E-ISSN: 2709-9407
P-ISSN: 2709-9393
JMPES 2021; 2(1): 27-35

www.mathematicaljournal.com
Received: 05-01-2021
Accepted: 07-03-2021

MU Bawa
Lecturer, Department of Mathematics, School of Science, Abubakar Tafawa Balewa University, Bauchi, Nigeria

Dr. Anil Shabri
Associate prof. lecturer in statistics Department of mathematical sciences, faculty of science UTM

Dr. HG Dikko
Reader Lecturer, Department of Statistics, Ahmadu Bello University, Zaria, Nigeria

Dr. J Garba
Senior Lecturer, Department of Statistics, Ahmadu Bello University, Zaria, Nigeria

Dr. S Sadiku
Senior Lecturer, Department of Statistics, Ahmadu Bello University, Zaria, Nigeria

Correspondence
MU Bawa
Lecturer, Department of Mathematics, School of Science, Abubakar Tafawa Balewa University, Bauchi, Nigeria

Forecasting performance of mixed data sampling (MIDAS) regressions, autoregressive distributed lag (ADL) model and hybrid of GARCH-MIDAS model: A comparative study

MU Bawa, Dr. Anil Shabri, Dr. HG Dikko, Dr. J Garba and Dr. S Sadiku

Abstract
This paper considers the Comparison of forecasting performance between Mixed Data sampling (MIDAS) Regressions model, Autoregressive distributed lag (ARDL) Model and hybrid of GARCH-MIDAS. The data employed for this study was secondary type in nature for all the variables and it is obtained from the publications of Central Bank of Nigerian bulletin, National Bureau of Statistics and World Bank Statistics Database dated, January, 2005 to Dec, 2019. The result of unit root test shows that all variables are stationary at level and after first differences at 5% level of significant. From the results we found that F-statistics 1.895554 is inside the regions defined as the lower and upper bound (3.62 and 4.16) at 5% level of significant, this implies that there's no long-run relationship between dependent variable (NSE) and independent Variable (CC). using forecasting evaluations with shows that that GARCH-MIDAS has a least value of RMSE and MAPE than ARDL and MIDAS model (1823.531 and 3.976542) is least than for MIDAS and Ardl models (2372.846, 4.765421 and 2134.732, 5.952348). Finally, we can conclude that GARCH- MIDAS model outperform MIDAS and ARDL models of Nigeria Stock Exchange.

Keywords: MIDAS Regression's, ARDL Model, GARCH-MIDAS, long-run relationships

1. Introduction

A recent survey by the Central Bank of Nigeria (CBN) informs that 99% customer activity in the Nigerian Banks is cash-related transactions (Simon Tomlinson, 2012). This is because majority of retail and service payments in Nigeria are made in cash. Therefore, it is imperative to analyze the impact of currency curriculum (CC) on economic performance. CC, which forms the most liquid money aggregate, refers to the amount of currency notes and coins in the hands of economic agents outside banking sector. Together with demand deposits, CC forms the two components of narrow money, which makes it a variable of interest to monetary economists. More so, CC dynamics are generally regarded as measure for monetization or demonetization of the economy. According to Stavreski, (1998), share of CC in money supply and ratio of CC to GDP are the two major indicator of the importance of CC in every economy (Satvreski, 1998). Just like CC, demand deposits (DD) are also universally accepted as money, as such, both from part of the money stock in the economy. However, problems arise regarding liquid monetary assets which fulfill the store-of-value functioning but not medium of exchange function (e.g. home deposit) and vice –versa (e.g. credit card).

The autoregressive distributed lag (ADL) is considered as the major workhouse in the dynamic single equation regressions whose optimal lags are determined on the basis of the Schwartz Information Criterion (SIC). In this criterion, the model with the least SIC value is regarded as parsimonious, while its corresponding lag structure is considered the optimal lag. Hence, such model is used for the in-sample and out-of-sample forecasts. The commonest approach of modeling and forecasting in econometrics involves the use of data sampled at the same frequency, restricting the choice of variables to those that meet this condition. This often affect thee forecast performance (Ghysels *et al*., 2015). However, the advent of mixed data sampling (MIDAS) regression and its variants has widened the scope of choosing variable allowing for macro-economic (low frequency) variables when forecasting (high frequency) financial series like board price, stock price, exchange rate and interest rate (Ghysels *et al.*, 2006); Clements and Galvao, 2010; Ghysels *et al.*, 2015). It is evident from the previous studies that, due to the dependence of the time-varying volatility analyses on high frequency data, the impacts of variables such as unemployment rate and inflation on volatility have not been thoroughly examined. Instead, studies are mostly limited to variables such as short-term interest rates, term premiums and default premiums. Mixed Data Sampling (MIDAS), a regression scheme developed by Ghysels *et al*., (2006), allows for the inclusion of data from different frequencies into the same model. This makes it possible to combine the high-frequency return data with macroeconomic data that are only observed in the lower frequencies such as monthly or quarterly.

