Journal of Physical Sciences and Innov... © 2012 Cenresin Publications www.cenresinpub.org ISSN 2277-0119 ...Volume 4, September 2012

SEASONAL BOX-JENKINS MODELLING OF NIGERIAN MONTHLY CRUDE OIL EXPORTS

Ette Harrison Etuk
Department of Mathematics/Computer Science,
Rivers State University of Science and Technology, Nigeria
Email: ettetuk@yahoo.com

ABSTRACT

The time plot of the original series NCOE reveals a negative secular trend between 2006 and 2009 after which the trend tends to be increasing up to 2011. Seasonality is not so evident. A twelve-month (i.e. seasonal) differencing was done to yield the series SDNCOE with a generally positive trend and not so regular seasonality. Further nonseasonal differencing yielded a series DSDNCOE with no trend. Its correlogram reveals a seasonality of order 12, a seasonal moving average component of order 1 and an autocorrelation structure of a $(0, 1, 1)x(0, 1, 1)_{12}$ model. Therefore the model was proposed and fitted to the series. Diagnostic checking results show that the model is adequate.

Keywords: *Seasonal Models, Box-Jenkins methodology, Crude Oil Exports, Nigeria.*

INTRODUCTION

A time series may be defined as a sequence of data collected at times often equally spaced. Such a sequence usually exhibits a tendency of having correlated neighbouring values. This tendency is called *autocorrelation.* A time series is said to be *stationary* if its mean and variance are invariant under a translation along the time axis. Moreover, the autocorrelation is a function of the time lag separating the correlated values. This is called the *autocorrelation function,* ACF. Nonstationary nature easily shows up in the time-plot and the plot of the ACF. Besides the requirement of stationarity is that of *invertibility* which refers to the property whereby corresponding to a model there is a unique covariance structure (Priestley, 1981).

A stationary time series $\{X_t\}$ is said to follow an autoregressive moving average model of order p and q (denoted by ARMA(p, q)) if it satisfies the following difference equation

$$X_t - \alpha_1 X_{t-1} - \alpha_2 X_{t-2} - \ldots - \alpha_p X_{t-p} = \varepsilon_t + \beta_1 \varepsilon_{t-1} + \beta_2 \varepsilon_{t-2} + \ldots + \beta_q \varepsilon_{t-q} \quad (1)$$

Or

$$(1 - \alpha_1 L - \alpha_2 L^2 - \ldots - \alpha_p L^p)X_t = (1 - \beta_1 L - \beta_2 L^2 - \ldots - \beta_q L^q)\varepsilon_t \quad (2)$$

Or

$$A(L)X_t = B(L)\varepsilon_t \quad (3)$$

where L is the backshift operator defined by $L^k X_t = X_{t-k}$ and the α's and the β's are constants such that for stationarity and invertibility A(L) and B(L) have zeroes outside the unit circle respectively. The sequence $\{\varepsilon_t\}$ called a *white noise process* involves random variables that are uncorrelated and have zero mean and constant variance.

If q = 0, model (1) is called an *autoregressive model of order* p (denoted by AR(p)). If, however, p = 0, the model is called a *moving average model of order q* (denoted by MA(q)). An AR(p) model may be more specifically written as

$$X_t - \alpha_{p,1}X_{t-1} - \alpha_{p,2}X_{t-2} - ... - \alpha_{p,p}X_{t-p} = \varepsilon_t + \beta_1\varepsilon_{t-1} + \beta_2\varepsilon_{t-2} + ... + \beta_q\varepsilon_{t-q}$$

