Journal of Sciences and Multidisciplinary Research
Volume 9, No. 2, 2017
ISSN: 2277-0135

EMPIRICAL DISTRIBUTION AND MODELLING OF SPOT PRICES OF NIGERIA CRUDE OIL

J. F. Ojo and R.O. Olanrewaju

Department of Statistics,
University of Ibadan, Ibadan, Nigeria.
E-mail: jfunminiyiojo@yahoo.co.uk

Abstract: This paper attempts to provide a suitable time series model for daily West Texas Intermediate (WTI) Spot prices for Nigeria crude oil having determined the appropriate probability distribution for the series. The identified probability distribution for the series formed an integral part of the model. Two thousand eight hundred and eighty (2880) sample points of the daily returns of the WTI Spot prices from June 14,2005 to October 17,2016 were considered. Sixty-one (61) probability distributions were fitted to the series. Generalized Error Distribution (GED) emerged as the ideal distribution that explained the abstraction of reality for the return series due to its smallest test statistic(0.04337, 9.1145, and 158.62) for Kolmogorov-Smirnov, Anderson-Darling and Chi-Squared, respectively, from all the considered distributions goodness-of-fittests. The Generalized Auto Regressive Conditional Heteroscedasticity (GARCH) volatility model with GED as the error term (GED-GARCH)was used to capture the leverage effects. GED-GARCH(1,1)with minimum Akaike Information Criteria(AIC),Bayesian Information Criteria(BIC),and Hannan and Quinn Information Criteria(HQIC) calculated tobe (7.573, 7.583, 7.577), respectively, was found better to explain the fluctuation characterized by the WTI Spot p rices. It was discovered that 19.2% of the present variance shock(either positively or negatively) was realized in succeeding period, and that the volatility clustering was the major leverage effects expect for years between 2007 and 2008; and 2014 and 2015 that experienced sharp deviation of leverage effects. The study concludes that GED is the appropriate distribution. We recommend GED-GARCH for modeling Spot Prices of Nigeria crude oil because of its ability to capture fluctuation inherent in Spot prices of Nigeria crude oil.

Keywords: Crude Oil Spot Prices, Leverage Effect, Volatility Clustering, Generalized Error Distribution, West Texas Intermediate.

INTRODUCTION

Oil is unarguably the life blood of modern economy and it has now become the most essential commodity in the world. Hence, no nation today can survive without oil that is why Smil (2008) describes it as the "lifeblood of modern world", adding that, "without oil, there would be no globalization, no plastic, little transport, and a worldwide landscape that few would recognize". Yergin (2008) also calls it "the world's most important resource". Odalonu (2015). Crude oil price has been the determining factor to some Governments' revenues, yearly budgets, expenditures, taxes, Balance of Payments, price level, and level of economic activities that their countries are actively involved in – in both exporting and importing of crude oil. Members of the Organization of the Petroleum Exporting Countries (OPEC) like Nigeria, Saudi Arabia, Angola, Kuwait etc. have been heavily relying on the revenues and profits from crude oil production (Chioma and Oyedele, 2017).

Crude oil price movement are constantly changing as the market reacts to new information, regarding current production, consumption and inventory levels of crude oil and petroleum products. Oil prices are also affected by changes in the market's expectations of the future supply and demand balance. Depending on market conditions and sentiment, different time periods can have news and events related to either supply or demand issues as the dominant factors dictate price movements Preciado (2012).

The large fluctuations in crude oil prices in recent years have put significant pressure on the fiscal balances of both oil exporting and importing countries, Husain *et al.*, (2015). Governments of oil exporting countries generally rely heavily on revenue from oil productions and therefore, tend to suffer financially from oil price decline. Reliable forecasts of the price of oil for oil exporting countries like Nigeria, Angola, Saudi Arabia etc. is of extreme importance (Chioma and Oyedele, 2017).

