FORECASTING THE TAIL DENSITY OF NIGERIAN EXCHANGE RATES WITH A MIXTURE, AUTOREGRESSIVE MODEL

***Akinyemi, M. I.[a], Boshnakov, G. N.[b], Rufai, A.[c]**
[a,c] Department of Mathematics, University of Lagos, Akoka, Lagos, Nigeria
[b] School of Mathematics, University of Manchester, Manchester M13 9PL, UK
**Corresponding author: makinyemi@unilag.edu.ng*

ABSTRACT
Density forecasts have become more popular as real life scenarios require not only a forecast estimate but also the uncertainty associated with such a forecast. The class of mixture autoregressive (MAR) models provide a flexible way to model various features of financial time series and are also suitable for density forecasting. This study forecasted the out-of-sample tail density of Nigerian foreign exchange rates using MAR models with Student-t innovations. The model parameters were estimated using the maximum likelihood method. The forecast results of the MAR model were compared with some competing asymmetric Generalised Autoregressive Conditional Heteroskedastic (GARCH) models. Comparisons were based on the Berkowitz tail test. The test results suggested that the MAR model provided the best out-of-sample tail-density forecasts. The findings support the suggestion that the MAR models are well suited to capture the kind of data dynamics present in financial data and provide a useful alternative to other models.

Keywords: Mixture Autoregressive Models, Density Forecasts, GARCH Models, Time Series Analysis, Exchange rate.

1.0 INTRODUCTION
Forecasting play a very significant role in economics and finance just as they do in any other field of science. Evaluating accurate or dependable predictions is of primary concern in practice: A large chunk of the existing forecast literature is focused on evaluating point forecasts, a smaller slice on interval forecasts and a much thinner slice on probability forecasts. Point forecasts have been noted to be generally unsuitable as forecasts generated from quite a number of financial and economic models are not readily summarized by point forecasts (Berkowitz, 2001). Density forecasts have become more popular as applications in real life scenarios require not only a forecast estimate but also the uncertainty associated with such a forecast (Akinyemi, 2013). Applications of density forecasts span across the field of macro-economics a popular example is the 'fan-chart' of inflation and GDP published by the Bank of England and by the Sveriges Risk

bank in Sweden in their quarterly inflation reports (Schultefrankenfeld, 2014). Further applications can be found in finance with the major area being in risk management. Distributional forecasts of a portfolio are issued with the purpose of tracking measures of portfolio risk such as Value at Risk (VaR) and Expected Shortfall (ES) and the generation of density forecasts from option pricing data (Diebold, Gunther, and Tay, 1998). Good forecasts based on a series are however dependent on selecting an appropriate model to capture the dynamics of that series.

Mixture Autoregressive (MAR) models belong to the class of finite mixture models, this class of models have a number of interesting properties that make them viable models for diverse time series data in real life scenarios. Some of these properties include their ability to model both unimodal and multimodal conditional distribution as well as capture conditional heteroskedasticity (a property that occurs in most financial time series). In addition, the MAR model lends a flexible approach for capturing multiple regimes in financial data. These properties have made the MAR models and its variations popular in modelling non-linear time series. The application of this class of models can be found in finance, medicine, engineering among others. An important application in finance is in modelling exchange rate data (Akinyemi, 2013).

The Foreign Exchange (FOREX) market is at the very core of international trade, as world economies interact directly or indirectly through buying and selling (export and import) of various good and services. Nigeria, one of the fastest growing economies in sub Saharan Africa, not only imports most of her raw material and machinery, but is heavily involved in the trade of crude oil and related products. Hence, the Nigerian foreign exchange market plays a vital role in the socio-economic dynamics of the country.

A good number of researchers have dedicated time to study the dynamics of the Nigerian foreign exchange market as well as the behaviour of the Nigerian Naira (NGN) in relation to other currencies such as the US Dollar (USD), British Pound Sterling (GBP) among others. Mordi(2006) examined the effect of exchange rate volatility on economic management in Nigeria while, Yaya & Shittu (2010) assessed the impact of inflation and exchange rate on conditional stock market volatility. Adeoye & Atanda (2012) examined the severity, consistency and persistency of exchange rate volatility in Nigeria and their results allude to the existence of volatility persistence in both the nominal and real exchange rates.

Also, Usman & Adejare(2013) studied the effect of exchange rate volatility on the Nigerian economy.

Similarly, Exchange rate volatility in Nigeria have been extensively studied using various times series models the most popular of which are the GARCH type models. Some of such studies include and are not limited to: Olowe (2009) who modelled Naira/Dollar exchange rate volatility using GARCH (1, 1), GJR-GARCH (1, 1), EGARCH (1, 1), APARCH (1, 1), IGARCH (1, 1) and TS-GARCH (1, 1), Bala and Asemota (2013) compared performance of variants of the GARCH models with and without the incorporation of exogenous breaks in model estimation. Their findings showed that performance of the models improved when volatility breaks were included in the estimated models. Isenah and Olubusoye (2016) forecasted exchange rate dynamics using the GO-GARCH approach, they explored generalized orthogonal GARCH (GO-GARCH) models for forecasting time-varying conditional correlations and variances.

This paper focuses on the application of the MAR model introduced by Wong & Li(2000). The flexibility of the class of models has made them increasingly preferred candidates for capturing stylized properties of different financial time series. An important property of the MAR model is that the shape of the conditional distribution of a forecast depends on the recent history of the process Boshnakov (2009).

The aim here is to evaluate the tail forecast density of exchange rates time series based on the class of MAR models (see Wong & Li (2000), Wong, Chan, & Kam (2009) and (Akinyemi (2013)). We compare these tail density forecasts to those based on selected asymmetric GARCH models. Evaluating and comparing the tail density forecasts, create an avenue to further assess, the overall quality of the conditional loss distributions of the parametric methods used in estimating risk measures like VaR and ES.

The rest of the paper is structured as follows: in Section 2, the data and methods used in the study are discussed including a detailed description of the class of MAR models. The main results of the study are presented in Section 3, in particular, The MAR (3; 2, 2, 1) model is applied to evaluating the density forecast of some selected exchange rate data. The accuracy of the density forecasts evaluated based on the MAR model is then compared some asymmetric GARCH models using the Berkowitz test. Section 4 provides a discussion on the findings of the study and Section 5 concludes.

