Time Series Modelling of Monthly WTI Crude Oil Returns

Derek LAM

Lincoln College

University of Oxford

A thesis submitted in partial fulfilment for the degree of Master of Science in Mathematical and Computational Finance

Trinity 2013

Acknolwedgements

First of all, I would like to express my sincere gratitude to my thesis supervisor, Dr. Patrick McSharry, for his valuable guidances and advices. I would also like to thank all of the lecturers I have met this year, for without them, I would never be able to learn so much this year. Lastly, I would like to thank my parents for their unlimited support all along.

Abstract

This paper examines the dynamics of the monthly WTI crude oil return for the past two decades. The data are divided into two ten-year periods, and we explore with two approaches. We first build univariate time series models using the Box-Jenkins methodology. Techniques such as stationarity tests and autocorrelation plots are used to determine the orders of the final ARIMA model. GARCH and APARCH are also used to model residuals. Then, we build regression models based on eight explanatory variables. They are consumption, production, ending stock, net import, refinery utilisation rate, U.S. interest rate, NYMEX oil futures contract 4 and S&P 500 index. Stepwise AIC method is employed to determine the optimal variables to be included. Multicollinearity is not evident in the reduced models. Residual analysis suggests that the assumptions of linear regression are not violated. Lastly, the forecasting powers of the models are compared. GARCH and APARCH perform the best in terms of forecasting accuracy, with APARCH performing the best in a turbulent market.

Keywords: Linear regression, ARIMA, GARCH, APARCH, time series forecasting, residual analysis

Contents

1 Introduction

Crude oil is one of the most important commodities in the world. Its applications are ubiquitous in daily life, from making detergents and plastic bags to fueling cars and ships. Despite being non-renewable, the world consumes crude oil every single minute as it is difficult to find an alternative source that can parallel its performances. With such a uniqueness in the world, it is vital for us to develop a better understanding of its price dynamics, so that the myriad industries which consume or supply oils can make more informed decisions.

In this paper, we will explore two feasible approaches towards building models. The first is a time series approach, aiming at building ARIMA-GARCH/APARCH models, using the Box-Jenkins methodology. The second is to seek a set of meaningful explanatory variables and build linear regression models. We will compare and contrast the forecasting accuracy of these models.

1.1 Brief History of WTI crude oil for the past 20 years

Modelling the dyanmics of crude oil is not easy because its price may fluctuate from time to time unpredictably and may also depend on a lot of factors. For the past two decades, oil has been experiencing some ups and downs. In early 1999 (See Figure 1), there was the Asian Financial Crisis, along with Iraq deciding to increase oil production, which caused oil prices to reach a bottom. But the market adjusted quickly and reached over U.S.$ 34 by late 2000. The dotcom bubble in 2001 caused another round of economic panic, which caused it to drop until early 2002. Then, the global economy had been regaining momentum which resulted in a few years of bullish state. Accordingly, oil prices had been spiking up. Other contributing factors included limiting amount of global oil supply and hostile relationships between the U.S. and a number of oil producer countries. Oil price touched an all-time high of U.S.$ 147.30 in which the housing bubble in the U.S. started to burst, and a unprecedented credit crisis was followed. The dramatic decline in oil price that followed was difficult to model. Even though the price had, in general, exhibited a steady level of recovery after the financial crisis in 2008, it still posts a great challenge to find a model that performs consistently well, when being confronted with such unpredictable circumstances.

2 Box-Jenkins approach

2.1 ARIMA

We begin by exploring the ARIMA-Box Jenkins methodology, proposed by G. Box and G. Jenkins[1], to build a univariate time series forecasting model. It is one of the most common ways to formulate a forecasting model for a univariate time series such as the monthly WTI crude oil spot prices we intend to model. To understand what this methodology entails, we first define what is meant by an ARMA (and hence ARIMA) model.

Definition 2.1 (ARMA)**.** An ARMA (Autoregressive Moving Average) (p, q) model of a time series Y_t has the following form:

$$Y_t = \phi_0 + \sum_{i=1}^{p} \phi_i Y_{t-i} + \sum_{j=1}^{q} \theta_j \epsilon_{t-j} + \epsilon_t \tag{1}$$

1. p refers to the number of autoregressive terms
2. q refers to the number of lagged error terms
3. ϕ refers to the coefficients of the autoregressive terms and the constant
4. θ refers to the coefficients of the moving average terms

Definition 2.2 (White noise)**.** The ϵ_t defined above is a white noise process, meaning that it is a sequence of variables which have mean zero, variance σ^2, and zero correlation across time, i.e. $E(\epsilon_u \epsilon_v) = 0$ if $u \neq v$. They are also independent and identically distributed.

An ARIMA model is an extension of the ARMA class.

Definition 2.3 (ARIMA)**.** A time series Y_t is said to be of the ARIMA (Autoregressive Integrated Moving Average) format if $\nabla^d Y_t$ is a stationary ARMA process, where d refers to the number of difference to be taken on the original series.

However, in reality, it is common to model log returns instead of prices in financial time series. This is defined as the difference between current and previous log prices. To determine whether a log transformation is needed, with aims such as stabilising variance or creating a more Gaussian distributed set of data, we have the following procedure:

Definition 2.4 (Box-Cox transformation)**.** For each y, we can use the Box-Cox transformation and convert it into $y^{(\lambda)}$, where

$$\frac{y^{(\lambda)} - 1}{\lambda}, \text{if } \lambda \neq 0 \tag{2}$$

$$log(y), \text{if } \lambda = 0 \tag{3}$$

Accordingly, we can plot, with respect to the 95% confidence interval, a log-likelihood graph for the feasible values of λ.

In the Box-Jenkins approach, we also need to ensure the series is stationary, i.e. it has approximately constant mean and variance. This can be done by differencing the series several times (at most twice usually). The Augmented Dickey-Fuller Test[9] will be performed to see if the transformed series is stationary. This provides an alternative methodology to check for stationarity other than observing the plotted series itself.

Definition 2.5 (Augmented Dickey-Fuller Test)**.** It is a test to see if a time series has a unit root. If it does, the series is considered non-stationary. The null hypothesis here is that the time series is non-stationary.

After that, we try to plot the autocorrelation and partial correlation functions against different lags in order to determine the appropriate orders of p and q for our model. The expected patterns are summarised in the following:

	AR(p)	MA(q)	ARMA(p,q)
ACF	Tails off	Cuts off after lag q	Tails off
PACF	Cuts off after lag p	Tails off	Tails off

Sometimes, there may be more than one potential candidate for the final model. In this case, we may want to use an alternative approach based on information criteria to select the most approriate model. One way is to employ the AIC (Akaike Information Criterion) measure [10].

Definition 2.6 (AIC)**.** AIC measures how well the estimated model fits with the data relative to the other models, and is calculated using the formula below:

$$AIC(p, q) = -2log(L) + 2(p + q), \tag{4}$$

where L is the maximum value of the likelihood function of the ARMA(p,q) model, and p+q is the total number of parameters found in it. The recommended model under this approach is then the one with the smallest AIC value. Therefore, we want p+q to be small and L to be large. In other words, the approach encourages both goodness of fit and parsimony.

Once the model is specified, we employ the maximum likelihood estimation method to estimate the respective coefficients for the autoregressive and moving average terms. Finally, we perform model diagnostics to see if the estimated model is consistent with the specifications of a univariate time series process with stationarity. Graphically, we can employ the following plot:

Definition 2.7 (Normal QQ plot)**.** This refers to a plot of the quantiles of the model residuals against the quantiles of a Normal Distribution. On this plot there is a QQ line which represents a perfect match between model residuals and the Normal distribution. If the model residuals are normally distributed, they should be plotted along this line.

Also, we can perform the McLeod-Li test and the Jarque-Bera test on the residuals.

Definition 2.8 (McLeod-Li test)**.** The McLeod-Li test[12] is based on the sample autocorrelations of the squared time series. It has the following form:
H_0: The residuals are independent.
H_1: The resdiauls are dependent.

$$L = (n^2 + 2n) \sum_{k=1}^{K} \frac{\hat{\rho}_k^2}{n-k}, \tag{5}$$

$$\hat{\rho}_k = \frac{\sum_{t=k+1}^{n} (\epsilon_t^2 - \hat{\sigma}_n^2)(\epsilon_{t-k}^2 - \hat{\sigma}_n^2)}{\sum_{t=1}^{n} (\epsilon_t^2 - \hat{\sigma}_n^2)}, \hat{\sigma}_n^2 = \frac{\sum_{t=1}^{n} \epsilon_t^2}{n},$$

Provided that Y_t is a series of identical and independently distributed sequence, the statistic L should have a distribution of $\chi^2(K)$ asymptotically. If we apply the test to the residuals of the fitted ARIMA model and that the returned p-values are below 0.05, we have strong evidence to reject the null hypothesis that all lagged autocorrelations up to lag k are zero.

Definition 2.9 (Jarque-Bera test)**.** The Jarque-Bera test[11] can be defined in the following with test statistic L.
H_0: The residuals are normally distributed.
H_1: The residuals do not follow a normal distribution.

$$L = \frac{n}{6}(S^2 + 0.25(K-3)^2), \tag{6}$$

where n is the number of data, S is the sample skewness and K is the sample kurtosis. If the data come from a normal distribution, the statistic L should have the asymptotic distribution of a $\chi^2(2)$ distribution.

If the residuals violate our assumptions, we will resort to the GARCH (Generalised Autoregressive Conditional Heteroskedasticity) methodology to model the residuals.

2.2 Innovation modeling

In residual analysis of financial time series, Generalised Autoregressive Conditional Heteroskedasticity (GARCH) models, proposed by Bollerslev in 1986[6], are widely used to specify and model innovations, i.e. the differences between the fitted values of the proposed model and the observed values. They have the ability to model the phenomenon of volatility clustering seen in many financial time series data. They evolve from the Autoregressive Conditional Heteroskedasticity (ARCH) model, proposed by Engle in 1982[5], which assumes the variance of the current innovation to be dependent on error terms of the previous time periods.

Definition 2.10 (ARCH)**.** ARCH(q) models, where q refers to the order of the lagged autoregressive terms of previous innovations, have the following term:

$$\epsilon_t = \sigma_t z_t, z_t \text{ white noise},$$
$$\sigma_t^2 = \omega + \sum_{i=1}^{q} \alpha_i \epsilon_{t-i}^2, \omega > 0, \alpha_i \geq 0, i > 0 \tag{7}$$

Definition 2.11 (GARCH)**.** On the other hands, GARCH models assume the conditional variances of innovations follow an ARMA model. In that case, a GARCH(p, q) model, where p and q are the orders of the GARCH and ARCH terms respectively, refers to σ_t having the following term:

$$\epsilon_t = \sigma_t z_t, z_t \text{ white noise},$$
$$\sigma_t^2 = \omega + \sum_{i=1}^{q} \alpha_i \epsilon_{t-i}^2 + \sum_{i=1}^{p} \beta_i \sigma_{t-i}^2, \omega > 0, \alpha_i \geq 0, \beta_i \geq 0, i > 0 \tag{8}$$

Also, we need to impose $\sum_{i=1}^{q} \alpha_i + \sum_{i=1}^{p} \beta_i < 1$ to ensure stationarity. We note that GARCH will collapse to ARCH if p=0.

Also, we can definte APARCH[8], which is another generalised conditional heteroskedasticity model with the special feature that can account for asymmetric effects of volatilities. It is defined in the following:

Definition 2.12 (APARCH)**.**

$$\epsilon_t = \sigma_t z_t, z_t \text{ white noise},$$

$$\sigma_t^\delta = \omega + \sum_{i=1}^{q} \alpha_i (|\epsilon_{t-i}| + \gamma_i \epsilon_{t-i})^\delta + \sum_{i=1}^{p} \beta_i \sigma_{t-i}^\delta, \omega > 0, \alpha_i \geq 0, \beta_i \geq 0, i > 0 \quad (9)$$

To seek for the right GARCH/APARCH model, we plot squared residuals of the fitted model and their acf and pacf. Apply the same procedure to find out until we find a GARCH/APARCH(p,q) model on the squared residuals such that the AIC value is the smallest. Altogether, we will have the fitted ARIMA-GARCH/APARCH models for forecasting future values.