The sequence of the last coefficients $\{\alpha_{ii}\}$ is called the *partial autocorrelation function* (PACF). While the PACF of an AR(p) cuts off at lag p, the ACF of an MA(q) cuts off at lag q. This serves as a guide for preliminary model identification. For a nonstationary series, Box and Jenkins(1976) proposed that the nonstationarity could be got rid of by differencing of the series to an appropriate order d. The d^{th} difference of a series $\{X_t\}$ is the series $\{\nabla^d X_t\}$ where $\nabla = 1 - L$. Differencing is usually done on the series progressively from order 1 until stationarity is attained. Suppose stationarity is attained at the d^{th} difference. If an ARMA(p, q) model is fitted to the differenced series, the original series $\{X_t\}$ is said to follow an *autoregressive integrated moving average model of order (p, d, q)*, denoted by ARIMA(p, d, q). Seasonality refers to a tendency for a time series to show periodic behaviour after regular intervals of time. A time series $\{X_t\}$ is said to follow a multiplicative $(p, d, q)x(P, D, Q)_s$ *seasonal autoregressive integrated moving average model* if

$$A(L)\Phi(L^s)\nabla^d\nabla^D_s X_t = B(L)\Theta(L^s)\varepsilon_t \quad (4)$$

where

$$\Phi(L) = 1 + \phi_1 L + ... + \phi_P L^P \quad (5)$$

$$\Theta(L) = 1 + \theta_1 L + ... + \theta_Q L^Q \quad (6)$$

and the coefficients ϕ's and θ's are such that the zeroes of the polynomials (5) and (6) are outside of the unit circle for stationarity and invertibility respectively. Seasonality is often evident from the time plot as a periodic movement. Moreover the *correlogram* (i.e. the ACF plot) of a seasonal series is such that there is a spike at the seasonal lag. Seasonal ARIMA models are extensively discussed in Box and Jenkins (1976) and Madsen (2008), to mention but a few publications.

MATERIALS AND METHODS

The data for this work is from the Central Bank of Nigeria publication on the website www.cenbank.org . The data are the seventy two monthly crude oil exports expressed in million barrels per day for the years 2006 to 2011 and published under the Data and Statistics heading.

Determination of the orders p, d, q, P, D and Q

Knowledge of the theoritical properties of the models provides basis for the determination of the orders. For instance, the cutting-off of the ACF is indicative of an MA model of order equal to the cut-off lag q whereas the cutting-off of the PACF indicates an AR model of order equal to the cut-off lag p. A negative spike at the seasonal lag indicates seasonality as well as the involvement of a seasonal MA component of order 1. A positive spike at the seasonal lag shows seasonality that involves a seasonal AR component of order 1. The nonstationary series shall be differenced once seasonally. That is, D = 1. If there is no stationarity the series shall be differenced once, nonseasonally. That is, d = 1. It is noteworthy that an autocorrelation that is within $\pm 2/\sqrt{n}$, where n is the sample size, is not considered statistically significant.

Model Estimation

The involvement of items of a white noise process in an ARIMA model calls for nonlinear optimization techniques for model estimation. An initial estimate is usually made and employed in an iterative convergent process until an optimum estimate results within the stipulated limit of accuracy. The optimization criterion adopted could be the least sum of squares approach, the maximum likelihood approach or the maximum entropy approach, to mention only a few. However linear optimization methods have been proposed and used for pure AR and pure MA models (See for example Box and Jenkins(1976) and Oyetunji(1985)). Attempts have also been made to apply linear optimization methods to estimate mixed ARMA models (See for example Etuk(1987, 1998)). In this work Reviews software which employs the least squares approach shall be used.

Diagnostic Checking: To check for goodness-of-fit of the model to the data some residual analysis shall be done. Under the assumption of model adequacy the residuals should be uncorrelated, have zero mean and constant variance and follow a normal distribution.