2.0 DATA AND METHODS

2.1 Data Description

The data used in this study is time series of FOREX rates of the Nigerian currency (Naira) against three major currencies (the US Dollar (USD), the European Euro (EUR), and the British Pound (GBP)) sourced from the Central Bank of Nigeria's statistical and annual reports. The time period covered is between 12/06/2010-12/06/2015 spanning 5years containing 1,836 daily rates. Table 1 describes the main characteristics of the empirical distributions of the compounded return series of the exchange rates over the evaluation period while Fig. 1a-4c reports time series plots of the actual data (Fig. 1a-1c), time series plot of the returns series (Fig. 2a-2c), histogram of the returns series (Fig.4a-4c) and Autocorrelation (ACF) plot (Fig. 3ai-3cii) of the returns series for the GBP/NGN, USD/NGN and EUR/NGN.

Table 1: Summary of descriptive statistics of USD/NGN, EUR/NGN and GBP/NGN rate of returns series

Statistic	N	Minimum	Maximum	1st Quartile	3rd Quartile	Mean	SD	Variance	Stdev	Skewness	Kurtosis
USD/NGN	1826	-0.0356	0.0435	-0.0036	0.0039	0.0002	0.0070	4.90E-05	0.007	0.0965	2.0773
EUR/NGN	1826	-0.0368	0.0483	-0.0039	0.0041	0.0001	0.0075	5.60E-05	0.008	0.1886	1.9193
GBP/NGN	1826	-0.0424	0.047	-0.0041	0.0042	0.0002	0.0073	5.40E-05	0.007	0.2125	2.3076

Kurtosis and skewness are of special interest when modelling extreme events in risk management. The kurtosis parameter is a measure of the combined weight of the tails relative to the rest of the distribution, most often, kurtosis is measured against the normal distribution. All 3 data sets have kurtosis less than 3, so that we suspect that the underlying distribution for the series are not normal, this we also observe from the histograms (Fig.4).

Skewness describes the degree of asymmetry of a distribution. Table 1 reveals that none of the data sets are symmetrical; however, all 3 data sets are positively skewed. In addition, the mean the rate of returns series hover between 0.0001 and 0.0002 and corresponding standard deviations around 0.007.

Furthermore, the Autocorrelation functions (ACF, Fig.3) indicate strong persistence across the lags in all 3 data sets. These properties are quite typical of most financial time series.

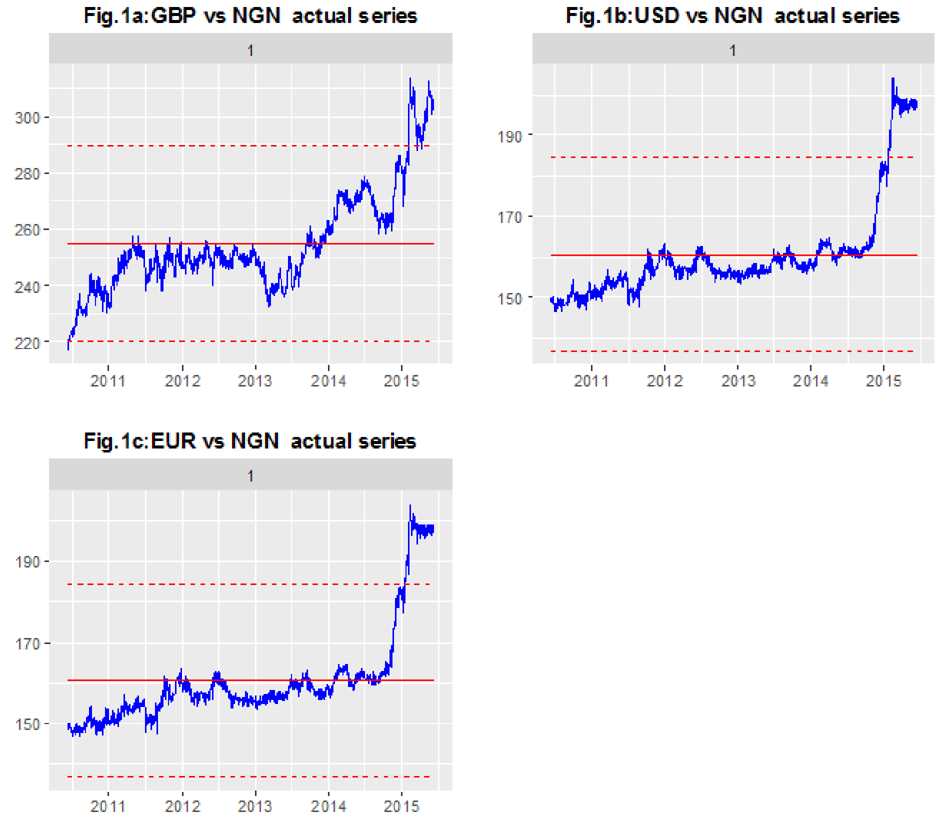

Figure 1: Time series plots of the actual series of GBP vs NGN, USD vs NGN and EUR vs NGN exchange rate data

For time series analysis, it is imperative to work with stationary process. Many of the formulated theorems in time series analysis assume a series to be stationary (at least in the weak sense). Processes whose Probability Density Functions do not change with time are termed stationary for analysis, the joint probability distribution must remain unchanged should there be any shift in the time series. Time series with persistence i.e. changing mean with time are non-stationary.

The time series plots of the actual exchange rate data are presented in Fig. 1a-c, they shows a steady rise in the series over the time period computed. We also observed that all 3 exchange rate series have exponential shape; it is best to thus turn to returns to analyze the series' behavior through time.

Figure 2: Returns series plot of GBP vs NGN, USD vs NGN and EUR vs NGN exchange rate data

The time plots of the returns series for each of the exchange rate data considered are presented in Fig. 2a-c above, we observed that the return series shows a high degree of variability, although this variability seem similar over several time periods. A sharp variation can also be observed towards the later part of 2011.