2.3 Forecasting

Definition 2.13 (n-step ahead forecast)**.** The one-step ahead forecast of Y_{t+1} based on an ARMA-GARCH model is defined as

$$\hat{Y}_t(1) = E(Y_{t+1}|Y_t, Y_{t-1}...) = \phi_0 + \sum_{i=1}^{p} \phi_i Y_{t+1-i} + \sum_{j=1}^{q} \theta_j \epsilon_{t+1-j}, \quad (10)$$

where the ϵs follow the stated GARCH model. Recursively, we can then define the n-step ahead forecast of $\hat{Y}_t(n)$ for any n.

We can employ two major statistical measures for comparing forecast accuracy. They are the MSE and the MAE respectively.

Definition 2.14 (Mean squared error)**.** It computes the squared difference between every forecasted value and every realised value of the quantity being estimated, and finds the mean of them afterwards. One can interpret it literally as the average of the squares of errors. Assume Y_j as the j-step ahead realised value, the error has the following formula:

$$MSE = \frac{1}{n} \sum_{j=1}^{n} ((\hat{Y}_t(j) - Y_j)^2) \quad (11)$$

Definition 2.15 (Mean absolute error)**.** It computes the mean of all the absolute, instead of squared, forecast errors. The formula is the following:

$$MAE = \frac{1}{n} \sum_{j=1}^{n} (|\hat{Y}_t(j) - Y_j|) \quad (12)$$

Although both ways are closely related, MAE is more popular in use as it does not require squaring, hence it is considered easier to implement. However, MSE may perform better in some occassions. For example, assume we have two different models: one has 10 error values with the first nine equal to 1 and the remaining

one equal to 11; the other also has 10 error values which are all equal to 2. In this particular case, we have equal MAE (both equal to 2) but the MSE are totally different (13 against 4). The second model is clearly preferred. Therefore, one should compute both metrics to obtain a better picture of the forecasting powers of various models.

3 Empirical results of univariate approach

WTI (West Texas Intermediate) crude oil, along with Brent crude oil, is widely considered as one of the benchmarks for understanding oil price dynamics. Many trades and derivatives are based on its prices. In this paper, we obtain the data from the U.S. EIA (Energy Information Administration) where monthly spot prices of the WTI oil are available. The time frame we choose spans between April 1993 and March 2013, i.e. 20 years of data. The monthly frequency is chosen as this is the smallest unit of frequency available for all variables. We then further divide the time frame into two 10-year periods, with the first nine years of data in each period used for model construction and the last year validating forecast accuracy. In particular, we can see that, although oil prices do experience ups and downs in the first period, the second period suffers a much larger scale of fluctuations, especially during 2008, the year when the financial crisis happened. Therefore, we should examine if there is a structual change in the oil market from one period to another.

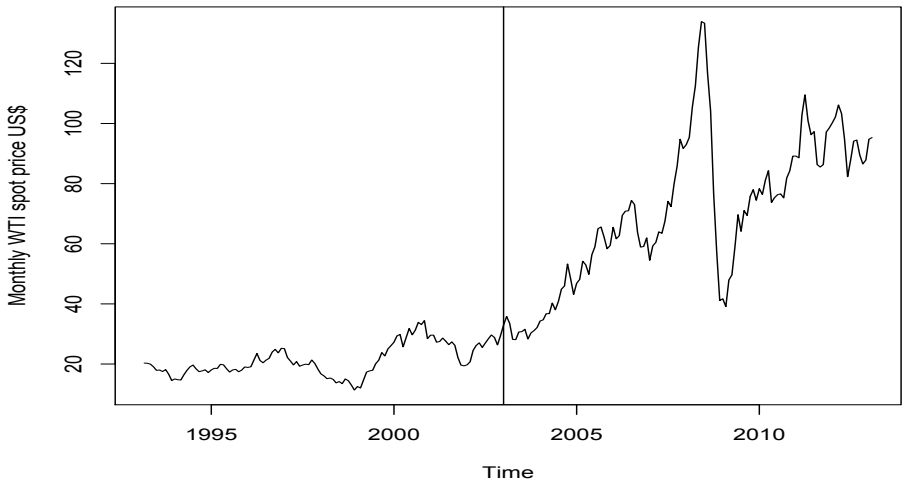


Figure 1: Time series plot of monthly WTI spot price from Apr 1993 to Mar 2002 (1st period, left) and from Apr 2003 to Mar 2012 (2nd period, right)

3.1 ARIMA fits

We first attempt to see if a log transformation is recommended by plotting the Box-Cox log likelihood graphs. As $\lambda = 0$ falls within the 95% confidence interval for both periods, we adopt such a transformation. From Figure 2, we can see

clearly that both series are still clearly non-stationary.

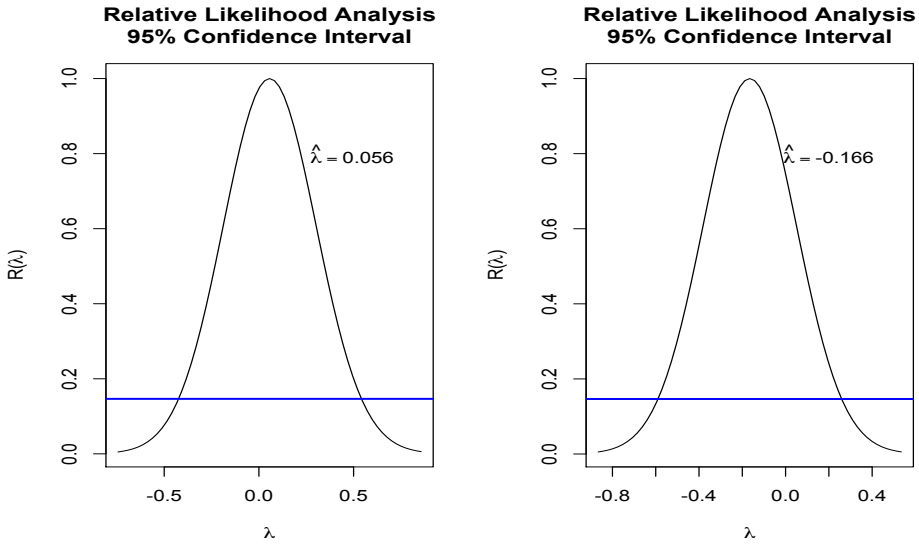


Figure 2: Box-Cox likelihood plots for period 1(left) and 2(right)

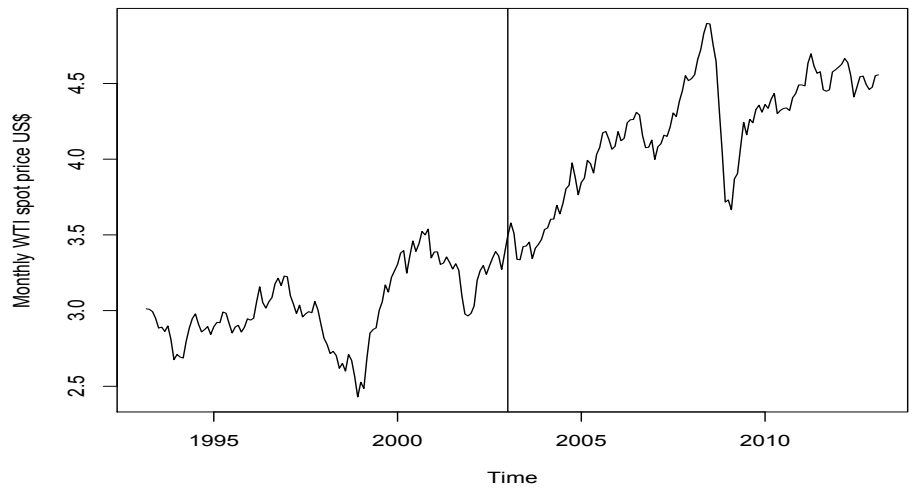


Figure 3: Time series plot of monthly WTI log price in period 1 (left) and 2 (right)

The Augmented Dickey-Fuller test for stationarity suggests a p-value of 0.47 and 0.09 respectively, hence we cannot reject the null hypothesis at the 5% significant level. We try working on log return instead. The formula for log return

is $L_t = log(\frac{P_t}{P_{t-1}})$. This time, we see that the ADF test provides p-values smaller than 0.01 in both periods, suggesting stationary in the series, which can also be confirmed by looking at Figure 4.

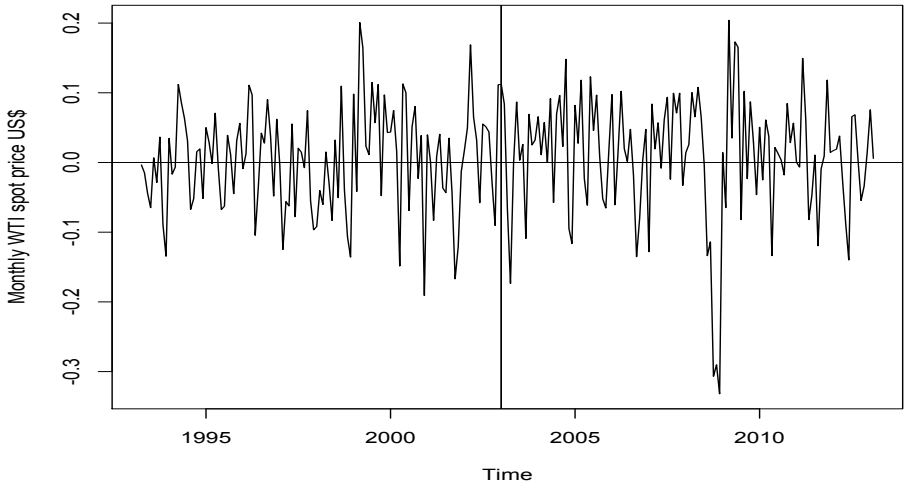


Figure 4: Monthly log returns of WTI crude oil in period 1 (left) and 2 (right)

Then, we start working on modeling the autoregressive (AR) and moving average (MA) orders respectively.

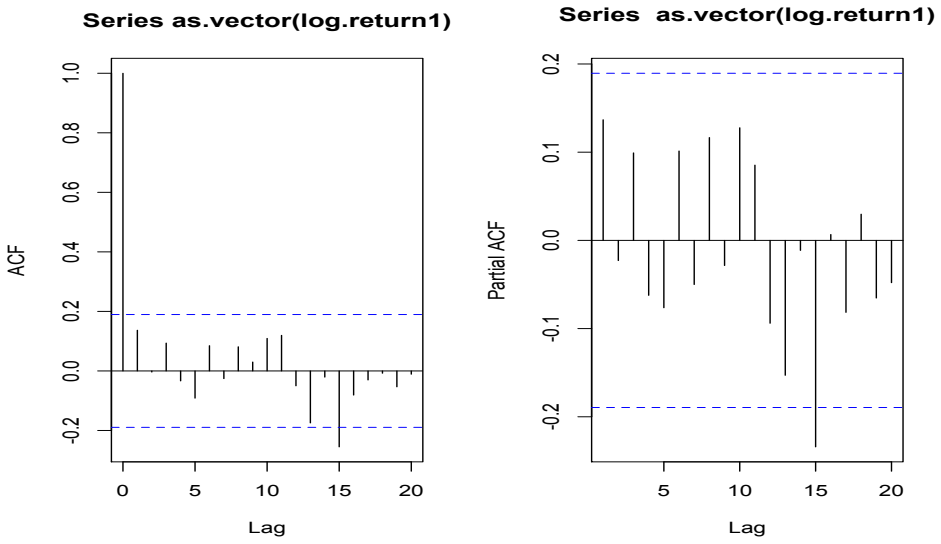


Figure 5: ACF and PACF plots for period 1

For period 1, we see no significant spikes in either plot, except at lag 15 where it is slightly below the lower confidence bound. ARMA(0,0) (random walk) model seems to be a potential candidate here.