1.1 Statement of the Problem

The GARCH Model is one of the most used model in estimating the volatility of a time series data. However, its main weakness is that it assumes that the conditional volatility remains over the entire period of time. This lead to spurious regression (Baillie and Morana, 2009). Also, the restrictions on the coefficients to be positive estimations to become possible is rarely met. In order to overcome these draw backs we intend to propose a hybrid model called GARCH-MIDAS, in the propose model our intent is to leverage on the unique advantage of the long-run component of the mixed data sampling (τ_t) which can change over a specific period of time, therefore The main advantage of the GARCH-MIDAS model is that it allows us to link the daily observations of stock returns with macroeconomic variables, in order to examine directly the macroeconomic variables impact on the stock volatility. Hence, also the model improves the prediction ability of the model, particularly for the long-term variance component.

1.2 Aim and Objectives of the study.

The aim of this paper is to propose a hybrid model for forecasting time series with an improved performance over the traditional forecasting model. Is achieved through the following specific objective:

I. Derivation of the hybrid model and implantations in R package
II. Comparing the performance of the proposed model with ARDL and MIDAS using forecast evaluation criteria

2. Review of Literature

2.1. Theoretical Review

2.1.1 Friedman's Theory of Demand for Money

Milton Friedman improved on Keynes liquidity preference theory by treating money like any other asset. He posits that individuals hold money for the services it provides to them in form of general purchasing power so that it can be conveniently used for buying goods and services. Friedman considers the demand for money merely as an application of a general theory of demand for capital assets in contrast to transactions and speculative demand for money. Friedman's theory is based on the premises that demand for money is affected by same factors as demand for any other asset. That individuals can hold their wealth in form of money, bonds, equity and real assets (e.g. cars, housing, land, etc.). That money yields return and provides services just like other capital assets. The various factors that determine the demand for money was used to derive the demand for money function. Equity (Shares) is another form of asset in which wealth can be held. The yield from equity is determined by the dividend rate, expected capital gain or loss and expected changes in the price level. (Jhingan, 2011).

2.1.2 Relationships between Stock Market Volatility and Currency Circulations in Nigerian Economy

Since the main purpose of CC is to provide for cash transactions within the economy, the development of CC should reflect both economic activity and changes in the price level. The gross domestic product, which is the widest indicator of economic activity, includes information on price developments among others. In terms of pace of growth, there is no first glance of clear relationship between CC and economic performance. Consequently, increasing consideration is now being dedicated to the impact of electronic payment system to economic performance. Several studies revealed that improving the electronic payment system by reducing the amount of CC ensured rapid growth in the economy (ref.). Even though, there have been a lot of empirical studies on the relationship between money supply/demand and economic performance in Nigeria, very few studies focus on the impact of CC on the economic growth. Such study is crucial because it deepens our understanding of monetary policy issues, while serving as a guide for policy makers in the developing countries.

2.2 Empirical Review

There are many empirical evidences across countries both the developed and the developing on the impact of money supply on real GDP; some are discussed in this work.

Obaid (2007) tests the causality relationship between money supply (M3) and real GDP in Egypt during the period (1970-2006) by using Granger test. Findings from his study revealed that there is no causality between the nominal money supply and nominal GDP during the study period, while when he used the real money supply and real GDP, he found that there exist mutual causality relationship between real money supply and real GDP in Egypt (non-neutral money), and thus the monetary policy is an effective policy on the real GDP in Egypt, the mutual causality relationship could help to forecast the GDP behavior within assumed volume of money supply by the economics policy making in Egypt. Ogunmuyiwa and Ekones (2010) investigated the impact of money supply on economic growth in Nigeria using annual data for the period 1980 to 2006. Applying Econometric technique (Ordinary Least Squares(OLS), Granger Causality test and Error Correction Model), the results revealed that although money supply is positively related to growth, the result is however insignificant in the case of GDP growth rates on the choice between contractionary and expansionary money supply.