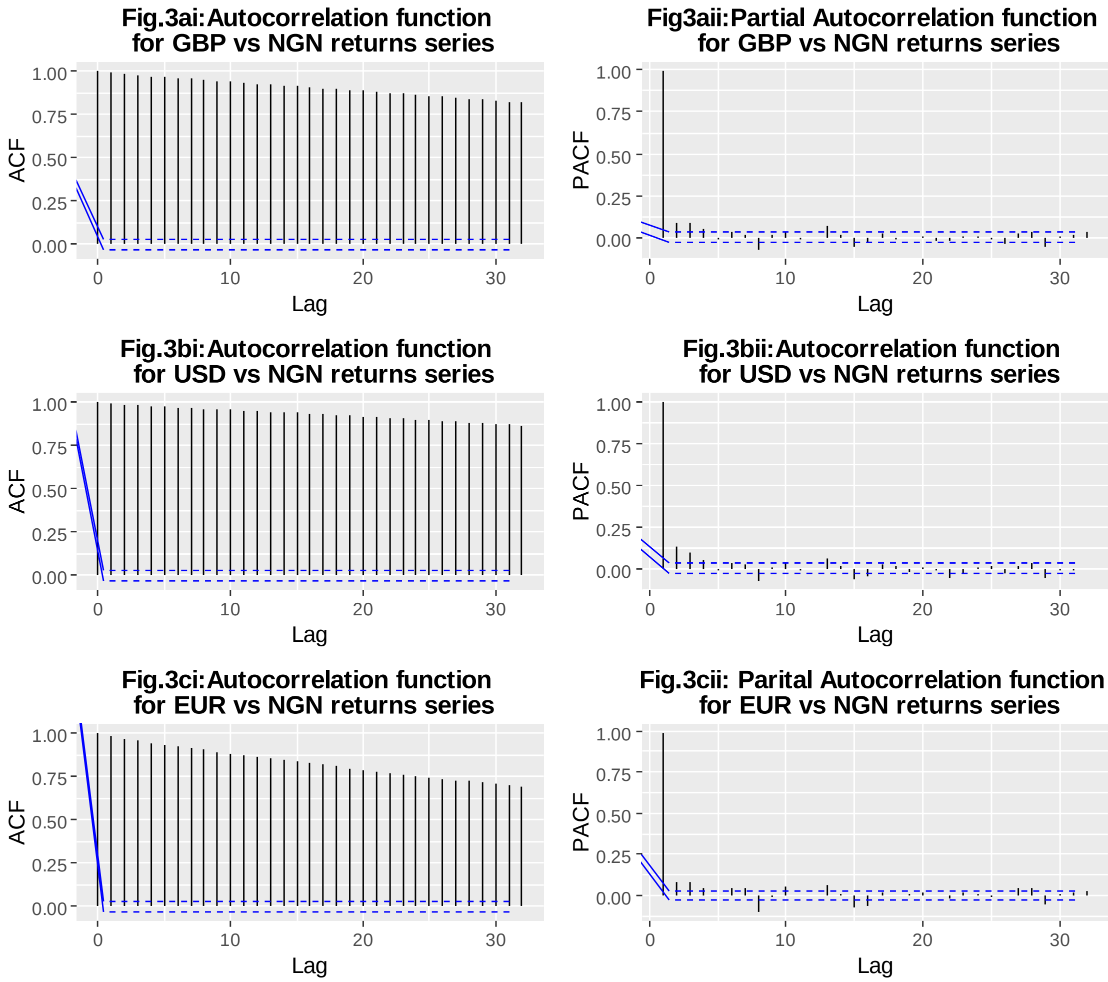

Figure 3: ACF and PACF plot of the returns series of GBP vs NGN, USD vs NGN and EUR vs NGN exchange rate data

Correlation of a time series with its own past and future values is called Autocorrelation. It is also referred as "lagged or series correlation". Positive autocorrelation is an indication of a specific form of "persistence", the tendency of a system to remain in the same state from one observation to the next (example: continuous runs of 0's or 1's). If a time series exhibits correlation, the future values of the samples probabilistically depend on the current and past samples. Thus, the existence of autocorrelation can be exploited in prediction as well as modeling time series. The Autocorrelation function (ACF) plot summarizes the correlation of a time series at various lags. It plots the correlation co-efficient of the series lagged by 1 delay at a time in the sample plot.

The partial autocorrelation function (PACF) gives the partial correlation of a time series with its own lagged values, controlling for the values of the time series at all shorter lags. It contrasts with the autocorrelation function, which does not control for lags. The ACF and PACF plots of all 3-exchange rate returns series considered indicate strong persistence across all the lags.

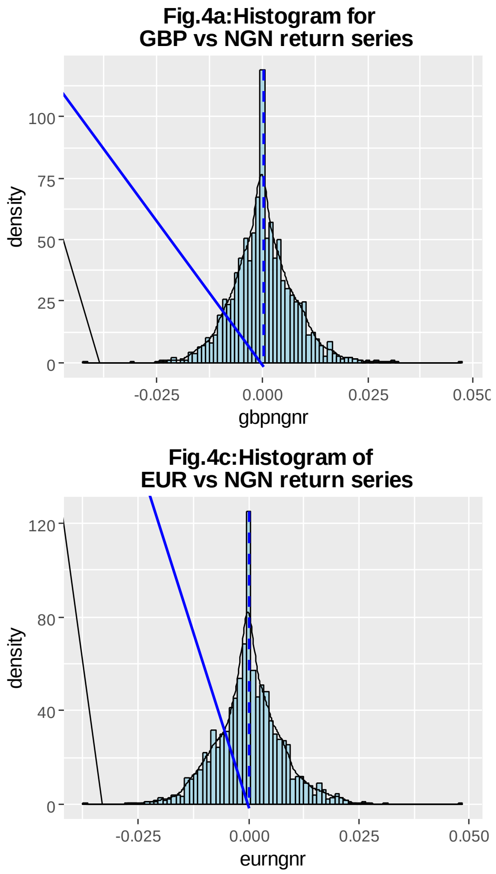

Figure 3: Histograms of the returns series of GBP vs NGN, USD vs NGN and EUR vs NGN exchange rate data

Plotting the histogram of the 3returns series (Fig. 4a-c), we can immediately reiterate that the returns series is now stationary and is suitable for further analysis.