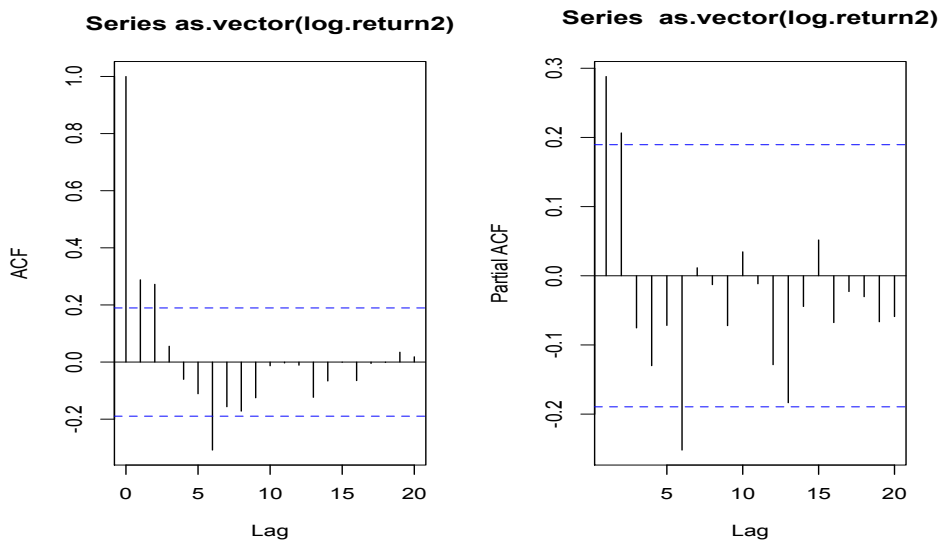


Figure 6: ACF and PACF plots for period 2

For period 2, spikes seem to be more significant up to lag 2 in both plots. This suggest an ARMA (2,2) model should be postulated as a potential fitted model. Then, we compare the various AIC values generated by fitting different p and q, where they range from 0 to 2, and conclude a ARMA(0,1) model and a ARMA(2,2) model for the monthly log return of WTI crude oil in the 1st and the 2nd period respectively. In particular, if we increase either order by 1 in the ARMA (2,2) model, we get higher AIC values (-220.9 and -214.7), resulting in overfitting. Below are the table of AIC values.

Order	(0,0)	(0,1)	(0,2)	(1,0)	(1,1)	(1,2)	(2,0)	(2,1)	(2,2)
AIC	-249.8	-250.0	-248.1	-249.9	-248.3	-246.3	-248.0	-245.9	-244.5
AIC	-208.9	-213.2	219.0	-216.1	-216.4	-217.7	-218.7	-217.0	-222.4

Table 1: AIC table for various orders

3.2 Residual analysis and GARCH/APARCH fits

We first look at the plots of model residuals across the two periods. They both look reasonably stationary and seem to evolve around a mean of zero. We should also note that, during the financial crisis, we see some relatively large model residuals. We then proceed to plot their ACFs and PACFs.

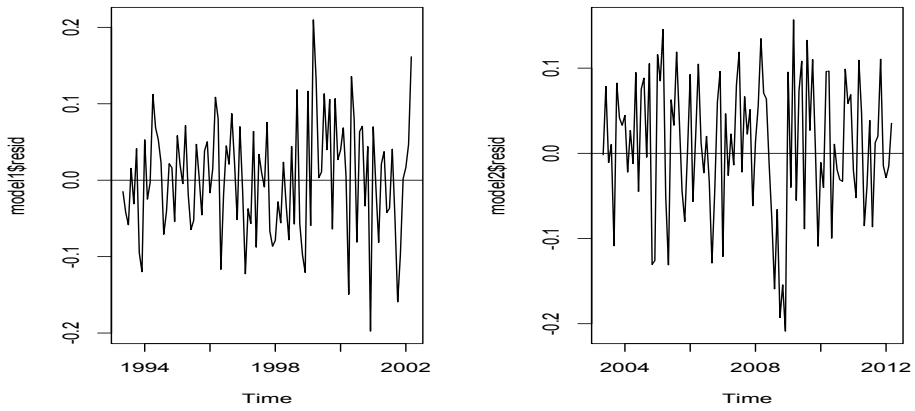


Figure 7: Residual plots for model 1 and 2

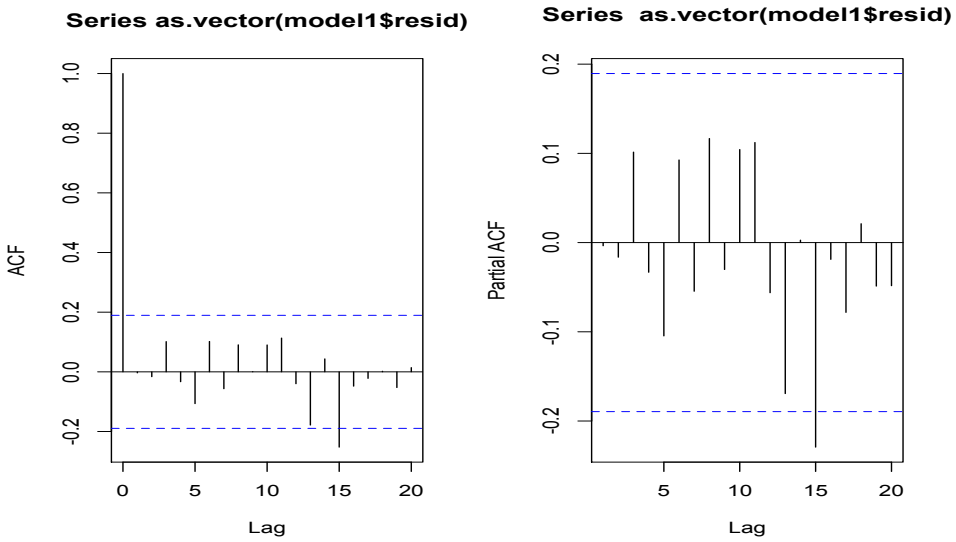


Figure 8: ACF and PACF plots for residuals of model 1

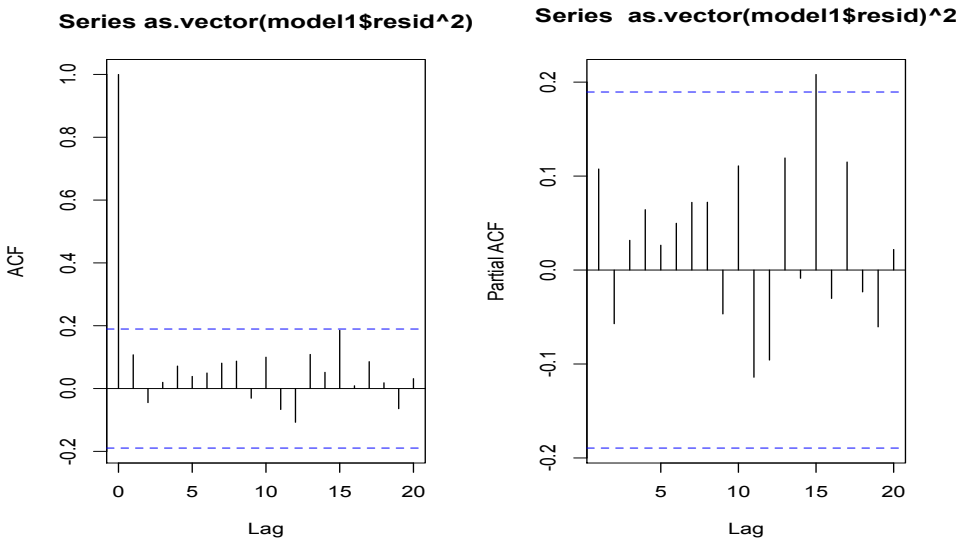


Figure 9: ACF and PACF plots for squared residuals of model 1

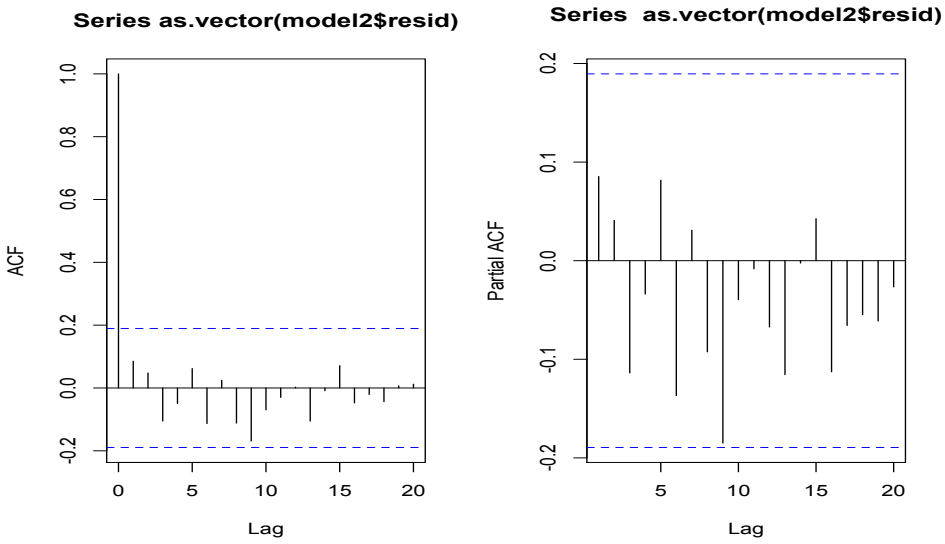


Figure 10: ACF and PACF plots for residuals of model 2

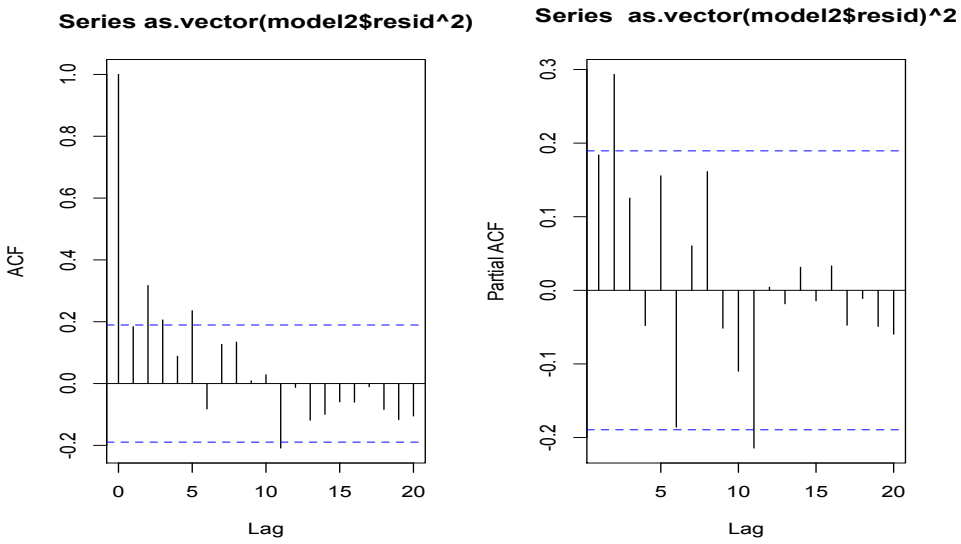


Figure 11: ACF and PACF plots for squared residuals of model 2

As expected, there is no significant autocorrelation for the residuals across any lag in either period. For period 1, there is no particularly significant spikes in any lag. For period 2, lag 2 seems to catch our attention for the ACF and PACF of the squared residuals.