The oil industry is very important to the Nigerian economy. It provides among other things the largest part of the foreign exchange earnings and total revenue needed for socio-economic and political development of the country. Thus, a small oil price change can have a large impact on the economy. The recent changes in oil prices in the global economy are so rapid and unprecedented, partly due to increased demand of oil by China and India. The bulk of Nigerian crude oil is sold unrefined. Changes in the prices of either the crude oil or any of the end products are expected to have impact on users and the nation at large. (Akomolafe and Jonathan, 2014).

Nigeria, being one of the major oil producing states, has been experiencing irregular changes in the prices of its unrefined product and this has had significant effect on her citizen's standard of living and Nigeria's economy at large (Akomolafe and Jonathan, 2014). It is an indubitable fact that the crude oil is the key determinant indexes to Nigeria's national budget. Hitherto, the Nigeria government, in collaboration with Nigerian National Petroleum Corporation (NNPC) and Budget and National Planning Office approved an oil price benchmark of $42.50 per barrel at a production assumption of 2.2 million barrels per day, for revenue calculation in its 2017 budget, in comparison with its 2016 budget, which had a price benchmark of $38/barrel at a 2.2 million barrels/day output figure. However, the government struggled to implement the 2016 budget due to, amongst other things, high volatility of crude oil prices. For instance, in January of 2016, oil prices went as low as below $25 per barrel. Since then, the prices have slowly surged to $40 and then to the present almost $50 per barrel. Given the social, political, and economic cost of volatile oil prices; different oil producing countries have tried to solve the problem of their oil price risk exposure in variety of ways (Chioma and Oyedele, 2017).

In order to make appropriate policy especially for the future, there is the need to know empirical distribution that the price of crude oil is following which will help to know the appropriate model that will capture the volatiles that must have arisen over some years in prices of the so called WTI crude oil Nigeria had sold out. Hence, this is the focus of this paper.

METHODOLOGY

Empirical Distribution

The financial time series data (daily observations) on West Texas Intermediate (WTI) Spot prices for Nigeria crude oil in Dollars per barrel from June 14, 2005 to Oct 17, 2016 were extracted from Thomson Reuter's daily oil and products database. Two thousand eight hundred and eighty (2880) sample points of the daily return of the spot prices were considered. The actual data was subjected to *Easy-Fit* software where sixty one (61) probability distributions were considered in order to know the stylized and distribution properties of the Spot prices for Nigeria crude oil. The financial returns were shown to have followed Generalized Error Distribution (GED).

The Generalized Autoregressive Conditional Heteroscedasticity (GARCH) with GED

Error Innovation

Bollerslev (1986) defined GARCH (p, q) model to be the series of returns r_t

$$r_t = \delta + \varepsilon_t = \delta + \eta_t \sqrt{h_t} \quad (1)$$

$$h_t = \alpha_0 + \sum_{i=1}^{p} \alpha_i \varepsilon_{t-1}^2 + \sum_{i=1}^{q} \beta_i h_{t-1} \quad (2)$$

such that $\alpha_0, \alpha_i \; and \; \beta_i \geq 0$ to ensure a positive conditional variance, with the error innovation of an independently and identically distributed $(0, \eta_t)$ times the square root of the conditional variance.

In literature, the empirical distribution of the financial returns (e.g. price, exchange rates, stocks, etc.) of the GARCH model is found to be affected by leptokurtic or platykurtic (Marcucci, 2005), so to such model, a robust error distribution needed to be substituted for the error term (ε_t). Hence, Generalized Error Distribution (GED) based on the empirical distribution of Spot price of crude oil was substituted. Diongue *et al.*(2008) defined the Probability Density Function (PDF) of the error term (ε_t) that takes the form

$$f(\varepsilon_t) = \frac{v \exp\left[-(1/2) \left| \frac{\varepsilon_t}{\lambda \sqrt{h_t}} \right|^v \right]}{\sqrt{h_t}\, \lambda^{2(1+v)} \Gamma(1/v)} \qquad -\infty < \varepsilon_t < \infty \quad (3)$$

such that $\lambda = \sqrt{2^{-2/v} \frac{\Gamma(1/v)}{\Gamma(3/v)}}$, $\Gamma(.)$ is the Gamma Euler function, $0 < v < \infty$ is tail thickness parameter. The GED becomes the standard normal distribution when $v = 2$, while the $v < 2 \; and \; v > 2$ the distribution has thicker and thinner tails than normal respectively. Also, it is to be noted that as $v \to \infty$, GED becomes uniform distribution with $\left[-\sqrt{3}, \sqrt{3}\right]$.