2.1 Mixture autoregressive (MAR) approach to density forecast

2.1.1 Mixture Autoregressive (MAR) Models ((Wong & Li, 2000))

A process $\{y_t\}$ is said to be a mixture autoregressive (MAR) process if the conditional distribution function of y_t given past information is given by

$$F_{t|t-1}(x) = \sum_{k=1}^{g} \pi_k F_k\left(\frac{x-\phi_{k,0} \sum_{i=1}^{p_k} \phi_{k,i} y_{t-i}}{\sigma_k}\right), \qquad (2.1)$$

where, g is a positive integer representing the number of components in the model and the k^{th} component of the model, for k = 1, . . . , g, is specified by its mixing proportion,$\pi_k > 0$, scale parameter,$\sigma_k > 0$, autoregressive order,p_k, intercept,$\phi_{k,0}$ autoregressive coefficients,$\phi_{k,i}, i = 1, \dots, p_k$, and cumulative distribution function,$F_k(.)$. The mixing proportions π_k define a discrete distribution π, so that $\sum_{k=1}^{g} \pi_k = 1$.

We denote by MAR$(g; p_1, \dots, p_g)$ a g-component MAR model whose components are of orders $p_1, \dots, p_g$. The noise distribution functions$F_k, k = 1, \dots, p$, are typically taken to be standard Gaussian (Wong & Li, 2000) or (standardized) Student-t (Wong, Chan, & Kam, 2009). We will denote by $f_k(.)$ the corresponding probability density functions. It is also convenient to set $p = max_{1 \le k \le g} p_k$ and $\phi_{k,i} = 0$ for $i > p_k$.

A useful interpretation of the MAR model is that at each time t, one of g autoregressive like equations is picked at random to generatey_t. Let $\{z_t\}$ be an independently and identically distributed (i.i.d.) sequence of discrete random variables with distribution π, then $y_t$can be written as

$$y_t = \phi_{z_t,0} + \sum_{i=1}^{p} \phi_{z_t,i} y_{t-i} + \sigma_{z_t} \epsilon_{z_t}(t), \qquad (2.2)$$

where $\{\epsilon_{z_t}\}$ are jointly independent and independent of past y_t and the probability density of $\{\epsilon_{z_t}\}$ is f_k(.) (for further details see Boshnakov (2009; 2011). Let $\{z_t\}$ be an i.i.d. sequence of random variables with distribution π such that$Pr\{z_t = k\} = \pi_k, k = 1, \dots, g$, define a vector $Z_t = [Z_{t,1}, \dots, Z_{t,g}]'$ such that,

$$Z_{t,k} = \begin{cases} 1 & if\ z_t = k \\ 0 & otherwise \end{cases}$$

Then, the process y_t can be written as in Boshnakov (2009),

$$y_t = \mu_{z_t}(y_t') + \sigma_{z_t}\epsilon_{z_t}(t) \tag{2.3}$$

Where

$$\mu_{z_t}(y_t') = \phi_{z_t,0} + \textstyle\sum_{i=1}^{p}\left(\phi_{z_t,i}y_{t-i}\right)\left(p = max_{1\le k\le g}P_k\right). \tag{2.4}$$

$\{Z_t, t > 0\}$ is a simple case of a hidden Markov chain on a finite state space,$\{Z_t, t > 0\}$ drives the dynamics of $Y_t = \left(y_t, \dots, y_{t-p+1}\right)'$. In this work we make the following five (5) assumptions:

i. For each $k \in \{1, \dots, g\}$, $\{Z_{t,k}: t \ge 0\}$ is an irreducible, aperiodic Markov chain on a finite space S with probability distribution $\pi_1, \dots, \pi_g$ and transition probability matrix $A = \{a_{ij}\}$, so that $Z_{t,k}$ inherits the properties of $\{Z_t\}$.

ii. The chain $\{Z_t\}$ is independent of the ϵ_t.

iii. $\{\epsilon_t\}$ are jointly independent and are independent of past ys.

iv. $\{\epsilon_t\}$ has a probability density function that is continuous and positive everywhere.

v. $f_{z_t}(y)$ is non-periodic and bounded on all compacts sets for all k and $z_t \in S$.

2.1.2 The MAR (3:2, 2, 1) Model

The MAR (3; 2, 2, 1) model is a mixture autoregressive model with three AR components. Two AR components are of order two and the third one is of order one, that is, p1 = p2 = 2, p3 = 1 and k = 3.

The MAR (3; 2, 2, 1) is such that,

$$y_t = \begin{cases} \phi_{1,0} + \phi_{1,1}y_{t-1} + \phi_{1,2}y_{t-2} + \sigma_1\varepsilon_1(t) & \text{with probabilit y } \pi_1 \\ \phi_{2,0} + \phi_{2,1}y_{t-1} + \phi_{2,2}y_{t-2} + \sigma_2\varepsilon_2(t) & \text{with probabilit y } \pi_2 \\ \phi_{3,0} + \phi_{3,1}y_{t-1} + \sigma_3\varepsilon_3(t) & \text{with probabilit y } \pi_3 \end{cases} \tag{2.7}$$

with conditional distribution given by:

$$F_{t|t-1}(x) = \pi_1 F_1\left(\frac{y_t - \phi_{1,1}y_{t-1} - \phi_{1,2}y_{t-2}}{\sigma_1}\right) + \pi_2 F_2\left(\frac{y_t - \phi_{2,1}y_{t-1} - \phi_{2,2}y_{t-2}}{\sigma_2}\right) + \pi_3 F_3\left(\frac{y_t - \phi_{3,1}y_{t-1}}{\sigma_3}\right)$$

The MAR(3;2,2,1) model for $F_i(.)$: i = 1,2,3 and Student-t with 3-degrees of freedom was investigated, the parameters of which were estimated using the maximum likelihood method, implemented using the R (R Core Team, 2013). The parameters estimated for each of the financial returns series based on the model are given in Tables 2.

