plot for MS (2)-AR (2) Model: 1988 to 2018.

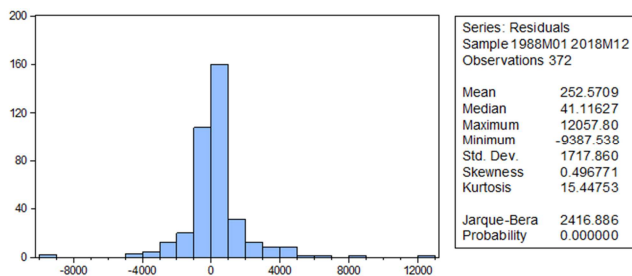


Figure 8. Histogram and Descriptive Statistics for Model Residuals: 1988 to 2018.

4.6. Diagnostic Tests for MS (2)-AR (1) Model

Heteroskedasticity Test

Table 8. Heteroskedasticity Test: ARCH.

Heteroskedasticity Test: ARCH			
F-statistic	0.097823	Prob. F (1,369)	0.7546
Obs*R-squared	0.098328	Prob. Chi-Square (1)	0.7538

H_0 : There is no ARCH effect in the Residuals.

H_1 : There is ARCH effect in the Residuals.

From the heteroskedasticity test carried out on the residuals of the estimated MS-AR model, we accept the null hypothesis at the 5% level of significance and conclude that the ARCH test reported homogeneity of variance across error term series.

Table 9. Correlogram of Residuals Q-Statistics.

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
. *	. *	1	0.084	0.084	2.6766	0.102
. .	. .	2	0.072	0.065	4.6176	0.099
. *	. *	3	0.109	0.099	9.1207	0.053
. .	. .	4	-0.031	-0.052	9.4785	0.058
. .	. .	5	-0.017	-0.025	9.5847	0.088
. .	. .	6	-0.027	-0.030	9.8605	0.131
. .	. .	7	0.049	0.066	10.764	0.149
. .	. .	8	0.044	0.043	11.498	0.175

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
. .	. .	9	0.066	0.058	13.143	0.156
. .	. .	10	-0.013	-0.044	13.203	0.213
. .	* .	11	-0.056	-0.069	14.396	0.212
. .	. .	12	0.033	0.038	14.807	0.252

H_0 : The error terms (Residuals) are independent.

H_1 : The error terms (Residuals) are not independent.

From the Q-Statistics tests above, it can be concluded that the error terms are not serially correlated since the P-values across the 12 lags are greater than 0.05, we therefore accept the null hypothesis and conclude that the residuals of the MS-AR models are independent. Therefore, the residuals of the MS (2)-AR (1) model has been diagnosed and examined to be white noise since they are not serially correlated, and have a constant variance.

5. Conclusion

This study examined the volatility on the Nigeria stock market by comparing two Markov regime switching Autoregressive (MS-AR) Models estimated at different lagged values using the Nigeria stock exchange monthly All Share Index data from 1988 to 2018. The preliminary analysis of the data obtained shows the presence of volatility in the stock market. The parameters of the MS-AR Models were estimated with most of the parameters significant at 5% level of significance. AIC and log-likelihood was used to select the best model that was used. From the AIC and log-likelihood, MS (2)-AR (1) was selected to be the best model since it has the smallest AIC and largest log-likelihood. The nonlinear time series model known as the Markov Regime Switching Autoregressive (MS-AR) model was then used to generate post sample forecasts for 2019 and 2020. The forecasts value indicates clearly that All Share Index prices will have a steady growth between 2019- 2020. From the trend analysis of the Nigerian All Share Index, the result shows that there is an unstable trend which indicates that Nigeria stock market is experiencing fluctuations in the price indices which shows that over the years, investors have been exposed to some certain risks in the stock market in the time past.

References

- [1] Benschop T. and Cabrera B. L., Volatility Modeling of CO2 EmissionAllowanceSpot Prices with Regime-Switching GARCH Models, 2014, pp1-11.
- [2] Cont, R. (2001). "Empirical properties of asset returns: stylized facts and statistical Issues", Quantitative Finance, 1, p. 223-236.
- [3] Hamilton, J. D. (1989). "A new approach to the economic analysis of nonstationary time series and the business cycle" Econometrica, 57, p. 357-384.
- [4] Lindgren, G. (1978). "Markov Regime Models for Mixed Distributions and Switching Regressions" Scandinavian Journal of Statistics, 5, p. 81-91.

- [5] Hamilton, J. D. (2005). "Regime-Switching Models" The New Palgrave Dictionary of Economics, Second Edition.
- [6] Bodie, Z., Kane, A., & Marcus, A. J. (1998). Essentials of Investment. New York: The McGraw-Hill Companies.
- [7] Central Bank of Nigeria statistical bulletin, (CBN, 2018).
- [8] Brock, W. A., Dechert, W. D., & Scheinkman, J. A. (1987). A Test for Independence Based on The Correlation Dimension. Department of Economics, University of Wisconsin at Madison, University of Houston, and University of Chicago.
- [9] Brock, W. A., Hsieh, D. A. and LeBaron, B. (1991) Nonlinear Dynamics, Chaos, and Instability: Statistical Theory and Economic Evidence. MIT Press, Cambridge, London.
- [10] Akaike, H. (1973), "Information theory and an extension of the maximum likelihood principle", in Petrov, B. N.; Csáki, F. (eds.), 2nd International Symposium on Information Theory, Tsahkadsor, Armenia, USSR, September 2-8, 1971, Budapest: the maximum likelihood principle", in Petrov, B. N.; Csáki, F. (eds.), 2nd International Symposium on Information Theory, Tsahkadsor, Armenia, USSR, September 2-8, 1971, Budapest: Decision 411 forecasting, 2010, What's the bottom line? How to compare models, Available at: file:///D:/Informatics/thesis/in%20KTM/error%20metrics/How%20to%20compare%20models.htm, [Accessed: 23 jan 2011].
- [11] Decision 411 forecasting, 2010, What's the bottom line? How to compare models, Available at: file:///D:/Informatics/thesis/in%20KTM/error%20metrics/How%20to%20compare%20models.htm, [Accessed: 23 jan 2011].
- [12] Armstrong, J. S., & Fildes, R. (1995). On the Selection of Error Measures for Comparisons Among Forecasting Methods. Retrieved from http://repository.upenn.edu/marketing_papers/63.
- [13] Armstrong, J. S. (2001), "Role-playing: A method to forecast decisions," in J. S. Armstrong (ed.), Principles of Forecasting. Norwell, MA: Kluwer Academic Press.