2.3. Theoretical Framework

This study is anchored on the theoretical framework of modern quantity theory of money (QTM) (Milton Friedman, 1960), of High level of CC in Nigeria has been attributed to inefficiency of central bank of Nigeria monetary policy, the underdeveloped nature of the Nigeria banking system and failure of the economy to develop the informal sector. The model estimated in this study follows modern quantity theory of money (QTM) (Milton Friedman, 1960) that includes the traditional variables such as the real interest rates, currency circulations, velocity of money, Gross Domestic Product (GDP), etc. Drawing from the works of Emerson (2006) and Akhtaruzzaman (2008), the quantity theory of money identity is written as:

$$GDP_t = CC_t + V_t \quad (1)$$

Where CC_t is the amount of currency circulations at a point in time and V_t is the velocity of money.

3. Methodology

3.1 Model Specification and Estimation Techniques

The model for this study is specified as follows:

$$NSE=F(CC) \quad (2)$$

Where NSE = daily of Nigerian Stock exchange and CC = monthly Currency circulations in order to examine the influence of the seasonal impact of currency circulations on real GDP, the model specification above is presented below in a Mixed Data Sampling (MIDAS) form developed by Ghysels, Sinko and Valkanov (2007). Mixed data sampling (MIDAS) regressions allow us to estimate dynamic equations that explain a low frequency variable by high frequency variables and their lags. (Foroni & Marcellino, 2013)

$$NSE_i = \beta_0 + \beta_1 \beta(L^{\frac{1}{m}};\theta) CC_t^m + \varepsilon_t \quad (3)$$

Where NSE_i is the dependent variable is a daily data while CC^t is the explanatory variable of a monthly data, (m) denotes the frequency, ε is the disturbance, and

$$\beta(L^{\frac{1}{m}}) = \sum_{j=0}^{j\max} \beta(j) L^{\frac{j}{m}} = 1$$

is an order of polynomial and $L^{1/m}$ is the lag operator. Midas is much more flexible than a fat-weighting scheme since it can nest the equal weighting scheme by setting $\theta_1 = \theta_2 = 0$ .Another important characteristic of the Midas approach is that the slope coefficient can be easily obtained from the regression as the weights attached to the high frequency data are normalized and sum to one. We use the exponential Almon lag polynomial proposed by Ghysels et al. (2007) which most general weighting scheme compared to others as it is very flexible and can take many shapes. The input variable (independent variable) of CC is the monthly data with (Xlag=3) and the output variable (dependent variable) of GDP was taken daily as (Ylag=1).

$$y_t = \sum_{i=1}^{p} a_i y_{t-i} + \sum_{i=0}^{n} c_i^1 x_{t-i} + \varepsilon_t, \quad (4)$$

where ε_t is a scalar zero mean error term and x_t is a K-dimensional column vector process. Typically, a constant is included in (1), which we neglect here for brevity. The coefficients a_i are scalars while c_i^1 are row vectors. Using the lag operator L applied to each component of a vector, $L^k x_t = x_{t-k}$, it is convenient to define the lag polynomial a(L) and the vector polynomial c(L),

$$a(L) = 1 - a_1 L - ... - a_p L^p, \quad (5)$$

$$C(L)= c_0 + c_1 L + ... + c_n L^n. \quad (6)$$

Now, it is straightforward to write (1) more compactly:

$$a(L) y_t = c^1(L) x_t + \varepsilon_t.$$

The ADL form of the above equation can be represented as below

$$a(L)NSE_i = c^1(L)CC_t + \varepsilon_t \tag{7}$$

3.2. Proposed Modifications Hybrid Model

In this paper, we use a new class of component GARCH model based on the MIDAS (Mixed Data Sampling) regression. MIDAS regression models are introduced by Ghysels et al. (2006). MIDAS offers a framework to incorporate macroeconomic variables sampled at different frequency along with the financial series. This new component GARCH model is referred as GARCH- MIDAS, where macroeconomic variables enter directly into the specification of the long- term component. The GARCH-MIDAS model can formally be described as below. Assume the return on daily i in month t follows the following process:

$$r_{i,t} = \mu + \sqrt{g_t \tau_i \varepsilon_{it}^2} \ \forall_i = 1...N_t \ \ \varepsilon_{i,t}^2 / \phi_{i-1,t} \approx N(0,1) \tag{8}$$