RESULTS AND DISCUSSION

The time plot of the series NCOE in Figure 1 shows an downward secular trend from 2006 to 2009 and an upward one from then to 2011. Seasonality is not obvious. Seasonal differencing once yielded SDNCOE with an overall upward trend as shown in Figure 2. Nonseasonal differencing of SDNCOE yielded DSDNCOE with no trend (See Figure 3) but with correlogram showing a negative spike at lag 12, depicting seasonality of order 12 and a seasonal MA component of order 1. A spike at lag 1 in the ACF suggests the involvement of a nonseasonal MA component of order one. A (0, 1, 1)x(0, 1, 1)$_{12}$ seasonal model

$$DSDNCOE_t = \beta_1\varepsilon_{t-1} + \beta_{12}\varepsilon_{t-12} + \beta_{13}\varepsilon_{t-13} \quad (7)$$

Is therefore proposed. By Eviews, the estimation of the model yielded the results as summarized in Table 1. It is noteworthy that of the coefficients only β_{13} = 0.2015 is not statistically significant, being less than twice its standard error. The adequacy of the model is not in doubt given the close agreement of the fitted model and the data (See Figure 5), and the normal distribution of the residuals with zero mean (See Figure 6). Moreover, the correlogram of the residuals in Figure 7 is such that none of the 24 autocorrelations is statistically significant.

CONCLUSION

It may be concluded that monthly Nigerian Crude Oil Exports follow a $(0, 1, 1)x(0, 1, 1)_{12}$ model. This has been shown to be adequate.

REFERENCES:

Box, G. E. P. and Jenkins, G. M. (1976). Time Series Analysis, Forecasting and Control, Holden-Day, San Francisco.

Etuk, E. H. (1987). *On the Selection of Autoregressive Moving Average Models.* An unpublished Ph. D. Thesis, Department of Statistics, University of Ibadan, Nigeria.

Etuk, E. H. (1998). *An Autoregressive Integrated Moving Average (ARIMA) Simulation Model: A Case Study*. Discovery and Innovation, Volume 10, Nos. 1 & 2: pp. 23 – 26.

Madsen, H. (2008). Time Series Analysis, Chapman & Hall/CRC, London.

Oyetunji, O. B. (1985). *Inverse Autocorrelations and Moving Average Time Series Modelling.* Journal of Official Statistics, Volume 1: pp. 315 – 322.

Priestley, M. B. (1981). Spectral Analysis and Time Series. Academic Press, London.