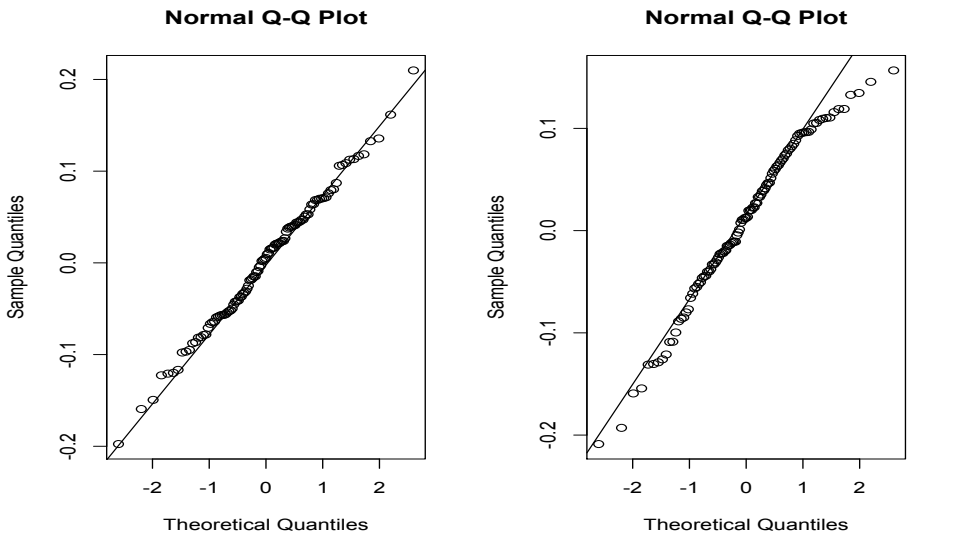


Figure 12: Normal QQ plots for model 1 and 2

The normal QQ-plots suggest strong evidence of normality for period 1, but

only a mild evidence for period 2. However, the Jarque-Bera test suggests a p-value of 0.9932 and 0.1492 respectively, so we cannot reject the null hypothesis that the residuals are normally distributed.

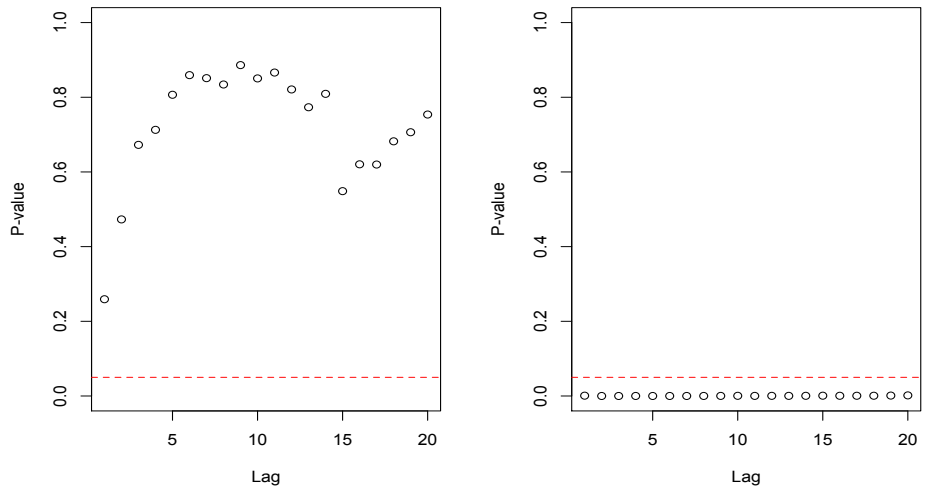


Figure 13: Mcleod.Li test results plotted for period 1(left) and 2(right)

On the other hand, the McLeod-Li test suggests a strong evidence that the residuals for period 2 are autocorrelated. The plots of the p-values against different lags indicate that from lag 2 onwards in the second period, all p-values lie below the 0.05 threshold. This is a stark contrast to what we see in the first period, where all values lie way above the threshold. Therefore, a GARCH model may be appropriate to fit the second set of data. The intuitive rationale behind that is possibly linked to the fact that the financial crisis has brought in some significant effects of volatility clustering. We then loop through a set of AIC values for various orders of GARCH model, and we see that a GARCH(1,1) model provides the lowest value (-234.5455). To validate the choice, we plot the residuals of the GARCH(1,1) model for period 2 in Figure 14 and 15. We notice that all the significant spikes are removed, confirming the adequacy of the model.

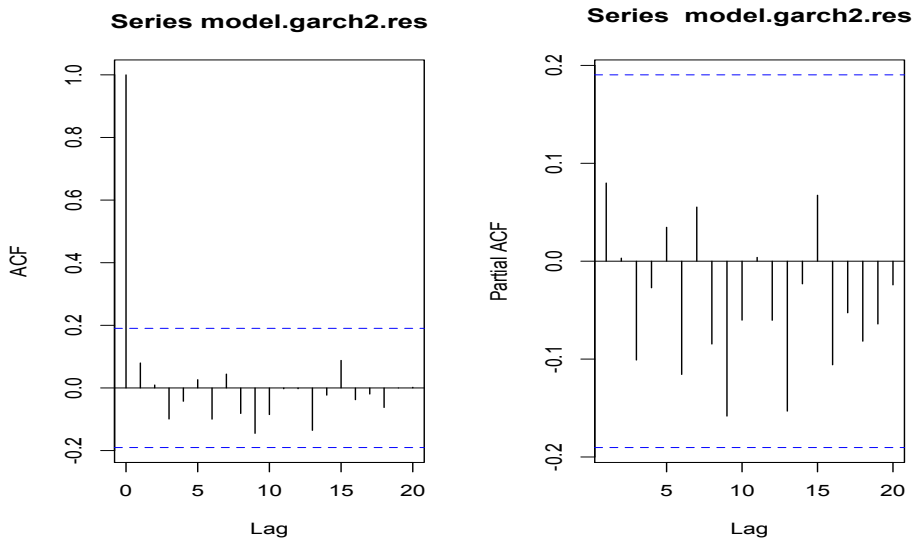


Figure 14: ACF and PACF plots of residuals of GARCH(1,1) model for period 2

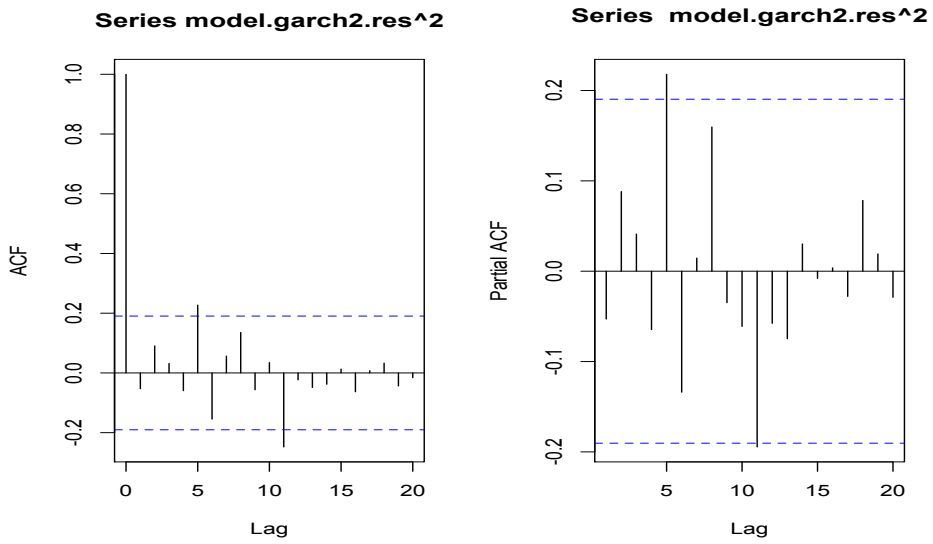


Figure 15: ACF and PACF plots of squared residuals of GARCH(1,1) model for period 2

Our finalised are therefore presented below. For comparison purpose, we have also included an APARCH(1,1) fit to the residuals. We will evaluate their forecasting powers later.

Order	AR(1)	AR(2)	MA(1)	MA(2)	GARCH(1)	ARCH(1)
1st period			0.15			
2nd period	0.62	-0.24	-0.53	0.38	0.18	0.57
Order	AR(1)	AR(2)	MA(1)	MA(2)	APARCH(1)	ARCH(1)
2nd period	0.65	-0.30	-0.58	0.41	0.07	0.58

Table 2: Parameters for fitted models

4 Linear regression approach

4.1 Stepwise AIC approach towards model selection

Linear models can be used both to describe the statistical relationships between response variable and a list of explanatory variables, and to make predictions regarding the response variable.

Let p be the number of explanatory variables and n be the number of observations for each variable. In a linear regression, we have that $y_i = \beta_1 x_{1,i} + \beta_2 x_{2,i} + ... + \beta_p x_{p,i} + \epsilon_i$, where y and x_is refer to response variable and explanatory variables respectively. It is also convenient to express the relationship in matrix form,

$$Y = X\beta + \epsilon, \tag{13}$$

$$Y = (y_1, ..., y_n)^T,$$

$$X = \begin{pmatrix} 1 & x_{1,1} & ... & x_{p,1} \\ 1 & x_{1,2} & ... & x_{p,2} \\ \vdots & \vdots & ... & \vdots \\ 1 & x_{1,n-1} & ... & x_{p,n-1} \\ 1 & x_{1,n} & ... & x_{p,n} \end{pmatrix},$$

$$\beta = (\beta_1, ..., \beta_p)^T,$$

$$\epsilon = (\epsilon_1, ...\epsilon_n)^T$$

In fitting a linear regression model, our goal here is to derive a set of estimated $\hat{\beta}$. First, we need to make a few major assumptions:

1. **Linearity.** This means that the mean of the response variable is a linear combination of the regression coefficients and the explanatory variables. In mathematical terms, this essentially means we need $E(Y) = X^T\beta$.

2. **Absence of multicollinearity among explanatory variables.** In mathematical terms, this means we need the columns of $\mathbf{X}$ to be linearly indepedent and form a basis of the matrix. This way, we can guarantee a vector of regression coefficients exists. However, this may be violated if we include variables that exhibit perfect correlation. In reality, it is rare to have a correlation number of 1 or -1 between two variables. We decide to follow [15], and define 0.75 as the threshold. Any absolute correlation number above that is considered as perfect or near-perfect correlation.

3. **Precision of data.** Here, we need our data collected for all variables to be accurate so that we can treat them as deterministic values. We believe EIA should be a reliable source of information and we assume the validility of this point.

4. **Homoscedasticity.** This means we have constant variances in residuals. A plot of standardised residuals and model fitted values should review a scattered distribution if this is the case.

5. **Uncorrelated errors.** This assumes that the errors of the response variables are indepedent of each other.

6. **Normality of errors with zero mean.** This can be verified by normal QQ-plots, just like what we did for the univariate time series analysis. As for the mean, we just sum all the residuals and divide them by the total number of errors to see if the value is close to zero.

Here, we recall Akaike Information Criterion (AIC)[10] as a measure to determine which orders of autoregressive and moving average terms we include in the final model. This can be similarly adopted in regression model selection. The definition here is slightly modified:

Definition 4.1 (Akaike Information Criterion for linear regression)**.** Let RSS be the residual sum of squares, i.e. RSS=$(Y - X\beta)^T(Y - X\beta)$. Then, we have

$$AIC_{LR} = n\log(\frac{RSS}{n}) + 2p, \tag{14}$$

Below are three approaches that make use of the metric:

1. **Forward selection:** fit the null model, add each term separately, keep the one that reduces AIC the most, use this as the base model, add each of the other terms separately, keep the one that reduces AIC the most, repeat until reduction in AIC is no longer significant.

2. **Backward elimination:** fit the full model, remove the term that reduces AIC the least, use this as the base model, remove the next term that reduces AIC the least, repeat until any further reduction in AIC is significant.