Now i is the short scale (for example, ε_{it}^2 is a daily return), and t is the long or aggregated scale (the unit is monthly). The short-run component g_t is measured in daily, whereas the long-run τ_i is available on monthly basis.Where $\phi_{i-1,t}$ is the information set up to $(i-1)^{th}$ day of period t. expresses the variance into a short term component defined by g_t and a long term component defined by τ_i

$$r_{i,t} = \mu + \sqrt{g_t \tau_i \varepsilon_{it}^2} \tag{9}$$

$$r_{it} - \mu = (g_t \tau_i \varepsilon_{it}^2)^{\frac{1}{2}} \tag{10}$$

Square both side of the equations

$$(r_{i,t} - \mu)^2 = (g_t \tau_i \varepsilon_{it}^2) \tag{11}$$

Divided both side by τ_i in equation (11) with the assumptions given in equation

$$\frac{(r_{i,t} - \mu)^2}{\tau_i} = (g_t \varepsilon_{it}^2) \tag{12}$$

is a covariance –stationary of GARCH process. Substitute this equation below

$$g_t = \alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j h_{t-j} = \alpha_0 + \alpha(L)\varepsilon_t^2 + \beta(L) g_t \tag{13}$$

$$g_t = \alpha_0 + \alpha(L) \frac{(r_{i,t} - \mu)^2}{\tau_i} + \beta(L) g_t \tag{14}$$

Where $\alpha_0 = \alpha + \beta \leq 1 = (1 - \alpha - \beta)$

$$g_t = (1 - \alpha - \beta) + \alpha(L) \frac{(r_{i,t} - \mu)^2}{\tau_i} + \beta(L) g_t = (1 - \alpha - \beta) + \alpha_{i-t} \frac{(r_{i,t} - \mu)^2}{\tau_i} + \beta_{j-t} g_{t-j} \tag{15}$$

And τ_i is defined as in the spirit of MIDAS regression, and definitions of the long-term volatility component

$$\tau_i = m + \beta_1 \beta(L^{\frac{1}{m}}; \theta) X_t^{(m)} + \varepsilon_t \tag{16}$$

Then

$$\beta(L^{\frac{1}{m}}) = \sum_{j=0}^{j\max} \beta(j) L^{\frac{j}{m}} = 1$$

is an order of polynomial and $L^{1/m}$ is the lag operator. Midas is much more flexible than a fat-weighting scheme since it can nest the equal weighting scheme by setting $\theta_1 = \theta_2 = 0$ .Another important characteristic of the Midas approach is that the slope coefficient can be easily obtained from the regression as the weights attached to the high frequency data are normalized and sum to one.

$$\tau_i = m + \beta_1 \beta(L^{\frac{1}{m}};\theta) X_t^m + \varepsilon_t = m + CC_t^m + \varepsilon_t \tag{17}$$

The GARCH-MIDAS model is

$$g_t = (1-\alpha-\beta) + \alpha_{i-t} \frac{(r_{i,t}-\mu)}{\tau_i} + \beta_{t-j} g_{t-j} \tag{18}$$

$$NSE_t = (1-\alpha-\beta) + \alpha_{i-t} \frac{(r_{i,t}-\mu)^2}{\tau_i} + \beta_{t-j} g_{t-j} \tag{19}$$

Where NSE ^{t} is the for the short-term forecast we use the daily observations. And τ_i is the predict of the long-term volatility with the monthly observations. How to estimated (eqn. 19) below.

Description
A dataset containing the NSE for all shares from 2005-2019. And A data frame with 5478 rows and one variable. This function estimates a multiplicative mixed-frequency GARCH model. For the sake of numerical stability, it is best to multiply log returns by 100. Arguments data. Data frame containing a column named date of type 'Date'.

NSE ^{t} - is the high frequency dependent variable in df.

μ -covariate employed in mfgarch.

K- an integer specifying lag length K in the long-term component.

CC ^{i} - a string of the low frequency variable in the df.

Var. Ratio. Freq specify a frequency column on which the variance ratio should be calculated. Gamma if TRUE, an asymmetric GJR-GARCH is used as the short-term component. If FALSE, a simple GARCH (1,1) is employed. Weighting specifies the weighting scheme employed in the long-term component. Options are "beta. Restricted" (default). broom. mfgarch - a broom-like data. Frame with entries 1) estimate: column of estimated parameters 2) rob. std. err - standard errors 3) p. value - p-values.
Based on this explanations we are going to estimate the following parameters ($\alpha, \beta,$ m μ and γ) and note that NSE_{ti} is the deterministic functions of $r_{i-1,t}$ and historical observations. For estimated ($r_{i-1,t} - \mu)^2$ is not available. In order to obtain the positive value, the sufficient conditions for the conditional variance are α i > 0, $\alpha i \geq 0$, $\beta j \geq 0$, $i = 1\ldots, p$ and $j = 1, \ldots, q$. When $p = 0$, the GARCH (p, q) model

3.3. Unit Root Test
Recently, testing for unit roots has already become a standard procedure in time series studies and the application of unit root test such as DF-GLS Test. Unit root test is important because the absence of unit root indicates that the series has some variances are not being determined by time and that the effects of shocks dissolve over time. Besides that, the existence of non-stationary variables will cause a spurious regression which has high R^2 and t-statistic is significant but the results would not consist of any economic meaning.