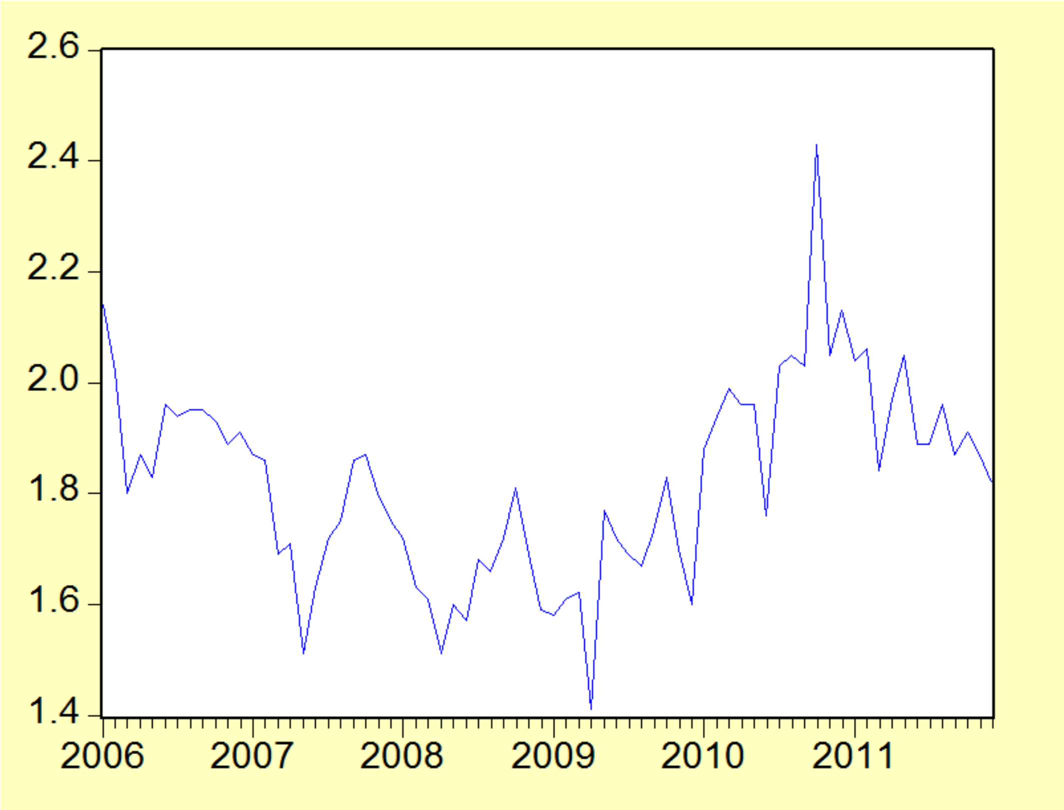


FIGURE 1: NCOE

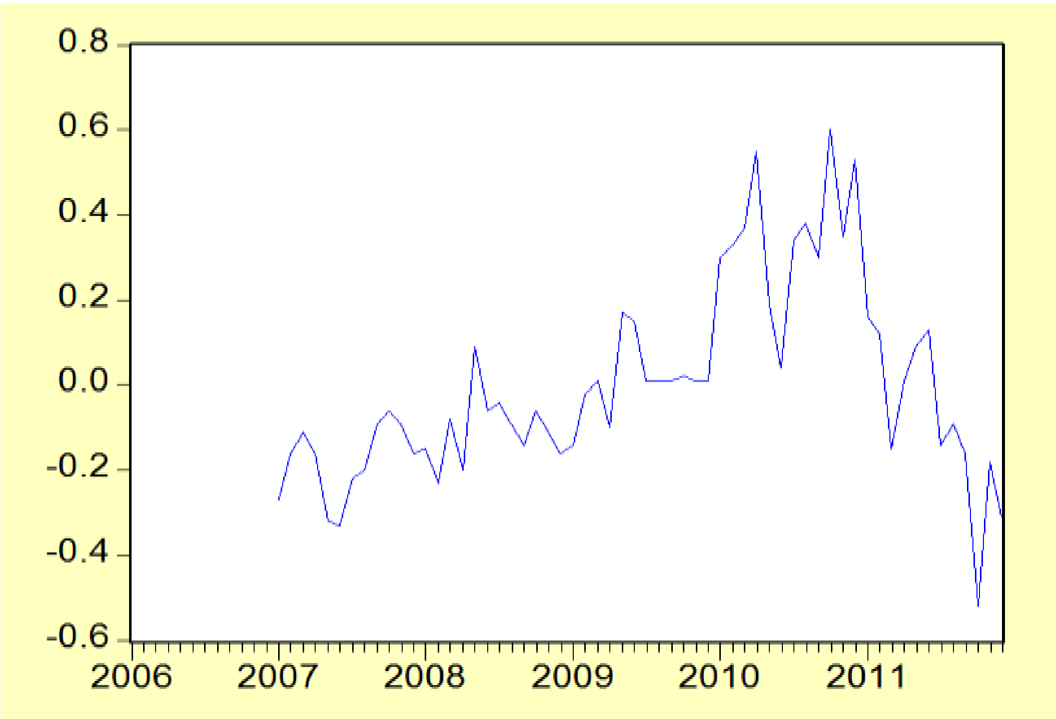


FIGURE 2: SDNCOE

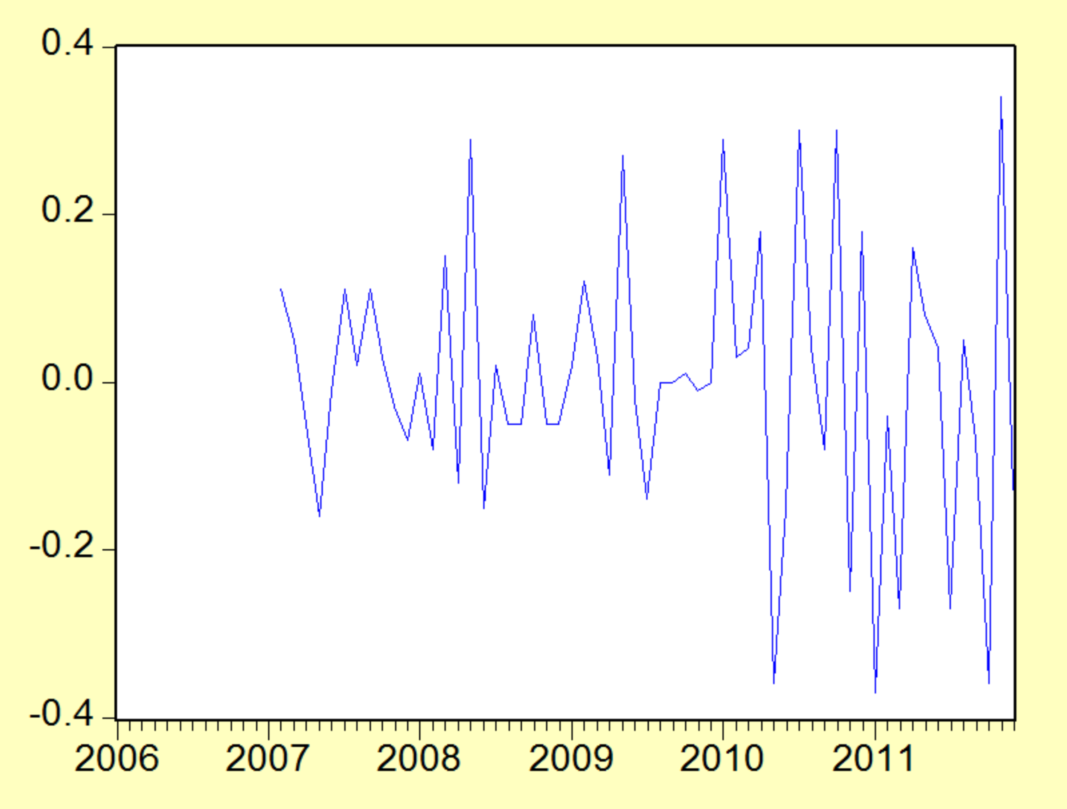


FIGURE 3: DSDNCOE

© 2012 Cenresin Publications
www.cenresinpub.org
ISSN 2277-0119

Date: 09/22/12 Time: 14:05
Sample: 2006:01 2011:12
Included observations: 59

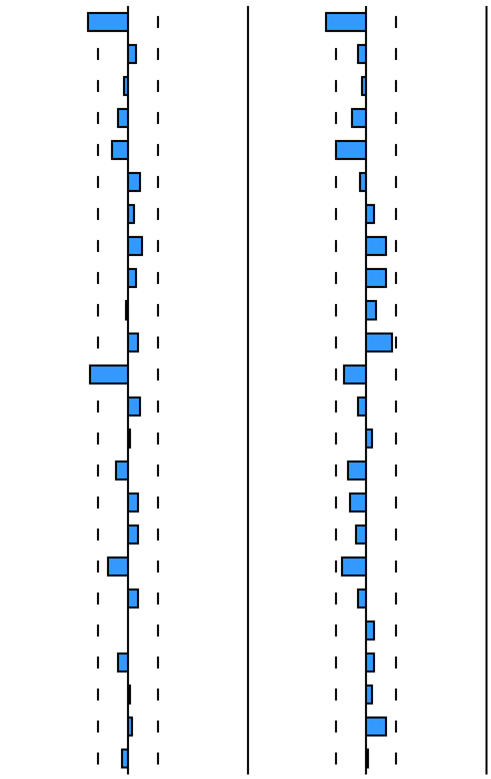
Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	-0.358	-0.358	7.9610	0.005
		2	0.068	-0.069	8.2542	0.016
		3	-0.026	-0.028	8.2974	0.040
		4	-0.089	-0.119	8.8131	0.066
		5	-0.146	-0.258	10.231	0.069
		6	0.109	-0.053	11.034	0.087
		7	0.047	0.078	11.187	0.131
		8	0.118	0.176	12.171	0.144
		9	0.072	0.178	12.545	0.184
		10	-0.018	0.088	12.570	0.249
		11	0.088	0.224	13.145	0.284
		12	-0.335	-0.193	21.738	0.041
		13	0.111	-0.063	22.705	0.045
		14	0.021	0.056	22.741	0.065
		15	-0.106	-0.153	23.654	0.071
		16	0.098	-0.132	24.451	0.080
		17	0.085	-0.092	25.064	0.093
		18	-0.172	-0.218	27.675	0.067
		19	0.091	-0.062	28.427	0.076
		20	0.004	0.078	28.428	0.100
		21	-0.084	0.063	29.091	0.112
		22	0.013	0.044	29.107	0.142
		23	0.040	0.167	29.267	0.172
		24	-0.057	0.018	29.598	0.198