Table 2: Estimated Parameters of the MAR (3;2,2,1) model with Student's t innovations for daily returns of USD/NGN, EUR/NGN and GBP/NGN

Series	π_1	π_2	π_3	σ_1	σ_2	σ_3	$\phi_{1,0}$	$\phi_{2,0}$	$\phi_{3,0}$	$\phi_{1,1}$	$\phi_{1,2}$	$\phi_{2,1}$	$\phi_{2,2}$	$\phi_{3,1}$
USD/ NGN	0.0581	0.2962	0.6456	0.0000	0.0064	0.0061	0.0000	0.0000	0.0000	0.0000	-0.8844	0.0000	-0.6313	-0.0675
EUR/ NGN	0.1462	0.7668	0.0870	0.0053	0.0075	0.0002	0.0003	0.0002	0.0000	-0.8238	-0.0143	-0.7336	-0.0041	0.0122
GBP/ NGN	0.2643	0.3280	0.4077	0.0097	0.0052	0.0048	0.0013	0.0020	-0.0018	-0.5461	0.1642	-0.4152	-0.1106	-0.2519

2.1.3. Asymmetric Generalized Autoregressive Conditional Heteroskedastic (GARCH) Models

Here we outline the asymmetric GARCH models considered in this study, for a full reposition on GARCH models and their variations see Bollerslev (1986), Hentschel (1995) and Peters (2001).For this study, the following model were compared to the MAR model:

In the following, σ^2 - Is the conditional variance and μ -is the residual

1. Exponential Generalized Autoregressive Conditional Heteroskedastic (EGARCH): $\ln(\sigma_t^2) = \chi + \beta \ln(\sigma_{t-1}^2) + \gamma \frac{u_{t-1}}{\sqrt{\sigma_{t-1}^2}} + \alpha \left[\frac{|u_{t-1}|}{\sqrt{\sigma_{t-1}^2}} - \sqrt{\frac{2}{\pi}} \right]$, where, σ^2=the conditional variance, β =the coefficient of lags of the natural logarithm of the conditional variance, μ =the innovation of variance, γ= the innovation of the second term in the innovation of variance and α = the coefficient of the absolute value of the lag of the innovation of the conditional variance.

2. Glosten-Jagannathan-Runkle Generalized Autoregressive Conditional Heteroskedastic (GJR-GARCH):
$$\sigma_t^2 = \alpha_0 + \alpha_1 u_{t-1}^2 + \beta \sigma_{t-1}^2 + \lambda u_{t-1}^2 I_{t-1},$$
Where σ^2=the conditional variance, β =the coefficient of lags of the conditional variance, α_0= the intercept, α_1= the coefficient of the absolute

value of the lag of the innovation of the conditional variance and I is a dummy variable that takes the value of 1 when the shock is less than 0 (negative) and 0 otherwise.

3. Component Standard generalized autoregressive conditional heteroskedastic (csGARCH): $\sigma_t^2 = q_t + \sum_{j=1}^{q} \alpha_j (\mu_{t-j}^2 - q_{t-j}) + \sum_{j=1}^{p} \beta_j (\sigma_{t-j}^2 - q_{t-j})$

$$q_t = \omega + \rho q_{t-1} + \phi(\mu_{t-1}^2 - \sigma_{t-1}^2)$$

The model decomposes the conditional variance into a permanent and transitory component, Where σ^2= the conditional variance, q_t= the transitory component, β = the coefficient of lags of the conditional variance less of the transitory component, α_j = the coefficient of the absolute value of the lag of the innovation of the conditional variance less of the transitory component, ρ = the coefficients of the lags of the transitory components.

4. Nonlinear Asymmetric generalized autoregressive conditional heteroskedastic (NGARCH): $\sigma_t^2 = \omega_0 + \alpha_1 (u_{t-1} - \theta\sigma_{t-1})^2 + \beta\sigma_{t-1}^2$ where, ω_0= the intercept, α_1= the coefficient of the absolute value of the lag of the innovation of the conditional variance and β =the coefficient of lags of the conditional variance

5. Asymmetric Power autoregressive conditional heteroskedastic (APARCH):

$$\sigma_t^2 = (\omega_0 + \sum_{j=1}^{m} \varsigma_j \upsilon_{jt}) + \sum_{j=1}^{q} \alpha_j (|\mu_{t-j}| - \gamma_j \mu_{t-j})^2 + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2$$

where, σ^2= the conditional variance, β = the coefficient of lags of the natural logarithm of the conditional variance, μ = the innovation of variance.

2.3. Density Forecasts

A density forecast is an estimate of the future probability distribution of a random variable, conditional on the information available at the time of the forecast. It gives a complete characterisation of the uncertainty associated with a prediction, as against the point forecast which does not provide information about the uncertainty of the prediction (Diebold et al., 1998).We obtained the one-step ahead density forecasts based on the MAR model by applying the following theorem.

Theorem 2.1. For each $h \geq 1$ the condition characteristic function, $\varphi_{t+h|t}(s) \equiv E\left(e^{isy_{t+h}}\middle|\mathcal{F}_t\right)$,of the h−step predictor at time t of the MAR process is given by

$$\varphi_{t+h|t}(s) \equiv E\left(e^{isy_{t+h}}\middle|\mathcal{F}_t\right) \tag{2.8}$$

$$= E\left(E\left(e^{isy_{t+h}}\middle|\mathcal{F}_t, z_{t+h}, \dots, z_{t+1}\middle|\mathcal{F}_{t-h}\right)\right.$$

$$= \sum_{k_1,\dots,k_h}^{g} \left(\pi_{k_1}, \dots, \pi_{k_h}\right) e^{is\left(\mu_{k_1,\dots,k_h}(t+h)\right)} \prod_{i=0}^{h-1} \varphi_{k+h-i}\left(\theta_{h-i}^{k_1,\dots,k_h} s\right),$$

where

$$\mu_{k_1,\dots,k_h}(t+h) = \sum_{i=1}^{p} \beta_i^{k_1,\dots,k_h} y_{(t+1-i)} + \beta_0^{k_1,\dots,k_h} \tag{2.9}$$

For the one step, ahead density forecast that is $h = 1$, this gives

$$\varphi_{t+1|t}(s) = \sum_{k=1}^{g} \pi_k e^{is\mu_k(t+1)} \varphi_k\left(\sigma_k, \phi_{k,1} s\right), \tag{2.10}$$

See Boshnakov (2009) for proof.