3.3.1 DF-GLS test
Adopted is for a unit root in a time series. It performs the modified dickey-fuller test (known as the DF-GLS test) proposed by

Elliott, Rothenberg, and Woodford (1996). Essentially the test is an augmented Dickey-fuller test, similar to the test performed, time series is transformed via a generalized least squares(GLS) regression before performing the test, Elliott, Rothenberg, and stock and later studies have shown that this test has significantly greater power than the previous versions of the augmented Dickey-fuller test. Hypothesis, otherwise it has to be differentiated until is stationary. This test is used to determine the order of integration of a variable. The test states that if a particular series say Y has to be differenced n times (number of times, 1, 2, 3… n) before it becomes stationary then Y is said to be integrated of order n (it is written as I(n. The regression model for the test is given as:

$$\Delta y_t = \beta_1 + \beta_2 t + \delta y_{t-1} + \alpha \sum_{i=1}^{m} \Delta y_{t-1} + \varepsilon_t \qquad (20)$$

$\Delta Y_t = difference\ between\ variable\ and\ its\ own\ lag$

β_1=constant trend

β_2=is the parameter of time trend

δ=unit root

4. Results

Table 1: DF-GLS test with constant only

Variables	level	First diff	5% crit value	Stationary status
CC	2.285727	-2.672058**	-1.942781	I(1) after first difference
NSE	-4.432058		-1.942781	I(0) after first difference

Note that** indicate significant at 5% level.

The outcome of the unit root test for DF-GLS. The result of unit root test shows that all variables are stationary at level and after first differences at 5% level of significant. That means the variables are integrated at level and order one i.e. 1(0) and 1(1).

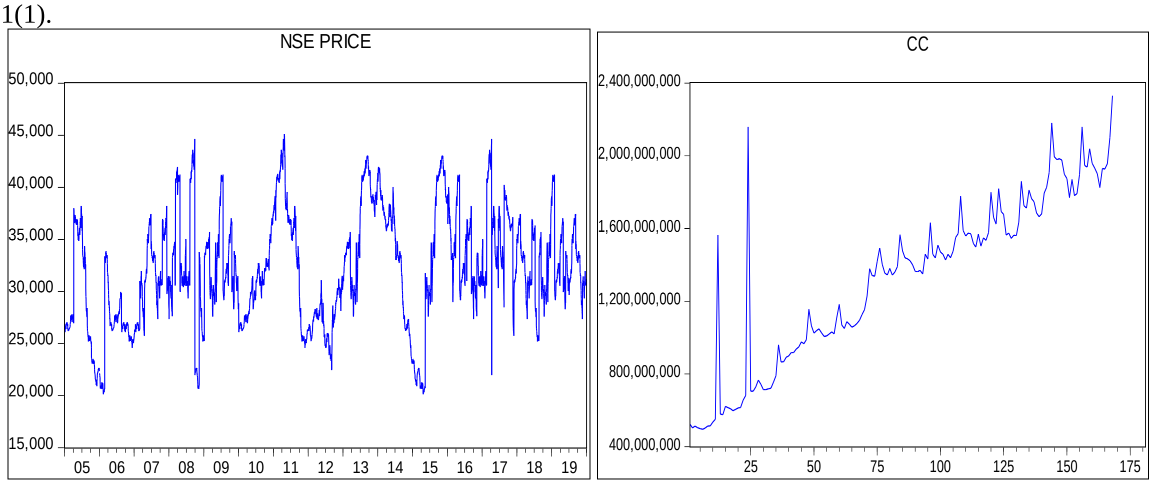

Fig 1: Graphs of variables showing test for stationarity.