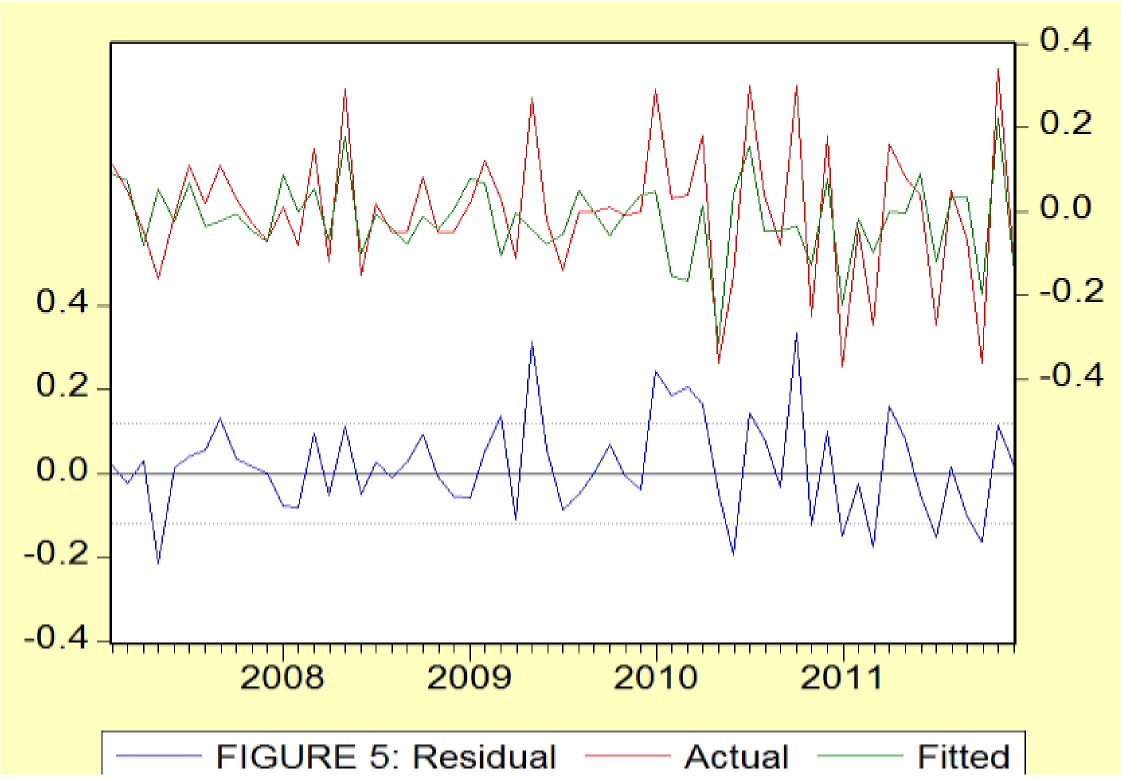
FIGURE 4: CORRELOGRAM OF DSDNCOE

Dependent Variable: DSDNCOE
Method: Least Squares
Date: 09/22/12 Time: 14:16
Sample(adjusted): 2007:02 2011:12
Included observations: 59 after adjusting endpoints
Convergence not achieved after 100 iterations
Backcast: 2006:01 2007:01

Variable	Coefficient	Std. Error	t-Statistic	Prob.
MA(1)	-0.431922	0.111051	-3.889394	0.0003
MA(12)	-0.702029	0.097761	-7.181099	0.0000
MA(13)	0.201479	0.112522	1.790585	0.0788

R-squared	0.462872	Mean dependent var	-0.000678