2.4. Density Forecast Evaluation

a. Diebold, Gunther, & Tay (1998) use the probability integral transforms (PITS) to evaluate density forecasts. They use graphical tools to test whether the resultant series consists of independently and identically distributed uniform random variables $U(0,1)$. They assess independence by examining the correlogram and plot of the probability density function (pdf) to assess uniformity. They argue that statistical tests do not give insight into the reasons for rejection.

The main idea behind the probability integral transform goes as far back as Rosenblatt (1952) and was made popular by Diebold, Gunther, & Tay (1998).However, non-parametric tests are quite data intensive. Research shows the need for at least 1000 observations for a relatively reliable conclusion Berkowitz (2001).

b. (Berkowitz, 2001) introduces an extension of the Rosenblatt transformation. He advocates for a simple transformation to normality and suggests working with the inverse normal CDF transformation. That is, rather than $\{z_t\}_{t=1}^{n}$, the observed portfolio returns are transformed to create a series, $z_t = \Phi^{-1}(\hat{F}(y_t))$.

Let $\Phi^{-1}(.)$ be the inverse of the standard normal distribution function, for a sequence of forecasts regardless of the underlying distribution of the portfolio returns.

Proposition 2.1If the series $r_t = \int_{-\infty}^{x_t} f(u)du$ is distributed as an i.i.d. $U(0,1)$, then

$$z_t = \Phi^{-1}[r_t] = \int_{-\infty}^{x_t} f(u)du \;\; is\ an\ i.i.d\ N(0,1) \qquad (2.11)$$

Eq. (2.11) is the inverse PIT. This transformation is widely used to generate random numbers in computations. For proof, see Berkowitz (2001).

Proposition 2.2. Let $h(z_t)$ be the density of z_t and let $\phi(z_t)$ be standard normal. Then $\log\,[f(y_t)/\hat{f}(y_t)] = \log[h(z_t)/\phi(z_t)]$. For proof, see Berkowitz (2001).

where $\hat{f}(y_t)$is the forecasted probability density function of the returns series and with corresponding forecasted distribution function $\hat{F}(y_t)$,$h(z_t)$ is the density of (z_t). Proposition 2.2 implies that the ratio between the actual probability density and the forecasted density of data y_t should be the same as the ratio between the density of the transformed variable z_t and the standard normal density. Eq. (2.11) and (2.12) above enable the use of Gaussian likelihood tools to test the null hypothesis that the data follows a normal distribution. They also establish that the inaccuracies in the density forecast will be preserved in the transformed data.

Proposition 2.1 and 2.2 are due to Berkowitz (2001) (For proof of the propositions see Berkowitz (2001)). He proposed Likelihood Ratio (LR) tests based on a censored likelihood where he compares the shape of the forecasted tail of the density to the observed tail. The test statistic is based on the difference between the constrained $(\mathcal{L}(0,1))$ and the unconstrained $(\mathcal{L}(\hat{\mu},\hat{\sigma}^2))$. He forms an LR tail test that tests the null that the mean and variance of the violations equal those implied by the model as follows,

$$LR_{tail} = -2(\mathcal{L}(0,1) - \left(\mathcal{L}(\hat{\mu},\hat{\sigma}^2)\right) \sim \chi^2(2). \qquad (2.12)$$

The test will not only reject if the tails are too large but will also asymptotically reject if the tail has excessively small losses relative to forecast.

In this paper, we consider the Berkowitz's tail test approach for testing density forecasts as we are interested in the tails of the forecasts.

3.0 DATA ANALYSIS RESULTS

The performance of the MAR(3:2,2,1) was compared against some popular GARCH models. The asymmetric models considered include exponential generalized autoregressive conditional heteroskedastic (EGARCH), Glosten-Jagannathan-Runkle generalized autoregressive conditional heteroskedastic (GJR-GARCH), Component Standard generalized autoregressive conditional heteroskedastic (csGARCH), Nonlinear Asymmetric generalized autoregressive conditional heteroskedastic (NGARCH), Asymmetric Power autoregressive conditional heteroskedastic (APARCH). The out of sample density forecasts for Nigerian Naira (NGN) against the US dollar(USD) and Great Britain Pound (GBP) are computed. The density forecasts from the models are derived from parameters estimated with the entire data. We calculate the density forecasts under the assumption that the estimated model parameters are the population values, thereby ignoring parameter estimation uncertainty (see West (1996) and McCracken (2000) for details about effects of parameter estimation uncertainty).

Tables 3-8 presented the results of the test at 95% and 99% confidence levels i.e. α = 5% and α = 1%. Where: uLL=The unconditional Log-Likelihood of the maximized values.

rLL = The restricted Log-Likelihood with zero mean, unit variance and zero coefficients in the autoregressive lags.

LR = The Likelihood Ratio Test Statistic with corresponding test statistic p-value, $LR_{p.}$

Table 3: Likelihood ratio tail test for one-step ahead density forecast for USD/NGN daily Returns