Parameter Estimation

Given a sample of $y_1, y_2, \cdots, y_n$ the conditional likelihood function for GED with v-degree of freedom.

$$\ell(w) = f(w / y_1, \cdots, y_n) = \prod_{t=1}^{n} f(w_i / y_1, \cdots, y_n) \quad (4)$$

for $w\{\alpha_0, \alpha_i, \beta_i, \eta_t\}$ such that $\alpha_i + \beta_i < 1$. The log-likelihood function for GED is $g(w) = \log \ell(w)$

$$= n\left\{ \log\left(\frac{v}{\lambda_v} \right) - (1 + 1/v)\log(2) - \log\Gamma(\frac{1}{2}) - \frac{1}{2} \sum_{t=1}^{n} \left(\log(h_t^2(w)) + h_t^{-v}(w) \left| \frac{\varepsilon_t}{\lambda_v} \right|^v \right) \right\} \quad (5)$$

To get the score functions and the Hessian matrix for the parameter

$$\frac{\partial g(w)}{\partial w} = \frac{1}{2} \sum_{t=1}^{n} \frac{v}{|\lambda_v|} \left(\frac{\varepsilon_t}{h_t(w)} \right)^v \frac{1}{\varepsilon_t} . \frac{\partial \varepsilon_t}{\partial w} + \frac{1}{2} \sum_{t=1}^{n} \frac{1}{h_t^2(w)} \left[\frac{1}{2} \frac{1}{|\lambda_v|^v} \left(\frac{\varepsilon_t}{h_t(w)} \right)^v - 1 \right] \frac{\partial h_t^2(w)}{\partial w} \quad (6)$$

$$\frac{\partial^2 g(w)}{\partial w \partial w^{'}} = \frac{-v}{|\lambda_v|^v}\left(\frac{v-3}{2}\right)\sum_{t=1}^{n}\frac{1}{h_t^2(w)}\left(\frac{\varepsilon_t^2}{h_t^2(w)}\right)^{\frac{v}{2}-1}\frac{\partial\varepsilon_t\partial\varepsilon_t}{h_t^2(w)} - \frac{1}{2}\frac{v}{|\lambda_v|^v}\sum_{t=1}^{n}\frac{\varepsilon_t}{h_t^2(w)}\left(\frac{\varepsilon_t^2}{h_t^2(w)}\right)^{\frac{v}{2}-1}.\left[1+\left(\frac{v}{2}-1\right)\frac{\varepsilon_t^2}{h_t^2(w)}\right] - \frac{1}{2}\frac{v}{|\lambda_v|^v}\sum_{t=1}^{n}\frac{\varepsilon_t}{h_t^2(w)}\left[1+\left(\frac{v}{2}-1\right)\frac{\varepsilon_t^2}{h_t^2(w)}\right] - \frac{1}{2}\sum_{t=1}^{n}\left[\frac{1}{4}\frac{v(v+2)}{|\lambda_v|^v}\left(\frac{\varepsilon_t}{h_t(w)}\right)^v - 1\right]\frac{1}{h_t^4(w)}\frac{\partial h_t^2(w)\partial h_t^2(w)}{h_t^4 \partial w \partial w^{'}} \quad (7)$$

Equation (6) and (7) requires the computation of

$$\frac{\partial h_t^2}{\partial w} = \left(1,\, \varepsilon_{t-i}^2,\, h_{t-i}^2,\, h_{-i}\varepsilon_{t-i}\right) + 2\alpha_i\varepsilon_{t-1}\frac{\partial\varepsilon_{t-i}}{\partial w} + \beta_i\frac{\partial h_{t-i}^2}{\partial w} + \eta_t\left(h_{t-i}\frac{\partial\varepsilon_{t-i}}{\partial w} + \frac{1}{2}\frac{\varepsilon_{t-i}}{h_{t-i}}\frac{\partial h_{t-i}^2}{\partial w}\right) \quad (8)$$