Adjusted R-squared	0.443689	S.D. dependent var	0.160987
S.E. of regression	0.120074	Akaike info criterion	-1.351907
Sum squared resid	0.807396	Schwarz criterion	-1.246269
Log likelihood	42.88124	F-statistic	24.12912
Durbin-Watson stat	2.078798	Prob(F-statistic)	0.000000

Inverted MA Roots	.99	.86+.48i	.86 -.48i	.50 -.84i
	.50+.84i	.29	.01+.97i	.01 -.97i
	-.48+.84i	-.48 -.84i	-.83+.48i	-.83 -.48i
	-.96			


0.4
0.2
0.0
-0.2
-0.4
2008
2009
2010
2011
FIGURE 5: Residual
Actual
Fitted

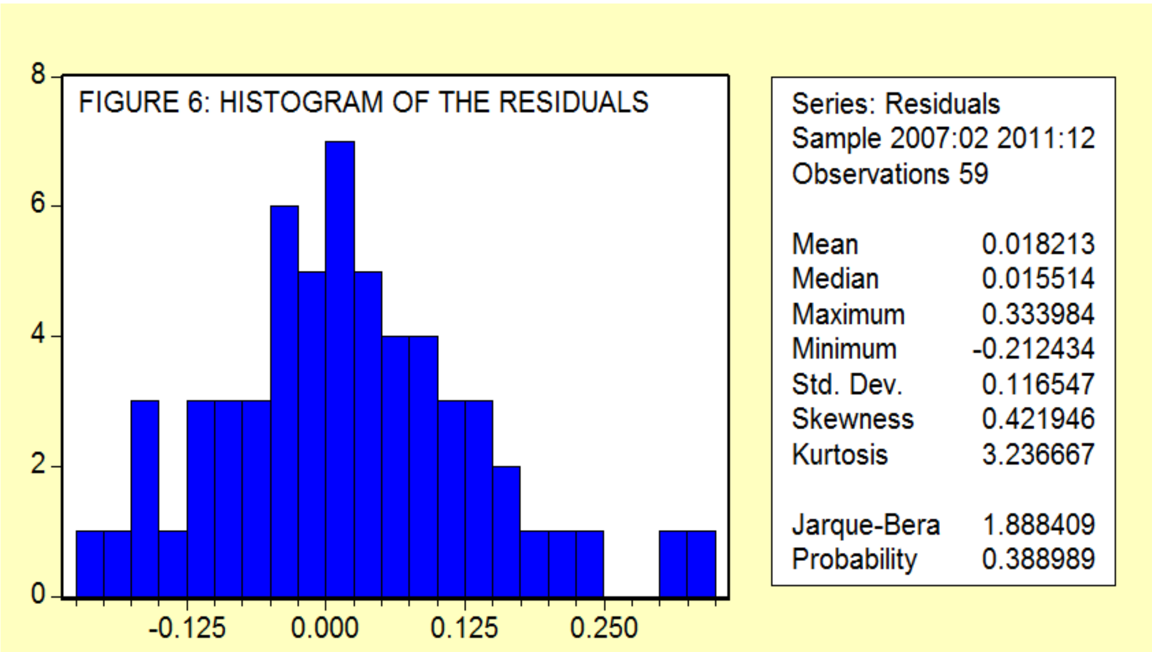
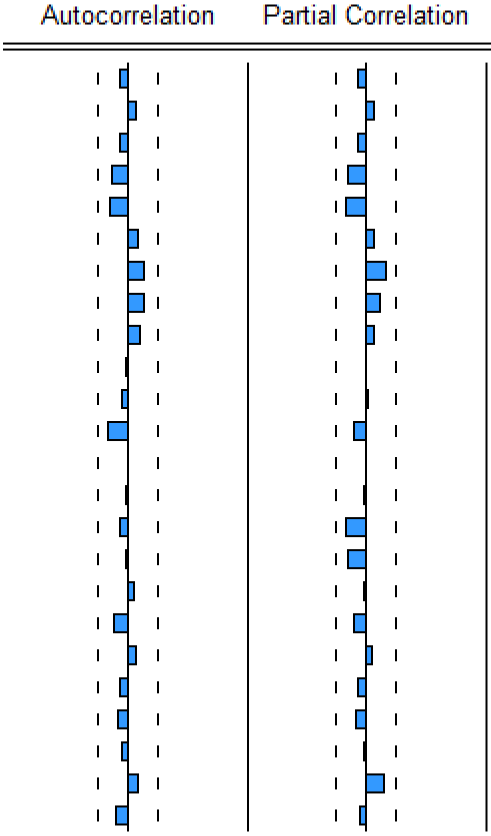