Models	uLL		rLL		LR		LR_p		Decision	
	1%	5%	1%	5%	1%	5%	1%	5%	1%	5%
95% Likelihood ratio tail test for one-step ahead density forecast for USD/NGN daily Returns										
eGARCH(1,1)	-155.03	-155.03	-158.55	-158.55	7.04	7.04	0.03	0.03	fail to reject NULL	reject NULL
GJR-GARCH(1,1)	-158.09	-158.09	-161.35	-161.35	6.52	6.52	0.04	0.04	fail to reject NULL	reject NULL
csGARCH(1,1)	-176.02	-176.02	-178.50	-178.50	4.95	4.95	0.08	0.08	fail to reject NULL	fail toreject NULL
NGARCH(1,1)	-166.41	-166.41	-169.48	-168.48	6.14	6.14	0.05	0.05	fail to reject NULL	reject NULL
APARCH (1,1)	-151.18	-151.18	-154.56	-154.56	6.76	6.76	0.03	0.03	fail to reject NULL	reject NULL
MAR-t(3:2,2,1)	-20.46	-127.09	-23.25	-130.07	5.96	5.96	0.06	0.05	fail to reject NULL	fail toreject NULL
99% Likelihood ratio tail test for one-step ahead density forecast for USD/NGN daily Returns										
eGARCH(1,1)	-20.29	-155.03	-23.69	-158.55	6.79	7.04	0.03	0.03	fail to reject NULL	fail to reject NULL
GJR-GARCH(1,1)	-20.49	-158.09	-23.89	-161.35	6.80	6.52	0.03	0.04	fail to reject NULL	fail to reject NULL
csGARCH(1,1)	-22.75	-176.02	-26.12	-178.50	6.72	4.95	0.03	0.08	fail to reject NULL	fail to reject NULL
NGARCH(1,1)	-13.90	-166.41	-18.74	-168.48	9.68	6.14	0.01	0.05	reject NULL	fail to reject NULL
APARCH (1,1)	-25.01	-151.18	-27.43	-154.56	4.85	6.76	0.09	0.03	fail to reject NULL	fail to reject NULL
MAR-t(3:2,2,1)	-20.46	-127.09	-23.25	-130.07	5.59	5.96	0.06	0.05	fail to reject NULL	fail to reject NULL

Here, the MAR(3:2,2,1) model and csGARCH model successfully failed to reject at both 1% and 5% levels of significance, whereas, all the other models failed to reject. Hence, the MAR (3:2,2,1)model with student-t innovations and the csGARCH model performed better than the other models in this instance.

Table 4: Likelihood ratio tail test for one-step ahead density forecast for EUR/NGN daily Returns

Models	uLL		rLL		LR		LR_p		Decision	
	1%	5%	1%	5%	1%	5%	1%	5%	1%	5%
95% Likelihood ratio tail test for one-step ahead density forecast for EUR/NGN daily Returns										
GJR-GARCH(1,1)	-190.35	-190.35	-193.59	-193.59	6.48	6.48	0.04	0.04	fail to reject NULL	reject NULL
csGARCH(1,1)	-221.60	-221.60	-228.10	-228.10	13.00	13.00	0.00	0.00	reject NULL	reject NULL
NGARCH(1,1)	-198.68	-198.68	-198.68	-198.68	14.07	14.07	0.00	0.00	reject NULL	reject NULL
APARCH(1,1)	-192.56	-192.56	-192.56	-192.56	11.32	11.32	0.00	0.00	reject NULL	reject NULL
MAR-t(3:2,2,1)	-18.10	-106.40	-18.10	-106.40	5.22	12.83	0.07	0.00	fail to reject NULL	reject NULL
99% Likelihood ratio tail test for one-step ahead density forecast for EUR/NGN daily Returns										
eGARCH(1,1)	-26.26	-193.10	-28.62	-197.42	4.71	8.85	0.10	0.01	fail to reject NULL	fail to reject NULL
GJR-GARCH(1,1)	-26.32	-190.35	-28.67	-193.59	4.71	6.48	0.10	0.04	fail to reject NULL	fail to reject NULL
csGARCH(1,1)	-32.23	-221.60	-33.88	-228.10	3.30	13.00	0.19	0.00	fail to reject NULL	reject NULL
NGARCH(1,1)	-26.24	-198.68	-27.84	-205.72	4.69	14.07	0.10	0.00	fail to reject NULL	reject NULL
APARCH (1,1)	-18.10	-192.56	-28.59	-198.22	4.71	11.32	0.10	0.00	fail to reject NULL	reject NULL
MAR-t(3:2,2,1)	-12.73	-107.23	-16.50	-116.91	7.53	19.35	0.02	0.00	fail to reject NULL	reject NULL

In Table 4, the MAR(3:2,2,1) model only successfully failed to reject 1% level of significance but rejected at 5% levels of significance, whereas, the other eGARCH(1,1) and the GJR-GARCH(1,1) model failed to reject at both 1% and 5% level of significance.

Table 5: Likelihood ratio tail test for one-step ahead density forecast for GBP/NGN daily Returns

Models	uLL		rLL		LR		LR_p		Decision	
	1%	5%	1%	5%	1%	5%	1%	5%	1%	5%
95% Likelihood ratio tail test for one-step ahead density forecast for GBP/NGN daily Returns										
eGARCH(1,1)	-174.90	-174.90	-29.29	-176.82	3.84	3.84	0.15	0.15	fail to reject NULL	fail to reject NULL
GJR-GARCH(1,1)	-186.61	-186.61	-29.50	-186.61	3.40	3.40	0.18	0.18	fail to reject NULL	fail to reject NULL
csGARCH(1,1)	-201.01	-201.01	-33.63	-203.56	5.10	5.10	0.08	0.08	fail to reject NULL	fail to reject NULL
NGARCH(1,1)	-184.38	-184.38	-28.62	-186.79	4.82	4.82	0.09	0.09	fail to reject NULL	fail to reject NULL
APARCH (1,1)	-182.99	-182.99	-29.48	-184.88	3.77	3.77	0.15	0.15	fail to reject NULL	fail to reject NULL
MAR-t(3:2,2,1)	-12.80	-133.25	-16.53	-143.13	7.45	19.76	0.02	0.00	fail to reject NULL	reject NULL
99% Likelihood ratio tail test for one-step ahead density forecast for GBP/NGN daily Returns										
eGARCH(1,1)	-26.80	-174.90	-29.29	-176.82	4.98	3.84	0.08	0.15	fail to reject NULL	fail to reject NULL
GJR-GARCH(1,1)	-26.97	-186.61	-29.50	-186.61	5.06	3.40	0.08	0.18	fail to reject NULL	fail to reject NULL
csGARCH(1,1)	-32.00	-201.01	-33.63	-203.56	3.26	5.10	0.20	0.08	fail to reject NULL	fail to reject NULL
NGARCH(1,1)	-25.92	-184.38	-28.62	-186.79	4.68	4.82	0.10	0.09	fail to reject NULL	fail to reject NULL
APARCH (1,1)	-26.95	-182.99	-29.48	-184.88	5.07	3.77	0.08	0.15	fail to reject NULL	fail to reject NULL
MAR-t(3:2,2,1)	-12.80	-133.25	-16.53	-143.13	7.45	19.76	0.02	0.00	fail to reject NULL	reject NULL

Table 5, reports that the MAR(3:2,2,1) model successfully failed to reject at only 1% level of significance but rejected at 5% levels of significance in both instances, whereas, the other models failed to reject at both 1% and 5% level of significance at all instances.