3. **Stepwise selection:** this is a combination of the above approaches. We determine at every time whether any term can be dropped or added. If the answer is no, the model is complete. This can be better than forward selection since the initial estimate is hardly correct. Also, it can be chosen over backward elimination, since the full model may sometimes be too complicated to fit.

4.2 Estimation of regression parameters

Once we select the right model, we proceed to estimate the regression parameters. The Least Squares approach contends that we measure a $\hat{\beta}$ that minimises the residual sum of sqaures, i.e.

$$\frac{\partial}{\partial \beta}(Y - X\beta)^T(Y - X\beta) = 0 \tag{15}$$

The partial differentiation leads to

$$-2X^TY + 2(X^TX)\hat{\beta} = 0 \tag{16}$$

Assuming X has rank p, i.e. full rank, we can reach a solution:

$$\hat{\beta} = (X^TX)^{-1}X^TY \tag{17}$$

To validate the use of this estimator, we need to check that such $\hat{\beta}$ is an unbiased estimate of the true β, i.e. $E(\hat{\beta}) = \beta$. This is shown in the following:

$$\begin{aligned} E(\hat{\beta}) &= E((X^TX)^{-1}X^TY) \\ &= (X^TX)^{-1}X^TE(Y) \\ &= (X^TX)^{-1}X^TX\beta \ (Assuming\ E(\epsilon) = 0) \\ &= \beta \end{aligned} \tag{18}$$

Once we have calculated these parameters, we can proceed to generate 12-month ahead predicted values for all explanatory variables. For example, if we have seven explanatory variables, we will a total of 96 predicted values. We then plug them into the regression model and predict the forecasted values of the response variable (monthly WTI crude oil log return in this case). Just like the univariate model, the model accuracy can be measured by MSE or MAE.

5 Empirical results of linear regression approach

5.1 Fitting explanatory variables

Many literatures have been dedicated to finding contributing factors to oil prices. We will include a list of eight fundamentals we deem relevant to explain the dynamics of oil price returns, and their respective time series plots. These data are all gathered from the EIA. After a literature review of several relevant papers, the following points are included, with their respective plots in Figure 16:

1. **Consumption(Thousand barrels per day):** in the field of economics, we know that equilibrium price of a particular asset, crude oil in this case, is determined by the intersection of its demand and supply curve. Therefore, we should include factors that may directly influence the demand curve, such as market consumption (thousand barrels per day). This is explored in [16] and [17].

2. **Production(Thousand barrels per day):** similarly, increase in production should impact the supply curve (by shifting it rightwards), hence it should play an important role in oil price determination. This is explored in [25] and [26].

3. **Ending Stock(Thousand barrels):** this refers to crude oil monthly ending stocks (thousand barrels) recorded by the EIA. Such stocks exist to avoid a short run of oil in case of unpredicted events, and the impact on prices should be two-folded. On one hand, it reduces dependence on current production; on the other hand, if people conceive that that there is excess supply at storage, they may be more willing to consume crude oil. This is explored in [15] and [25].

4. **Net import(Thousand barrels per day):** as one of the largest oil consumption countries in the world, the U.S. has long been importing oils from other countries. Net import should serve as an indicator on how much the U.S. is in need of extra oils apart from internal supply, hence helping contribute to the determination of oil prices. This is explored in [27].

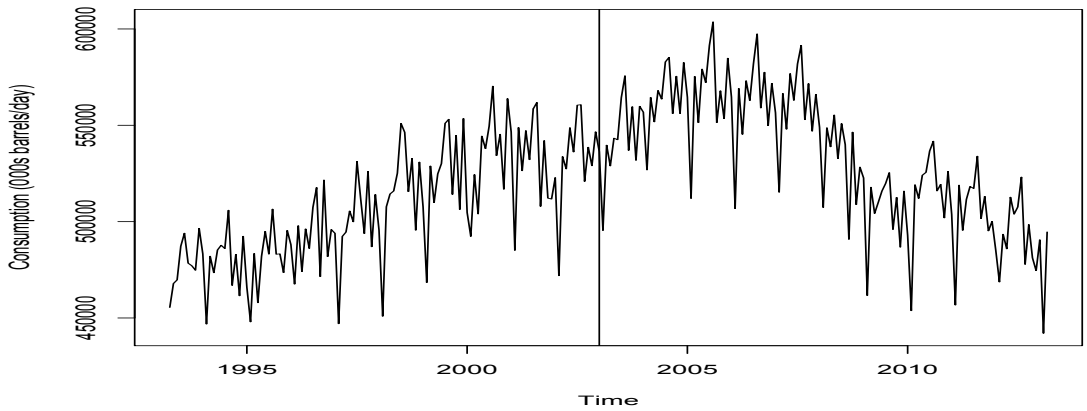
5. **Refinery utilisation rate(%):** here we refer to the utilisation of refinery, which is used to process crude oil into petroleum products. The number represents the utilization of all crude oil distillation units. It is calculated by dividing gross inputs to these units by the operable refining capacity of the unit, which is in turn defined as the amount of capacity that, at the beginning of the period, is in operation, or those which are not in operation and not under active repair, but capable of being placed in operation within 30 days.

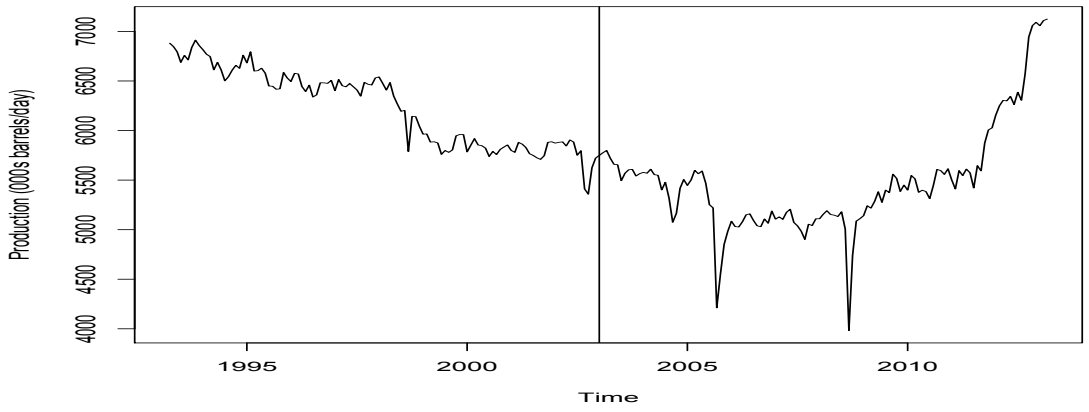
It is essentially the total of operating and idle capacity for oil production. Ideally, a higher rate indicates a more effective use of crude oil. This has been considered in [15] and [24].

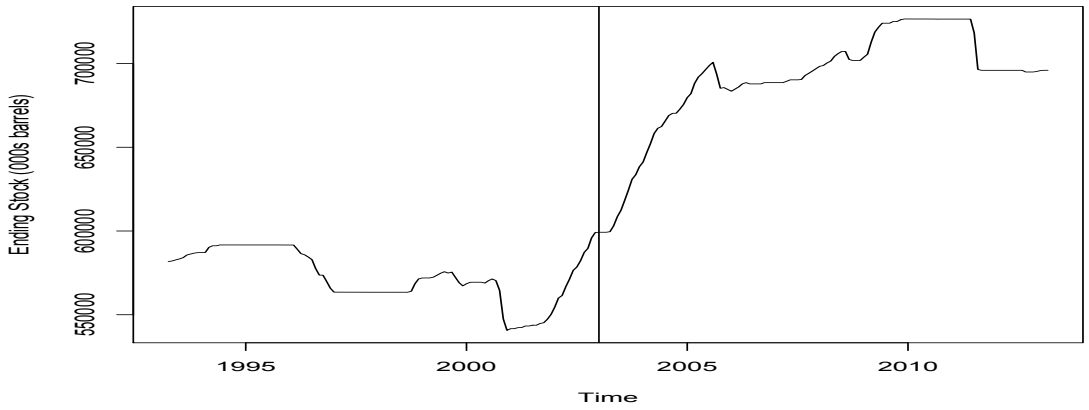
6. **Interest rate(%):** the US Federal Funds effective monthly rates is used for this important macroeconomic indicator. Interest rate has always been one of the most popular monetary tools used by government to boost the economy. Reducing its level implies a lower cost of borrowing, thereby encouraging people to consume and invest more. This has been considered in [18], [19] and [21].

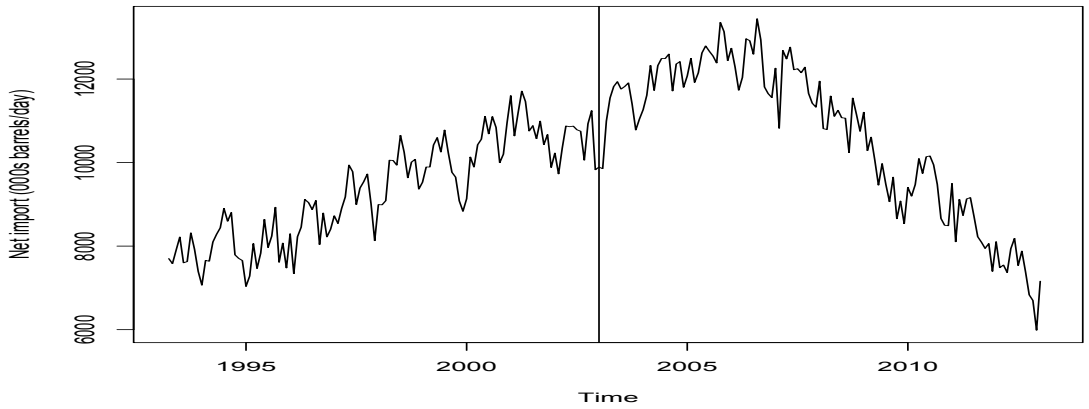
7. **NYMEX Oil Futures Contract 4 (US dollars per barrel):** The market's prediction of what crude oil should trade in the future should be an indicator of how it is trading at the moment. We use the NYMEX Contract 4 to reflect this view since this is the longest-dated price data available for access. The contract expires at the 4th earliest delivery date, which is usually the 25th day of a month. The inclusion of oil futures is explored in [15] and [24].

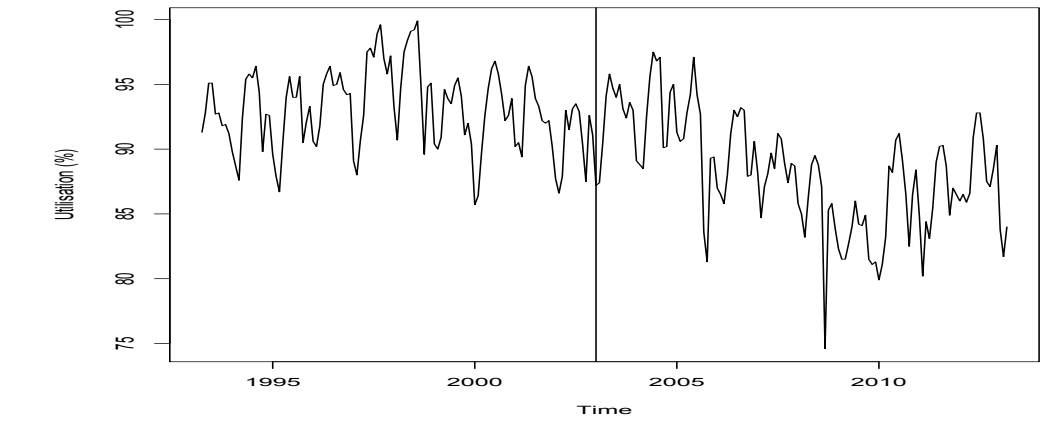
8. **S&P 500 index:** as oil is arguably one of the most important commodities in the U.S., it should be very sensitive to the macroeconomic enviroments. The monthly closing of the S&P 500 index, one of the most important indicators of the U.S. stock market, is included for this reason. This has been considered in [18] and [23].