FIGURE 6: HISTOGRAM OF THE RESIDUALS
8
6
4
2
0
-0.125
0.000
0.125
0.250
Series: Residuals
Sample 2007:02 2011:12
Observations 59
Mean 0.018213
Median 0.015514
Maximum 0.333984
Minimum -0.212434
Std. Dev. 0.116547
Skewness 0.421946
Kurtosis 3.236667
Jarque-Bera 1.888409
Probability 0.388989

© 2012 Cenresin Publications
www.cenresinpub.org
ISSN 2277-0119

Date: 09/27/12 Time: 20:39
Sample: 2007:02 2011:12
Included observations: 59
Q-statistic probabilities adjusted for 3 ARMA term(s)



	AC	PAC	Q-Stat	Prob
1	-0.065	-0.065	0.2640	
2	0.064	0.060	0.5199	
3	-0.068	-0.061	0.8203	
4	-0.139	-0.152	2.0814	0.149
5	-0.151	-0.167	3.6051	0.165
6	0.085	0.079	4.0968	0.251
7	0.146	0.171	5.5775	0.233
8	0.139	0.124	6.9402	0.225
9	0.103	0.073	7.7100	0.260
10	-0.008	0.003	7.7153	0.358
11	-0.046	0.024	7.8711	0.446
12	-0.178	-0.108	10.291	0.327
13	-0.002	0.005	10.291	0.415
14	-0.012	-0.015	10.302	0.503
15	-0.074	-0.164	10.746	0.551
16	-0.020	-0.152	10.781	0.629
17	0.049	-0.021	10.984	0.687
18	-0.121	-0.110	12.275	0.658
19	0.062	0.051	12.620	0.700
20	-0.074	-0.059	13.121	0.728
21	-0.085	-0.094	13.807	0.742
22	-0.044	-0.024	13.991	0.784
23	0.094	0.158	14.872	0.784
24	-0.115	-0.051	16.220	0.757

FIGURE 7: CORRELOGRAM OF THE RESIDUALS