4.1 Optimum lags Test Result:

Table 2: Result of Optimum Lags Test

Lag	LogL	LR	FPE	AIC	SC	HQ
0	-4918.003	NA	1.75e+24	61.50004	61.53848	61.51565
1	-4593.687	636.4696	3.20e+22	57.49609	57.61141	57.54292
2	-4581.803	23.02593	2.90e+22	57.39754	57.58974*	57.47558*
3	-4576.529	10.08691*	2.86e+22*	57.38161*	57.65069	57.49087
4	-4574.443	3.936079	2.92e+22	57.40554	57.75150	57.54602
5	-4573.699	1.387233	3.05e+22	57.44623	57.86907	57.61793
6	-4573.273	0.781642	3.19e+22	57.49092	57.99063	57.69383
7	-4571.787	2.693897	3.29e+22	57.52234	58.09893	57.75647
8	-4570.007	3.182163	3.38e+22	57.55008	58.20356	57.81544

* indicates lag order selected by the criterion

LR: sequential modified LR test statistic (each test at 5% level)

FPE: Final prediction error

AIC: Akaike information criterion
SC: Schwarz information criterion
HQ: Hannan-Quinn information criterion

Conducted the DF-GLS test, the result confirms the stationarity of the series variable at the level 1(0) and first different I (1), so it is important to determine the number of lag to be included in the regression. Since the variables are stationary at different order we are going to conduct Bounds test in order to test the long-run exist between the variables.

As presented in table 2, optimum lag order selection was carried out to determine the number of lag(s) to be included in the model prior to Bond test. The maximum lag for the model was selected based on the five different information criteria. It evident that from the table 2 that LR, FPE and AIC which agreed at 3 lag, while SC and HQ agrees at lag 2. Hence, the study adopted 3 as the maximum for the model.

Table 3: Statistics Test for Ardl Regression model

Variable	**Coefficient**	**Std. Error**	**t-Statistic**	**Prob.***
NSE_PRICE(-1)	0.984761	0.079798	12.34069	0.0000
NSE_PRICE(-2)	-0.062609	0.111931	-0.559355	0.00267
NSE_PRICE(-3)	0.021827	0.079415	0.274844	0.00388
CURRENCY_CIRCULATION	5.25E-07	3.85E-07	1.363395	0.00217
CURRENCY_CIRCULATION(-1)	-4.01E-08	4.04E-07	-0.099129	0.00922
CURRENCY_CIRCULATION(-2)	1.39E-07	4.04E-07	0.344891	0.04236
CURRENCY_CIRCULATION(-3)	-8.11E-08	3.90E-07	-0.207909	0.00156
C	1077.464	536.4123	2.008649	0.00023
R-squared	0.967611	Mean dependent var		31092.08
Adjusted R-squared	0.966167	S.D. dependent var		4756.174
S.E. of regression	874.8401	Akaike info criterion		16.43323
Sum squared resid	1.20E+08	Schwarz criterion		16.58382
Log likelihood	-1347.741	Hannan-Quinn criter.		16.49436
F-statistic	670.0461	Durbin-Watson stat		1.999834
Prob(F-statistic)	0.000000			

The table 3: above tells us R-square is 96.76% explain the variables in the model of ARDL and at lag 3 the probability value is significant for all the variables and also AIC is the best for estimating ARDL model statistics with has a least value than BIC, SC and HQC.

Table 4: ARDL Long Run Form and Bounds Test

Levels Equation				
Case 2: Restricted Constant and No Trend				
Variable	**Coefficient**	**Std. Error**	**t-Statistic**	**Prob.**
CURRENCY_CIRCULATION	9.70E-06	2.82E-06	3.434698	0.0008
C	19233.20	3874.976	4.963436	0.0000
EC = NSE_PRICE - (0.0000*CURRENCY_CIRCULATION + 19233.1972)				
F-Bounds Test		Null Hypothesis: No levels relationship		
Test Statistic	Value	Signif.	I(0)	I(1)
Asymptotic: n=1000				
F-statistic	1.895554	10%	3.02	3.51
K	1	5%	3.62	4.16
		2.5%	4.18	4.79
		1%	4.94	5.58
Actual Sample Size	165		Finite Sample: n=80	
		10%	3.113	3.61
		5%	3.74	4.303
		1%	5.157	5.917

The result in table 4 shows that the value of F-statistics 1.895554 is inside the regions defined as the lower and upper bound (3.62 and 4.16) at 5% level of significant, this implies that there's no long-run relationship between dependent variable (NSE) and independent Variable (CC). Therefore, we can't proceed to Cointegration test since there's no existing of long-run relationships between the variables.

Table 5: Summary of Candidate Models Identification of the Selections for lags.