$w = \{\alpha_0,\, \alpha_i, \beta_i, \eta_t\}$ cannot be solved by normal equation of equating the derivation of equation (8) of $\alpha_0, \alpha_i, \beta_i \& \eta_t$ to zero. An iterative procedure of Weighted Least Square (WLS) would be ideal.

Findings and Discussion

Empirical Distribution

In the table below, the best twenty (20) probability distributions were reported, having subjected sixty one (61) probability distributions to easy fit software. The analysis was based on distributional ranking via Kolmogorov-Smirnov, Anderson-Darling and Chi-Squared Test Statistic.

Table 1: The Distributional Ranking via Kolmogorov, Anderson and Chi-Squared

S/N	Probability Distributions	Kolmogorov-Smirnov		Anderson-Darling		Chi-Squared	
		Statistic	Rank	Statistic	Rank	Statistic	Rank
1	Beta	0.06133	10	14.902	6	279.75	9
2	Burr	0.05883	7	14.631	5	271.36	8
3	Burr (4P)	0.06809	15	17.781	10	338.16	14
4	Cauchy	0.10974	43	69.538	41	881.28	42
5	Chi-Squared	0.20587	49	655.94	55	2848.9	49
6	Chi-Squared (2P)	0.0873	28	24.556	21	362.68	18
7	Dagum	0.06385	11	18.189	11	203.27	4
8	Dagum (4P)	0.46119	57	1506.9	58	3619.4	50
9	Erlang	0.15629	47	112.68	44	684.84	36
10	Erlang (3P)	0.0696	18	20.274	17	379.71	20
11	Generalized Error	0.04337	1	9.1145	1	158.62	1
12	Error Function	0.92522	59	15551.0	59	1.9066E+5	56
13	Exponential	0.36453	54	648.41	54	5245.6	52
14	Exponential (2P)	0.25427	50	370.55	50	2384.4	48
15	Fatigue Life	0.09409	36	41.612	33	627.79	34
16	Fatigue Life (3P)	0.07081	19	19.288	15	344.62	15
17	Frechet	0.13689	46	126.85	45	1472.1	44
18	Frechet (3P)	0.08495	27	39.244	31	639.86	35
19	Gamma	0.08901	29	34.885	28	555.24	31
20	Gamma (3P)	0.07785	23	22.125	19	391.41	22

Table 2: The Goodness-of-fit Details for the Generalized Error Distribution

GED σ=22.559 μ=77.167					
Kolmogorov-Smirnov					
Sample Size Statistic P-Value Rank	2880 0.04337 4.1523E-5 1				
Sig. Level	0.2	0.1	0.05	0.02	0.01
Critical Value	0.02007	0.02288	0.02541	0.0284	0.03048
Reject H_0	Yes	Yes	Yes	Yes	Yes
Anderson-Darling					
Sample Size Statistic Rank	2880 9.1145 1				
Sig. Level	0.2	0.1	0.05	0.02	0.01
Critical Value	1.3749	1.9286	2.5018	3.2892	3.9074
Reject H_0	Yes	Yes	Yes	Yes	Yes
Chi-Squared					
Deg. of freedom Statistic P-Value Rank	11 158.62 0 1				
Sig. Level	0.2	0.1	0.05	0.02	0.01
Critical Value	14.631	17.275	19.675	22.618	24.725
Reject H_0	Yes	Yes	Yes	Yes	Yes

It was noted from table 1 that out of the probability distributions fitted; the Generalized Error Distribution (GED) was ranked first and has the distribution with the smallest test statistic of 0.04337, 9.1145 and 158.62 for Kolmogorov-Smirnov, Anderson-Darling and Chi-Squared test, respectively to ideal explained the abstraction of the real system (actual data). Table 2 expounded that the null hypotheses of "WTI Spot prices do not followed Generalized Error Distribution" subjected to each goodness-of-fit test was rejected at 20%, 10%, 5%, 2% and 1% level of significant respectively, because the critical values for each level of significant aforesaid under Kolmogorov-Smirnov, Anderson-Darling and Chi-Squared test was less than their test statistic.