4.0 DISSCUSION

The results presented here showed that only the MAR (3:2, 2, 1) model failed to reject at both 1% and 5% level of significance at more instances than all the other methods. Also, all the other models including the MAR (3:2, 2, 1) model failed to reject at 1% level of significance. Hence, the MAR (3:2,2,1) model with student-t innovations tend to perform at par if not better than most of the other models, as we fail to reject the null hypothesis in most instances. The MAR (3:2, 2, 1) with Student-t innovations performed well in all instances of the 99% tail tests but not so well at the 95% tail tests for EUR/NGN and GBP/NGN.

The results of the 95% and 99% Berkowitz test indicate that the models with Student-t innovations do give a viable alternative or complementary model to asymmetric models for out of sample tail density forecasts for exchange rate returns.

We also observed from the tables that, in general, the MAR models tend to outperform the other models in most instances. These findings are consistent with the findings of De Raaij (2005) and Diebold, Gunther, & Tay (1998) that fat-tailed conditional distributions generally give more satisfactory density forecasts for financial time series.

5.0 CONCLUSION

Obtaining good density forecast relies heavily on the ability to make proper distributional assumptions and adequate modelling of the dynamics of the relevant conditional moments of financial returns. From the set of competing models investigated, the MAR(3;2,2,1) model with Student-t innovations delivered the best out-of-sample density forecasts in more cases than the asymmetric GARCH models considered. Hence, the MAR model is a promising candidate for complementing existing models.

REFERENCES

Adeoye, B., & Atanda, A. (2012). Exchange rate volatility in Nigeria: consistency, persistency and severity analysis. *CBN Journal of Applied Statistics, 2*(2), 29-49.

Akinyemi, M. (2013). *Mixture Autoregressive Models; Asymptotic Properties and Applications to Financial Risk.* Thesis, University of Manchester, School of Mathematics.

Anderson, T. W. (1954). A test of goodness of fit. *Journal of the American Statistical Association*, *49*(268), 765-769.

Bala, D., & Asemota, J. (2013). Exchange rates volatility in Nigeria: application of GARCH models with exogenous break. *CBN Journal of Applied Statistics, 4*(1), 89-116.

Berkowitz, J. (2001). Testing density forecasts, with applications to risk management. *Jornal of Business and Economic Statistics*, *19*(4), 465-474.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics, 31*(3), 307-327.

Boshnakov, G. N. (2009). Analytic expressions for predictive distributions in mixture autoregressive models. *Stat. Prob. Let*, *79*(15), 1704-1709.

Boshnakov, G. N. (2011). On first and second order stationarity of random coefficient models. *Linear Algebra Appl., 434*(2), 415-423.

Cramer, H. (1945). *Mathematical methods of statitics. Princeton Landmarks in Mathematics and Physics Series.* Princeton University Press.

De Raaij, G. R. (2005). Evaluating density forecasts from models of. *The European Journal of Finance, 11*(2), 151-166.

Diebold, F., Gunther, T., & Tay, A. (1998). Evaluating density forecasts with applications to financial risk management. *International Economic Review, 39*(4), 863-883.

Hentschel, L. (1995). All in the family nesting symmetric and asymmetric garch models. *Journal of Financial Economics, 39*(1), 71-104.

Isenah, G. M., & Olubusoye, O. E. (2016). Empirical Model for Forecasting Exchange Rate Dynamics: the GO-GARCH Approach. *CBN Journal of Applied Statistics, 7*(1b), 179-208.

Kuiper, N. H. (1962). Tests concerning random points on a circle. *Nederl. Akad. Wetensch. Proc. Ser. A, 63,* 38-47.

McCracken, M. W. (2000). Robust out-of-sample inference. *Journal of Econometrics, 99*(2), 195 - 223.

Mordi, C. (2006). Challenges of exchange rate volatility in economic management in Nigeria. *CBN Bullion, 30*(3), 17-25.

Neave, H. W. (1988). *Distribution-Free Tests.* Unwin Hyman.

Olowe, R. (2009). Modeling Naira/Dollar exchange rate volatility: application of GARCH and asymmetric models. *International Review of Business Research Papers, 5*(3), 377–398.

Peters, J.-P. (2001). *Estimating and forecasting volatility of stock indices using asymmetric GARCH models and (Skewed) Student-t densities.* Preprint, University of Liege, Belgium.

R Core Team. (2013). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.R-project.org/

Rosenblatt, M. (1952). Remarks on a multivariate transformation. *The Annals, 23*(3), 470-472.

Schultefrankenfeld, G. (2014). *Essays on central bank forecasting.* Thesis, University of Siegen, Economics, Business and Business Law.

Usman, O., & Adejare, A. (2013). Effect of exchange rate volatility on Nigeria economy (1991-2010). *International Journal of Academic Research in Economics and Management Sciences, 2*(6), 172–184.

West, K. (1996). Asymptotic Inference about Predictive Ability. *Econometrica, 64*(5), 1067-84.

Wong, C., & Li, W. (2000). On a mixture autoregressive model. *Journal of the Royal Statistical Society. Series B: Statistical Methodology, 62*(1), 95-115.

Wong, C., Chan, W., & Kam, P. (2009). A student t-mixture autoregressive model with applications to heavu-tailed financial data. *Biometrica, 96*(3), 751-760.

Yaya, O., & Shittu, O. (2010). On the impact of inflation and exchange rate on conditional stock market volatility: a re-assessment. *American Journal of Scientific and Industrial Research., 1*(2), 115-117.