Methods	**Goodness of fit**	**Noise Variance**	**Log Likelihood**	**Akaike Criteria**	**Bayesian Criteria**
Beta density	0.63605	480.7407	-45.2361	69.2556	71.2451
Step function	-419.65	8214.5088	-148.2692	277.3217	279.7091
U –Midas	0.83791	214.106	-40.7874	61.3518	63.3476
Almon Lag	0.82751	162.7415	-41.1293	57.042	58.2356

After carefully examining Table 5: above the top competing model with the least information criteria is Almo lag in bold with AIC of 57.042 with we can applying in this study.

3.3: Statistics Test for Midas Regression model

The NSE data is taken Xlag365 as the daily 365 representing a daily data whereas, the CC Xlag12 as 12 representing monthly data and 3 resenting lags:

Table 6: MIDAS: Normalized exponential Almon lag polynomial

Formula y ~ trend + mls (GDP, 1:3, 365) + mls (CC, 1:3, 12)				
	Estimate	**Std. Error**	**t value**	
Parameters:	2.191e+00	1.374e+00	1.595e+00	0.117
(Intercept) trend	8.923e-02	6.671e-03	1.338e+01	<2e-16 ***
x1	4.008e-01	1.355e-09	2.958e+08	<2e-16 ***
x2	2.431e-01	1.366e-09	1.779e+08	<2e-16 ***
x3	1.474e-01	1.311e-09	1.125e+08	<2e-16 ***
z1	3.504e-01	5.986e-11	5.854e+09	<2e-16 ***
z2	4.280e-01	5.439e-11	7.869e+09	<2e-16 ***
z3	4.280e-01	5.572e-11	7.682e+09	<2e-16 ***

Based on the result of table 6 it shows that after simulated the data all variable estimated are statistical significant since the probability value are less than 0.05 level of significant.

Table 7: Descriptive Statistics of GARCH-MIDAS OF NSE ON CC

variables	Min-value	Max-value	1st quarter	3rd -quarter	Mean	Median
NSE	5877641	5877641	5877641	5877642	5875644	5674321
CC	8280970	42961403	16952391	29490685	23037526	21673951

The descriptive statistics of table 7: tell us Nigeria stock exchange of the first and third quarter the value are very high than the first and third quarter of currency circulation. And the mean and median of currency circulations is less than NSE with it shows that the data of currency circulations have better descried the statistics than NSE because the NSE data are very large.

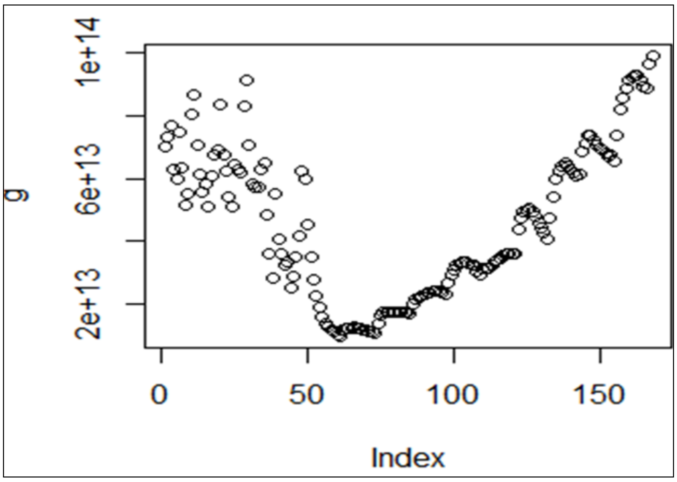

Fig 4: ACF graph of GARCH-MIDAS Model

Based on the figure 4 above the ACF plot shown that the data are stationary.

Table 8: Statistics Test for the Estimated of GARCH-MIDAS Model

	Coefficient est.	Std. Error	t.stat	Prob.
Mu	2.1690528	2.792425e-03	6.2308	0.000000e+01
Alpha	0.3018519	1.236056e-01	12.7654	1.460388e-02
Beta	0.7086918	1.213638e-01	43.8756	5.238916e-09
Theta	-1.4468365	1.154482e+04	5.8765	9.999000e-01
M	2.5979581	2.135216e-01	16.6541	0.000000e+00

Method: maximum likelihood
Sample size: 5478
Log-like hood- -1409.24
Akaike infor Criteria 2825.48
Bayesian info criterion 3452.64

The result of table 8 it tells us how to estimated GARCH-MIDAS of Nigerian stock exchange of daily observations at high frequency dependent variable with represent GARCH model while currency circulation of monthly a string of the low frequency variable. The estimate of alpha and beta are representing GARCH component while theta and m representing MIDAS component. The best infor criterions of GARCH-MIDAS is AIC because it has less value than BIC.