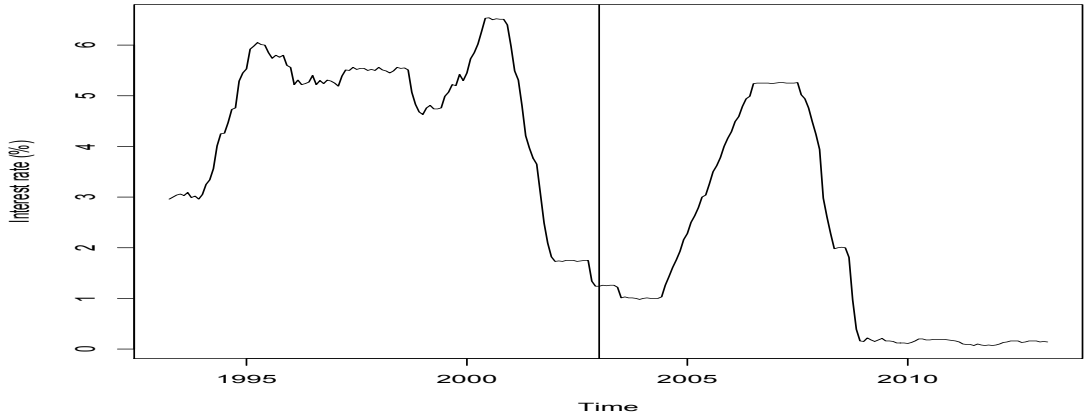


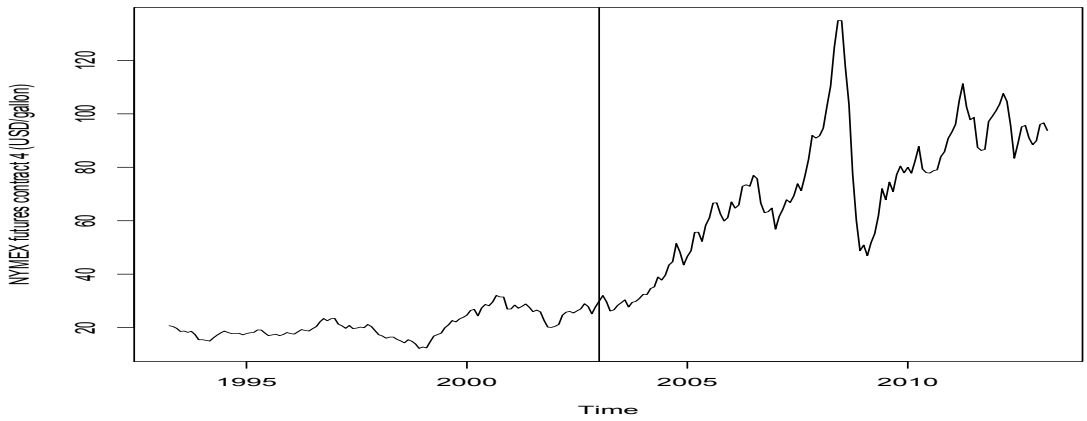

Production (000s barrels/day)
4000
4500
5000
5500
6000
6500
7000
1995
2000
2005
2010
Time


Ending Stock (000s barrels)
550000
600000
650000
700000
1995
2000
2005
2010
Time


Net import (000s barrels/day)
6000
8000
10000
12000
1995
2000
2005
2010
Time


Utilisation (%)
100
95
90
85
80
75
1995
2000
2005
2010
Time


Interest rate (%)
0
1
2
3
4
5
6
1995
2000
2005
2010
Time


NYMEX futures contract 4 (USD/gallon)
20
40
60
80
100
120
1995
2000
2005
2010
Time

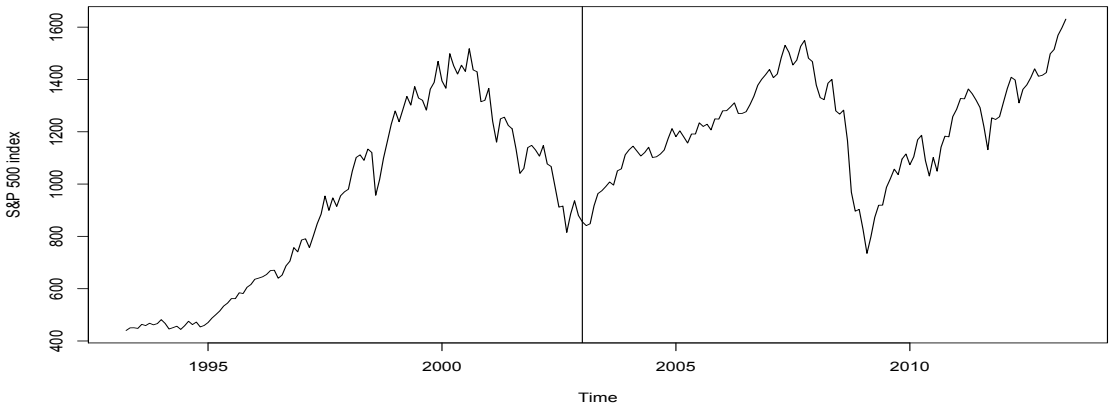


Figure 16: Plots of various explanatory variables

We first try to fit two linear regression models based on these data. For period 1, only ending stock is a significant factor at the 5% confidence level. Period 2 is not much better either. Only consumption and production are considered significant in the full model. We have possibly included too many variables.

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	-0.5297	0.3246	-1.63	0.1059
Consumption	-0.0000	0.0000	-0.19	0.8507
Production	-0.0000	0.0000	-1.25	0.2154
Ending Stock	0.0000	0.0000	3.49	0.0007
Net imports	-0.0000	0.0000	-0.12	0.9057
Refinery utilisation rate	0.0002	0.0013	0.16	0.8764
Interest rate	-0.0022	0.0031	-0.70	0.4829
Futures	0.0014	0.0009	1.59	0.1157
S&P	0.0000	0.0000	0.58	0.5616

Table 3: Regression model for period 1 before stepwise AIC

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	-0.4113	0.2211	-1.86	0.0658
Consumption	0.0000	0.0000	2.10	0.0384
Production	0.0001	0.0000	2.71	0.0079
Ending Stock	0.0000	0.0000	0.50	0.6193
Net imports	0.0000	0.0000	0.35	0.7264
Refinery utilisation rate	-0.0031	0.0017	-1.83	0.0702
Interest rate	0.0017	0.0053	0.33	0.7427
Futures	0.0001	0.0003	0.37	0.7096
S&P	0.0000	0.0000	0.19	0.8496

Table 4: Regression model for period 2 before stepwise AIC

Below are two table demonstrating the correlations among various explanatory variables. The pairs exhibiting an absolute correlation number above the threshold value of 0.75 are highlighted. Our conjecture is that under the stepwise AIC approach, some of these variables will be removed in the final models. In particular, net import seems to be correlated with a number of variables, so it may be deemed redundant.

	Oil.return	Consumption	Production	Stock	Net.import	Utilisation.rate	Interest.rate	Futures.Contract.4	S.P.closing
Oil.return	1.00	0.05	-0.13	0.17	0.02	-0.08	0.06	0.10	0.09
Consumption	0.05	1.00	-0.73	-0.59	0.76	0.31	0.17	0.46	0.72
Production	-0.13	-0.73	1.00	0.69	-0.87	0.03	-0.11	-0.56	-0.91
Stock	0.17	-0.59	0.69	1.00	-0.79	-0.07	0.06	-0.54	-0.71
Net.import	0.02	0.76	-0.87	-0.79	1.00	0.23	0.11	0.53	0.85
Utilisation.rate	-0.08	0.31	0.03	-0.07	0.23	1.00	0.28	-0.05	0.04
Interest.rate	0.06	0.17	-0.11	0.06	0.11	0.28	1.00	0.20	0.24
Futures.Contract.4	0.10	0.46	-0.56	-0.54	0.53	-0.05	0.20	1.00	0.51
S.P.closing	0.09	0.72	-0.91	-0.71	0.85	0.04	0.24	0.51	1.00

Table 5: Correlation table for variables in period 1

	Oil.return	Consumption	Production	Stock	Net.import	Utilisation.rate	Interest.rate	Futures.Contract.4	S.P.closing
Oil.return	1.00	0.08	0.17	0.06	-0.05	0.02	0.05	0.08	0.16
Consumption	0.08	1.00	-0.34	-0.44	0.80	0.67	0.64	-0.37	0.26
Production	0.17	-0.34	1.00	-0.10	-0.55	0.21	-0.57	-0.05	-0.16
Stock	0.06	-0.44	-0.10	1.00	-0.51	-0.62	-0.20	0.71	0.14
Net.import	-0.05	0.80	-0.55	-0.51	1.00	0.54	0.73	-0.50	0.13
Utilisation.rate	0.02	0.67	0.21	-0.62	0.54	1.00	0.25	-0.47	0.05
Interest.rate	0.05	0.64	-0.57	-0.20	0.73	0.25	1.00	-0.07	0.64
Futures.Contract.4	0.08	-0.37	-0.05	0.71	-0.50	-0.47	-0.07	1.00	0.48
S.P.closing	0.16	0.26	-0.16	0.14	0.13	0.05	0.64	0.48	1.00

Table 6: Correlation table for variables in period 2

Now that we understand we may have included too many variables, we try the stepwise AIC procedure to select a reduced model. The result shows that under the models with the lowest AIC values, all variables are significant now. In particular,

net import is removed, as expected. Note also that the remaining variables do not exhibit multicollinearity anymore, as seen in the correlation tables. We therefore conclude them as the final models.

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	-0.4470	0.1372	-3.26	0.0015
Production	-3.642e-05	0.0000	-3.21	0.0018
Ending stock	1.140e-06	0.0000	4.23	0.0001
Interest rate	1.184e-03	0.0008	1.47	0.0145

Table 7: Regression model for period 1 after stepwise AIC

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	-0.2615	0.1037	-2.52	0.0132
Consumption	5.890e-07	0.0000	2.94	0.0041
Production	4.789e-05	0.0000	3.29	0.0014
Refinery utilisation rate	-3.477e-03	0.0014	-2.47	0.0151

Table 8: Regression model for period 2 after stepwise AIC

5.2 Residual analysis

For both periods, we present a normal QQ plot and a plot of residuals against fitted values. The second plots for both periods look reasonaly scattered. The normal QQ plot in period 2 looks a bit off the line at the bottom left corner. A closer look reveals that those points (refering to data points no. 67, 68 and 69) all evolve around the financial crisis, where data of almost all variables experience abnormally large fluctuations. If we ignore these three points, the normality assumption of residuals is still valid. Hence, treating these points as outliers, we do not have enough evidences to say the two models fitted violate the regression assumptions made before.