Time Plot of WTI Spot Price of Nigeria Crude Oil

Figure 1: The Time Plot of the WTI Spot Prices Series

It is obvious from the time plot in figure 1 that there is proclivity of large variations in the WTI spot price returns that must have come-off of the tracks from large fluctuations. The proclivity of large variations is another word for volatility, which can be a clustering or a pooling volatility. Furthermore, wide sharp margin volatility was between 2006 and 2007; and between 2014 and 2015, such that a clustering volatility was experienced among the remaining years.

Summary Statistics

Table 3: Summary Statistics of the Daily WTI Spot Price of the Sampled Period

Series	Mean	Std. error	skewness	Kurtosis	Shapiro test	JB test
WTI Spot prices	77.167	22.559	0.0293	-0.6205	0.9815 (< 2.2e-16)	46.3520 (0.0001)

The WTI spot prices sampled are positively skewed whereas the excess kurtosis indicated a thicker-tailed distribution (playkurtic) that is rightly skewed. This reveals that the WTI spot price did not follow a normal distribution that is, the prices over the 12years period are not normally distributed, and the non-normality was also confirmed Shapiro-Wilk test and Jarque Bera test.

Volatility Model for the WTI Spot Prices

Box and Jenkins (1970) proposed that AR (0) is nothing but the white noise (ε_t), such that when it is assumed that there is no dependency between terms then the error term (ε_t) will contribute to the process. In other words, the serial correlation coefficients between the model residual series should not be statistically different from zero. Engle (1982) proposed the Lagrange Multiplier (LM) test for conducting the effect of AR (0) model residual series in the presence of volatility. In line with Aghayev and Rizvanoghlu (2014) that used the same test to know the distribution of the white noise (ε_t), an ARCH LM test was conducted in table 4 for GED.

Table 4: Autoregressive Conditional Heteroscedasticity Lagrange Multiplier for GED Parameter

Return Series	Parameter	90% quantile	95% quantile	99% quantile
WTI Spot price	1.4072	0.0231	0.0501	0.0834

Since the GED parameter for the Spot price series in table 4 is $1.4072 < 2$, this buttress the point made that the return series is not normally distributed with thicker-tailed distribution that might be rightly skewed and that the ideal distribution for the return series is GED. The 90%, 95% and 99% quantiles for the WTI spot price return series based on the Probability Density Function (PDF) confirmed the smaller values of each of the quantile than those of the standard normal distribution with 1.645, 1.96 and 2.576, respectively, which suggested the obvious thin tail of the Spot prices.

Table 5: Information Criterion (IC) Statistics for the Best Order

Models	AIC	BIC	SIC	HQIC
GED-GARCH(1, 1)	7.573002	7.583362	7.572996	7.576737
GED-GARCH(1, 2)	7.573902	7.586334	7.573893	7.578383
GED-GARCH(2, 1)	7.573975	7.586406	7.573966	7.578456
GED-GARCH(2, 2)	7.632572	7.647076	7.632560	7.637800

From table 5 it was observed that the Information Criteria (IC) increases as the both the autoregressive and conditional heteroscedasticity effect increases. It was noted that GED-GARCH (1, 1) will be the ideal system to explain the fluctuation characterized by WTI spot prices

Table 6: Coefficients for the GED-GARCH (1, 1) Variance Equation

Parameter	Estimate	t-value	Pr(>|t|)
μ	71.32106 (0.21263)	335.422	< 2e-16 ***
α_0	4.46243 (0.38031)	11.734	< 2e-16 ***
α	0.43288 (0.01999)	21.656	< 2e-16 ***
β	0.19233 (0.03492)	5.507	3.64e-08 ***
λ	10.00000 (0.31732)	31.514	< 2e-16 ***
v	1.4072		
Log-Likelihood	-10896.34		
AIC	7.573002		

Values asterisk and bracket in table 6 are the significant level and standard error.