4. Evaluations of the Forecasting Performance of the models:

Given already preselected and evaluated models, user can use the function average forecast to evaluate the forecasting performance: Steps for in-sample and out-sample.

1. Firstly, I use the midas_r ( ) Almon lag command for the forecasting of in-sample and out-sample with the use of mls argument in as the midas lag structure (lag3) for CC variable (monthly data).
2. Secondly, use the average forecast ( ) command for the forecast of in-sample and out-sample with the use midas_r ( ). I use the in-sample as 5478.
3. For the GARCH-MIDAS A function from the mfgarch package in R was develop and use to spawn the MIDAS-GARCH model in conjunction with the stationary data provided which was verified using the ACF plot. The square return of the model was then used to investigate the performance of the model using the accuracy () function from the R basic function.

Table 9: Forecast evaluations of in and Out-Sample of Midas models

Model	RMSE	MAPE	MAE	MASE
ARDL	2372.846	5.952348	1874.036	
MIDAS	2215.461	4.765421	2431.61	
GARCH-MIDAS	1823.531	3.976542	3216.41	6.76531

Based on the results obtaining in table (9) using forecasting evaluations it tells us that GARCH-MIDAS has a least value of RMSE and MAPE than ARDL and MIDAS model. It shows that GARCH-MIDAS outperform the ARDL and MIDAS model.

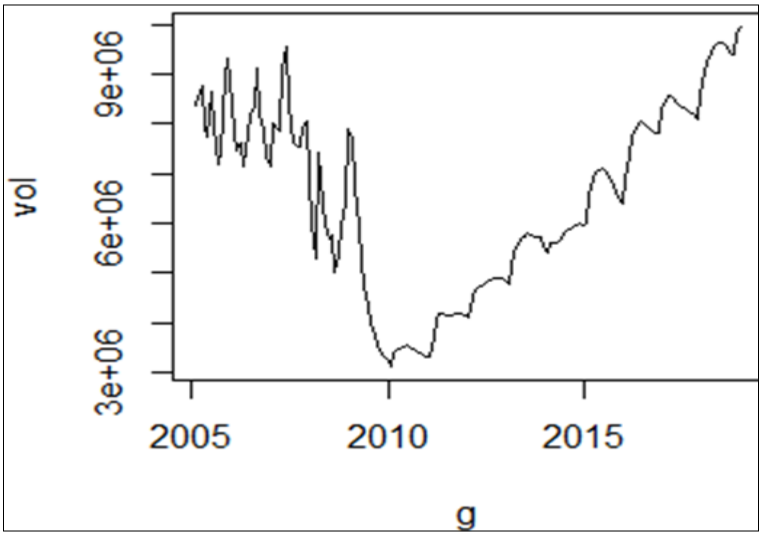

Fig 3: Forecasting performance graph of GARCH-MIDAS Model

Conclusions

The result of unit root test shows that all variables are stationary at level and after first differences at 5% level of significant. That means the variables are integrated at level and of order one i.e. 1(0) and 1(1). By implications there's no long-run relationships exist between variables. Result shows that in table (10) GARCH-MIDAS Model outperform Midas and Ardl models. So finally we concluded that GARCH-MIDAS model outperform Midas Regressions and Ardl model of Nigeria stock exchange of all shares by used of Forecasting evaluations.

References

1. Phillips PCB, Durlauf SN. Multiple Time Series Regressions with Integrated Processes. Review of Economic Studies LIII 473-495. Phillips, P.C.B., and M. Loretan 1986-91.
2. Estimating Long-Run Economic Equilibria. Review of Economic Studies. Phillips, P.C.B., and JY Park 1988;58:407-436.
3. Asymptotic Equivalence of Ordinary Least Squares and Generalized Least Squares in Regressions with Integrated Regressors. Journal of the American Statistical Association 83, 111-115.
4. Sargan JD. Wages and Prices in the United Kingdom: A Study in Econometric Methodology. In Econometric Analysis for National Economic Planning (R.E. Hart, G. Mills, J.K. Whittaker, eds.), Butterworths, London. Stock, J.H 1964-87, 25-54.
5. Asymptotic Properties of Least-Squares Estimators of Cointegration Vectors. Econometrica 55, 1035-1056.
6. Wickens MR, Breusch TS. Dynamic Specification, the Long-Run and the Estimation of Transformed Regression Models. The Economic Journal, 1988.
7. Giraitis L, Koul HL, Surgailis D. Large Sample Inference for Long Memory Processes, Imperial College Press, 2012.