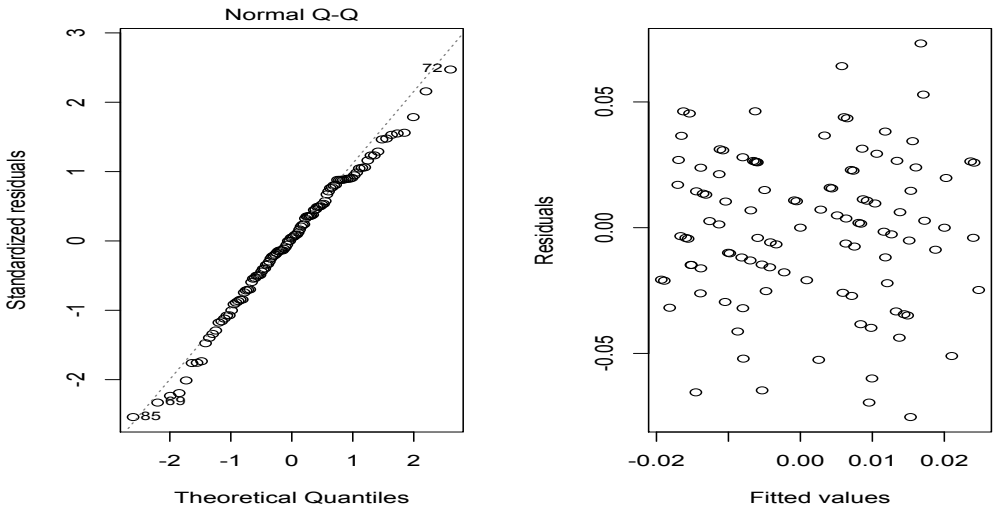


Figure 17: Residual plots of linear model 1 in the first period

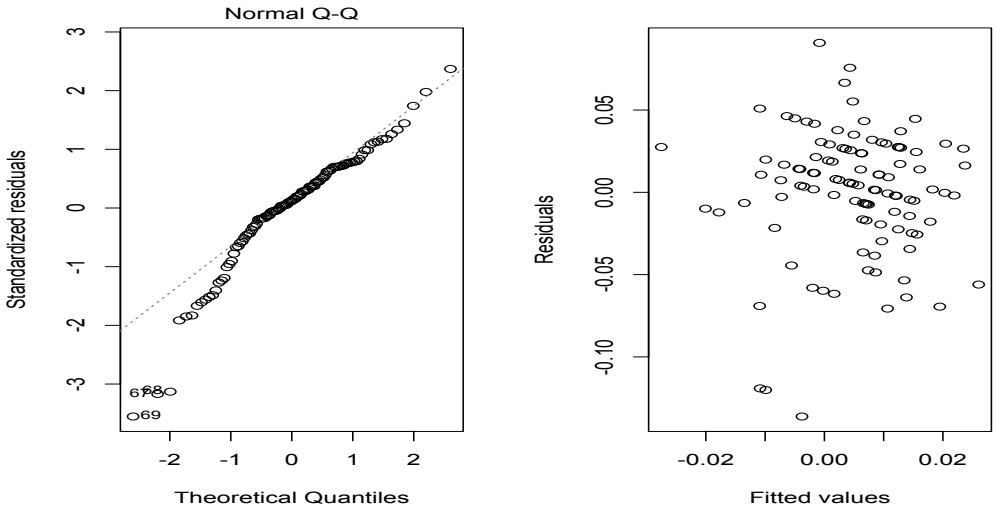


Figure 18: Residual plots of linear model 2 in the second period

6 Conclusion

6.1 Forecasting comparison

Now that we have produced several feasible models, we aim to analyse their forecasting accuracies. Here, we use a similar approach as in [15], where multiple n-step ahead (n=1,3,12 in this case) MSE and MAE are generated. This way, we get a better picture of how well these models predict the future as time evolves. The lowest value in each row has been highlighted in read, indicating it as the best model in terms of forecasting power.

Period 1 (1M)	Pure ARIMA(0,0,1)	GARCH(1,1)	APARCH(1,1)	Regression
MSE	0.0008273015	1.508008e-06	3.151164e-05	0.03918434
MAE	0.02793794	0.001228010	0.005613523	0.0271045
Period 2 (1M)	Pure ARIMA(2,0,2)	GARCH(1,1)	APARCH(1,1)	Regression
MSE	0.0003327995	0.0001827951	0.0001540007	0.1051161
MAE	0.01497401	0.01352018	0.01240970	0.3240156
Period 1 (3M)	ARIMA(2,0,2)	GARCH(1,1)	APARCH(1,1)	Regression
MSE	0.0005606348	3.338360e-04	3.438372e-04	0.0008208638
MAE	0.02127127	0.013742670	0.015204508	0.05710662
Period 2 (3M)	Pure ARIMA(2,0,2)	GARCH(1,1)	APARCH(1,1)	Regression
MSE	0.0019537375	0.0016644569	0.0016768143	0.02703593
MAE	0.03824697	0.03653158	0.03638651	0.3506823
Period 1 (12M)	ARIMA(2,0,2)	GARCH(1,1)	APARCH(1,1)	Regression
MSE	0.0009523015	9.501257e-04	9.526260e-04	0.001311681
MAE	0.02710461	0.026769001	0.027134460	0.0324022
Period 2 (12M)	Pure ARIMA(2,0,2)	GARCH(1,1)	APARCH(1,1)	Regression
MSE	0.0007042058	0.0007023213	0.0007060481	0.6945591
MAE	0.02148884	0.02090367	0.02085000	0.8328744

Table 9: Comparison of forecast accuracy (1M, 3M and 12M ahead)

6.2 Concluding remarks

1. **Univariate approach versus linear regression approach:** The univariate GARCH and APARCH clearly outform the other two models, as they always produce the lowest MAE and MSE for all three forecasting periods. We note that the regression model is not as effective as the univariate approach. This is in line with the results produced in [15], in which univariate

models, APARCH in particular, outperform the model with a list of explanatory variables chosen based on economic intuitions. This suggests that the variables included are not relevant enough to accurately drive oil prices.

2. **APARCH versus GARCH:** As for the two ARCH models we propose, APARCH is more accurate than GARCH in terms of reducing forecast errors in the second period and vice versa. This suggests APARCH is more effective in capturing assymetric volatility dynamics, which is the fact that bear markets tend to spike up volatility more than bull markets, a phenomenon which is more obvious in period 2. Again, this is in line with the mathematical feature of APARCH suggested in [8].

3. **Period 1 versus period 2:** In the second period, the market is evidently under a more turbulent condition, especially during the financial crisis. We can evidently say that there is a structural change from period 1 to period 2. The linear regression model that is fitted provides less satisfactory residual plots. In particular, the normal distribution assumption of residuals looks less convincing, which suggests linear regression models fail to capture tail risks or to respond effectively to tailed events.

4. **Using GARCH/APARCH versus using pure ARIMA:** We note that variance of log return in the second period amounts to 595.99, while that of the first period is only 31.58. Though we also see price fluctuations in the first period, such scale is definitely larger in the second period, especially during the financial crisis where see a significant decline in oil prices from historical highs. In particular, this volatile period makes a significant contribution to the much larger variance of the second period. Accordingly, this provides intuitive reasoning on why GARCH/APARCH is necessary in the second period while a pure ARIMA model is sufficient for the first period.