$h_t = 4.4624 + 0.4329\varepsilon_{t-1}^2 + 0.19233h_{t-1}$

Since $\alpha_1 + \beta_1 = 0.43288 + 0.19233 = 0.62521 < 1$, it implies that there is presence of conditional variance (volatility effect) in the return series and it is positive. It is obvious that the effect of (ε_t) that follows GED on h_t is $\alpha + v = 0.43288 + 1.4072 = 1.84008$. The coefficient of h_{t-1} = 0.19233, means that 19.2% of the present variance shock (either positively or negatively) was realized in succeeding period. In other words, the volatility shock is pretty faster and the volatility clustering is apparently observed, that is leverage effects of the series are close expect for great gap leverage effect that was experienced between 2007 and 2008; and between 2014 and 2015.

CONCLUSION

The paper had provided a suitable time series model for daily West Texas Intermediate (WTI)Spot prices for Nigeria crude oil having determined the appropriate distribution for the Spot prices of Nigeria Crude Oil. GED was discovered as the ideal distribution for the data. The time plot revealed the high level of variability within the data such that the error term (ε_t) of the GARCH model suitable for estimating the volatility was substituted for GED. It was noted that 19.2% of the present variance shock was realized in succeeding period, and that the volatility clustering was the major leverage effects expect for years between 2007 and 2008; and 2014 and 2015 that experienced sharp deviation of leverage effects.

REFERENCES

Akomolafe, K.J. and Jonathan, D. D. (2014): Oil Price Dynamics and the Nigerian Stock Market: An Industry Level Analysis. *International Journal of Economics, Finance and Management.* Vol.3 (6), pages 308-316.

Bollerslev, T. (1986).Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics.*Vol.31, pages307-327.

Box, G. E., Jenkins, G., (1970).*Time Series Analysis:* Forecasting and control. Holden-Dayiv, San Francisco.

Chioma, N. N. and Oyedele, A.A. (2017): Simulation and Hedging Oil Price with Geometric Brownian Motion and Single-Step Binomial Price Model. *European Journal of Business and Management*, Vol.9 (9), pages 68-815.

Diongue, A.K., Guegan, D. and Wolf, R.C. (2008). Exact Maximum Likelihood Estimation for the BL-GARCH Model under Elliptical Distributed Innovations. *CES WORKING PAPER.*

Engle, R. F., (1982). Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Nation. *Econometrica: Journal of the Econometric Society*, pages987-1007.

Husain, A. M., Arezki, R., Breuer, P., Haksar, V., Helbling, T., Medas, P., Sommer, M., (2015), "Global Implications of Lower Oil Prices", *IMF Staff Discussion Note.* Available at https://www.imf.org/external/pubs/ft/sdn/2015/sdn1515.pdf.

James Preciado (2012). Implications of Changing Correlations between WTI and Other Commodities, Asset Classes, and Implied Volatility. *Working Papers – Independent Statistics and Analysis.*

Available at https://www.eia.gov/workingpapers/pdf/correlation_paper.pdf

Odalonu Happy Boris (2015).The Upsurge of Oil Theft and Illegal Bunkering in the Niger Delta Region of Nigeria: Is There a Way Out? *Mediterranean Journal of Social Science.* Vol. 6 No. 3 s2,pages 563-573.

Smil, V. (2008).*Oil: A Beginner's Guide.* Oxford: One world Publications.

Yergin, D. (2008).*The Prize: The Epic Quest for Oil Money and Power.* New York: Free Press

Reference to this paper should be made as follows: J. F. Ojo and R.O. Olanrewaju (2017). Empirical Distribution and Modelling of Spot Prices of Nigeria Crude oil. *J. of Sciences and Multidisciplinary Research,* Vol. 9, No. 2, Pp. 66-77