7 R codes

```
library(TSA); library(MASS)
library(FitAR); library(tseries)
library(timeSeries); library(fGarch)
library(forecast); library(reporttools)

data<-read.csv('~Oilfactors1.csv', header=TRUE)
data
##Univariate approach##
plot(ts(data[,11][1:240],start=c(1993,3),frequency=12),
ylab="Monthly WTI spot price US$")
abline(v=2003)

#Divide it into two 10-year horizons
oilspot.ts1<-ts(data[,11][1:109],start=c(1993,3),frequency=12)
oilspot.ts2<-ts(data[,11][122:230],start=c(2003,3),frequency=12)

#Set up realised values for validation
realised.ts1<-diff(log(ts(data[,11][1:121],start=c(1993,3),frequency=12)))
realised.ts2<-diff(log(ts(data[,11][122:242],start=c(2003,3),frequency=12)))
realised.ts1; realised.ts2

oilspot.ts1;oilspot.ts2;
summary(oilspot.ts1); summary(oilspot.ts2);
head(oilspot.ts1);head(oilspot.ts2);

#BoxCox plots to show the need of log transformation
BoxCox.ts(oilspot.ts1); BoxCox.ts(oilspot.ts2);
plot(log(ts(data[,11][1:240],start=c(1993,3),frequency=12)),
ylab="Monthly WTI spot price US$"); abline(v=2003)

#ADF test indicating non-stationarity
adf.test(log(oilspot.ts1),alternative=c('stationary'))
adf.test(log(oilspot.ts2),alternative=c('stationary'))

#ADF test indicating non-stationarity
adf.test(diff(as.vector(log(oilspot.ts1))),alternative=c('stationary'))
adf.test(diff(as.vector(log(oilspot.ts2))),alternative=c('stationary'))

#log return plot
```

```
plot(diff(log(ts(data[,11][1:240],start=c(1993,3),frequency=12))),
ylab="Monthly WTI spot price US$")
abline(v=2003,h=0)

log.return1<-diff(log(oilspot.ts1))
log.return2<-diff(log(oilspot.ts2))
log.return1;log.return2

#ACF and PACF plots
par(mfrow=c(1,2));
acf(as.vector(log.return1),drop.lag.0=FALSE)
pacf(as.vector(log.return1))

acf(as.vector(log.return2),drop.lag.0=FALSE)
pacf(as.vector(log.return2))

#Write a function to compute AIC of various ARMA orders
AICfn<-function(N,K)
{
for(i in 1:N)
{for(j in 1:N)
{
print(AIC(arima(K,order=c(i-1,0,j-1))))
}}}
AICfn(3,log.return1)
AICfn(3,log.return2)

#Models determined
model1<-arima(log.return1,order=c(0,0,1),include.mean=FALSE)
model2<-arima(log.return2,order=c(2,0,2),include.mean=FALSE)

#Residual analysis
plot(model1$resid);abline(h=0)
plot(model2$resid);abline(h=0)
mean(model1$resid); mean(model2$resid)

#Residual ACF and PACF plots
acf(as.vector(model1$resid),drop.lag.0=FALSE)
pacf(as.vector(model1$resid))
acf(as.vector(model1$resid^2),drop.lag.0=FALSE)
```

```
pacf(as.vector(model1$resid)^2)

acf(as.vector(model2$resid),drop.lag.0=FALSE)
pacf(as.vector(model2$resid))
acf(as.vector(model2$resid^2),drop.lag.0=FALSE)
pacf(as.vector(model2$resid)^2)

#Normality tests
qqnorm(residuals(model1)); qqline(residuals(model1))
qqnorm(residuals(model2)); qqline(residuals(model2))
jarque.bera.test(model1$resid); jarque.bera.test(model2$resid)

#Independence tests
McLeod.Li.test(,model1$resid,gof.lag=20)
McLeod.Li.test(,model2$resid,gof.lag=20)

#Write a function to compute AICs of various GARCH orders
AICfn2<-function(N,K)
{
for(i in 1:(N+1))
{ for(j in 1:N)
{
print(AIC(garch(residuals(K),order=c(i,j-1),trace=FALSE)))
}}}
AICfn2(3,model1); AICfn2(3,model2)

#GARCH model fitted
model.garch2<-garch(model2$resid,order=c(1,1),trace=F)
model.garch2.res<-resid(model.garch2)[-1]
acf(model.garch2.res,drop.lag.0=FALSE)
pacf(model.garch2.res)
acf(model.garch2.res^2,drop.lag.0=FALSE)
pacf(model.garch2.res^2)

#For comparing forecast accuracy, fit GARCH/APARCH models
gfit1<-garchFit(formula=~arma(0,1)+garch(1,1),
data=log.return1,trace=FALSE,include.mean=FALSE)
gfit11<-garchFit(formula=~arma(0,1)+aparch(1,1),
data=log.return1,trace=FALSE,include.mean=FALSE)
gfit2<-garchFit(formula=~arma(2,2)+garch(1,1),
```

```
data=log.return2,trace=FALSE,include.mean=FALSE)
gfit22<-garchFit(formula=~arma(2,2)+aparch(1,1),
data=log.return2,trace=FALSE,include.mean=FALSE)

##Linear regression approach##
cor(data[,2:10][1:108,])
cor(data[,2:10][121:228,])
par(mfrow=c(2,8))

x11<-data[,c(3:10)][(1:108),]
x22<-data[,c(3:10)][(121:228),]
y1<-data[,2][(1:108)]
y2<-data[,2][(121:228)]
x1.ts<-ts(x11,start=c(1993,4),frequency=12)
x2.ts<-ts(x22,start=c(2003,4),frequency=12)
for (i in 1:8) {plot(x1.ts[,i],ylab=i); abline(v=2003)}
for (i in 1:8) {plot(x2.ts[,i],ylab=i); abline(v=2003)}

lm1<-lm(y1~x11[,1]+x11[,2]+x11[,3]+x11[,4]+x11[,5]+x11[,6]+x11[,7]+x11[,8])
lm2<-lm(y2~x22[,1]+x22[,2]+x22[,3]+x22[,4]+x22[,5]+x22[,6]+x22[,7]+x22[,8])
summary(lm1);summary(lm2);

#Carry stepwise AIC for model selection
new.lm1<-stepAIC(lm1); new.lm2<-stepAIC(lm2)
summary(new.lm1);
summary(new.lm2);

#No evidence of multicollinearity from correlation table
cor(x22[,c(2,3,7)][(1:108),])
cor(x11[,c(1,2,5)][(1:108),])
par(mfrow=c(1,2))

#Residual analysis, regression assumptions not violated
plot(new.lm1,c(2))
plot(new.lm2,c(2))
plot(fitted.values(new.lm1),new.lm1$residuals)
plot(fitted.values(new.lm2),new.lm2$residuals)

c<- as.vector(new.lm1$coefficients)
d<- as.vector(new.lm2$coefficients)
```

```
#Generate new matrix containing predicted values of variables

#1-month ahead for period 1
new.z1<-matrix(nrow=1,ncol=3)
new.z1[,1]<-as.vector(predict(x1.ts[,2],h=1)$mean)
new.z1[,2]<-as.vector(predict(x1.ts[,3],h=1)$mean)
new.z1[,3]<-as.vector(predict(x1.ts[,7],h=1)$mean)
#1-month ahead for period 2
new.k1<-matrix(nrow=1,ncol=3)
new.k1[,1]<-as.vector(predict(x2.ts[,1],h=1)$mean)
new.k1[,2]<-as.vector(predict(x2.ts[,2],h=1)$mean)
new.k1[,3]<-as.vector(predict(x2.ts[,5],h=1)$mean)
#3-month ahead for period 1
new.z2<-matrix(nrow=3,ncol=3)
new.z2[,1]<-as.vector(predict(x1.ts[,2],h=3)$mean)
new.z2[,2]<-as.vector(predict(x1.ts[,3],h=3)$mean)
new.z2[,3]<-as.vector(predict(x1.ts[,7],h=3)$mean)
#3-month ahead for period 2
new.k2<-matrix(nrow=3,ncol=3)
new.k2[,1]<-as.vector(predict(x2.ts[,1],h=3)$mean)
new.k2[,2]<-as.vector(predict(x2.ts[,2],h=3)$mean)
new.k2[,3]<-as.vector(predict(x2.ts[,5],h=3)$mean)
#12-month ahead for period 1
new.z3<-matrix(nrow=12,ncol=3)
new.z3[,1]<-as.vector(predict(x1.ts[,2],h=12)$mean)
new.z3[,2]<-as.vector(predict(x1.ts[,3],h=12)$mean)
new.z3[,3]<-as.vector(predict(x1.ts[,7],h=12)$mean)
#12-month ahead for period 2
new.k3<-matrix(nrow=12,ncol=3)
new.k3[,1]<-as.vector(predict(x2.ts[,1],h=12)$mean)
new.k3[,2]<-as.vector(predict(x2.ts[,2],h=12)$mean)
new.k3[,3]<-as.vector(predict(x2.ts[,5],h=12)$mean)

#Compute the forecast errors for regression approach
re.forecast.error11<-new.z%*%c[2:4]+c[1]-data[,2][109:109]
re.forecast.error21<-new.k%*%d[2:4]+d[1]-data[,2][229:229]
re.forecast.error13<-new.z%*%c[2:4]+c[1]-data[,2][109:111]
re.forecast.error23<-new.k%*%d[2:4]+d[1]-data[,2][229:231]
re.forecast.error112<-new.z%*%c[2:4]+c[1]-data[,2][109:120]
```

```
re.forecast.error212<--new.k%*%d[2:4]+d[1]-data[,2][229:240]

#Compute MSE and MAE
mean(abs(re.forecast.error11));mean(abs(re.forecast.error21))
mean(abs(re.forecast.error11)^2);mean(abs(re.forecast.error21)^2)
mean(abs(re.forecast.error13));mean(abs(re.forecast.error23))
mean(abs(re.forecast.error13)^2);mean(abs(re.forecast.error23)^2)
mean(abs(re.forecast.error112));mean(abs(re.forecast.error212))
mean(abs(re.forecast.error112)^2);mean(abs(re.forecast.error212)^2)

#Next, compute the forecast errors for univariate approach
forecast.error1<-vector(length=12); forecast.error2<-vector(length=12);
forecast.error11<-vector(length=12); forecast.error22<-vector(length=12);
forecast.error111<-vector(length=12); forecast.error222<-vector(length=12);

mse.forecast.error1<-vector(length=12); mse.forecast.error2<-vector(length=12);
mse.forecast.error11<-vector(length=12);
mse.forecast.error22<-vector(length=12);
mse.forecast.error111<-vector(length=12);
mse.forecast.error222<-vector(length=12);

for(i in 1:12)
{forecast.error1[i]<-mean(abs(predict(gfit1,n.ahead=i)
$meanForecast-data[,2][109:(108+i)]));
forecast.error11[i]<-mean(abs(predict(gfit11,n.ahead=i)
$meanForecast-data[,2][109:(108+i)]));
forecast.error111[i]<-mean(abs(forecast(Arima(log.return1,
order=c(0,0,1),include.mean=FALSE),
h=12)$mean-data[,2][109:(108+i)]));}

for(i in 1:12)
{mse.forecast.error1[i]<-mean((predict(gfit1,n.ahead=i)
$meanForecast-data[,2][109:(108+i)])^2);
mse.forecast.error11[i]<-mean((predict(gfit11,n.ahead=i)
$meanForecast-data[,2][109:(108+i)])^2);
mse.forecast.error111[i]<-mean((forecast(Arima(log.return1,
order=c(0,0,1),include.mean=FALSE),
h=12)$mean-data[,2][109:(108+i)])^2);}

for(i in 1:12)
```

```
{forecast.error2[i]<-mean(abs(predict(gfit2,n.ahead=i)
$meanForecast-data[,2][229:(228+i)]));
forecast.error22[i]<-mean(abs(predict(gfit22,n.ahead=i)
$meanForecast-data[,2][229:(228+i)]));
forecast.error222[i]<-mean(abs(forecast(Arima(log.return2,
order=c(2,0,2),include.mean=FALSE),
h=12)$mean-data[,2][229:(228+i)]));}

for(i in 1:12)
{mse.forecast.error2[i]<-mean((predict(gfit2,n.ahead=i)
$meanForecast-data[,2][229:(228+i)])^2);
mse.forecast.error22[i]<-mean((predict(gfit22,n.ahead=i)
$meanForecast-data[,2][229:(228+i)])^2);
mse.forecast.error222[i]<-mean((forecast(Arima(log.return2,
order=c(2,0,2),include.mean=FALSE),
h=12)$mean-data[,2][229:(228+i)])^2);}

#Compute MSE and MAE
mean(abs(forecast.error1)); mean(abs(forecast.error11));
mean(abs(forecast.error111)); mean(abs(forecast.error2));
mean(abs(forecast.error22)); mean(abs(forecast.error222));
mean((forecast.error1)^2); mean((forecast.error11)^2);
mean(abs(forecast.error111)^2); mean((forecast.error2)^2);
mean((forecast.error22)^2); mean(abs(forecast.error222)^2);
```

References

[1] G. Box and G. Jenkins, (1970) Time series analysis: Forecasting and control. San Francisco: Holden-Day

[2] P. Brockwell and R. Davis, (1991) Time Series: Theory and Methods. Springer-Verlag

[3] G. M. Ljung and G. Box (1978) On a Measure of a Lack of Fit in Time Series Models. Biometrika 65 (2): 297–303

[4] M. Arranz, (2005) Portmanteau Test Statistics in Time Series. Tol-Project.Org

[5] R. Engle, (1982) Autoregressive Conditional Heteroscedasticity with Estimates of Variance of United Kingdom Inflation". Econometrica 50:987-1008

[6] T. Bollerslev, (1986) Generalized Autoregressive Conditional Heteroskedasticity. Journal of Econometrics, 31:307-327

[7] R. Engle, (2001) GARCH 101: The Use of ARCH/GARCH Models in Applied Econometrics. Journal of Economic Perspectives 15(4):157-168

[8] Ding, Z., C. W. J. Granger, and R. F. Engle (1993). A Long Memory Property of StockMarket Returns and a New Model," Journal of Empirical Finance, 1, 83-106.

[9] D.A. Dickey and W.A. Fuller (1979) Distribution of the estimators for autoregressive time series with a unit root. J. Amer. Statist. Assoc., 74:427–431

[10] H. Akaike (1974) A new look at the statistical model identification. IEEE Trans. Automat. Control, 19(6):716–72

[11] C. Jarque and A. Bera (1980) Efficient tests for normality, homoscedasticity and serial independence of regression residuals. Economics Letters 6 (3): 255–259

[12] A. McLeod and W. K. Li, (1983) Diagnostic checking ARMA time series models using squared residual autocorrelations. Journal of Time Series Analysis, 4, 269273

[13] N. Meade and A. Cooper, (2007) An Analysis of Oil Price Behaviour. Quantitative and Qualitative Analysis in Social Sciences Volume 1, Issue 2, 2007, 33-54

[14] F. Bosler (2010) Models for Oil Price Prediction and Forecasting. San Diego State University

[15] T. Gileva (2010) Econometrics of Crude Oil Markets. Universit´e Paris 1 Panth´eon-Sorbonne

[16] C. Tsirimokos, (2011) Price and Income Elasticities of Crude Oil Demand, the case of ten IEA countries. Swedish University of Agricultural Sciences, Faculty of Natural Resources and Agricultural Sciences, Master Degree Thesis No. 705, ISSN 1401-4084

[17] R. Pirog, (2005) World Oil Demand and its Effect on Oil Prices. CRS Report for Congress, order code RL32530

[18] E. Papapetrou, (2001) Oil Price Shocks, Stock Market, Economic Activity and Employment in Greece. Energy Economics, Vol. 23, pp. 511-532.

[19] Huang, B.N., M.J. Hwang, and Peng, H.P., (2005) The Asymmetry of the Impact of Oil Price Shocks on Economic Activities: An Application of the Multivariate Threshold Model. Energy Economics, Vol. 27, pp. 455-476.

[20] Cunado J. and de Gracia F., (2005) Oil prices, economic activity and inflation: evidence for some Asian countries. Energy Economics, Vol. 45, pp. 65–83.

[21] Guo, H. and. Kliesen, K (2005). 'Oil Price Volatility and U.S. Macroeconomic Activity', Federal Reserve Bank of St. Luis, Review, Vol. 87, No.6, pp. 669-683.

[22] Lehmann, E. L. and Casella, G. (1998) Theory of Point Estimation (2nd ed.). New York: Springer. ISBN 0-387-98502-6. MR 1639875.

[23] Huang, R.D., Masulis, R.W. and Stoll, H.R. (1996) Energy shocks and financial markets. Journal of Futures Markets, 16, 3956.

[24] Kaufmann, R.K., Dees, S., Karadeloglou and P., Sanchez, M. (2004) Does OPEC matter? An econometric analysis of oil prices. The Energy Journal, 25(4), 67-90.

[25] Pindyck, R. S., (2004) Volatility and Commodity Price Dynamics. Journal of Futures Markets 24, 1029–1047.

[26] A. Deng, K. King and D. Metz (2012) An Econometric Analysis of Oil Price Movements: The Role of Political Events and Economic News, Financial Trading, and Market Fundamentals. Bates White Economic Consulting

[27] X. Mu and H. Ye (2010) Understanding the Crude Oil Price: How Important is the China Factor. USAEE-IAEE WP 10-050