Malaysian Tapis: A Closer Look into Additive Outliers and Persistence Volatility

Ida Normaya Mohd Nasir$^{1,2}$, Mohd Tahir Ismail$^1$ and Samsul Ariffin Abdul Karim$^3$

$^1$ School of Mathematical Sciences, Universiti Sains Malaysia, 11800 USM Penang, Malaysia
$^2$ Faculty of Computer Science and Mathematics, Universiti Teknologi MARA (UiTM) Kedah, 08400 Merbok, Kedah, Malaysia
$^3$ Fundamental and Applied Sciences Department, Universiti Teknologi PETRONAS, Bandar Seri Iskandar, 32610 Seri Iskandar, Perak Darul Ridzuan, Malaysia

m.tahir@usm.my

Abstract. As Malaysia is one of the countries, which relies on oil-related revenues for its economic planning, the modelling volatility in crude oil price may therefore have some implications for its policies which the present study intends to look into. Specifically, the main objective of the present study is to examine the impact of outliers, which are identified based on the method proposed by Doornik and Ooms (2005) on the volatility persistence of Malaysian Tapis. The outliers of Tapis have reportedly been associated with unexpected events such as wars, financial crises, etc. The salient findings of the present study revealed that: 1) Tapis may have relatively higher degree of volatility persistence 2) Outliers may potentially bias the estimation of GARCH (1,1) model and 3) Outliers may correct Student-$t$-based specifications and outperform the other models in capturing the volatility. It was on such findings; it can be recommended that outliers should be considered in improving the volatility modelling in Tapis oil markets.

1. Introduction

It is noteworthy that energy commodity is the fundamental driver of most economic activities, which potentially have direct impact on major macroeconomic variables such as gross domestic product (GDP), employment opportunities and tax revenues. In the context of Malaysia, the decline of crude oil price has in turn, reduced the number of associated dividends paid to the government by Petroleum Nasional Berhad (Petronas) from a whopping RM26 billion in the year of 2015 to RM16 billion in the year of 2016 and to an even lower amount i.e., RM13 billion in the following year. To ameliorate this, the government of Malaysia has announced some adjustments in relation to its annual budget and as a result, implemented the 6% Goods and Services Tax (GST) on April 1, 2015.

Tapis crude is a Malaysian crude oil which is mainly produced for export purposes whereas the Malaysian government opts for crude oil which is relatively lower in quality for domestic use. In this regard, Tapis blend is internationally considered as the most expensive oil with the highest quality, extra light and with relatively lower sulphur crude. However, the unpredictable and fluctuating oil price on a global scale especially when the price falls, it affects Tapis. Specifically, during the first half of the year of 2014, the price of Malaysia Tapis was considered high at an estimated USD116 per barrel before falling to USD50 by December of the same year. Such a declining pattern has remained...
in the following year when Tapis recorded its lowest price i.e., an estimated USD38 in December of the year of 2015 (see Figure 1).

It is noteworthy that the variability observed in crude oil prices may depend largely on various geopolitical and economic factors such as wars, financial crisis and natural disasters. In this regard, such unexpected events are often considered as outliers. The presence of such outliers may affect the volatility model, in relation to; the estimation of parameters \([1-3]\), the test of conditional homoscedasticity \([1,4]\) and the out-of-sample volatility forecasts \([2,4,5]\). To the best of the researchers’ knowledge, no related studies have to date been carried out to look into the impact of outliers on the modelling of Tapis crude oil volatility. Thus, modelling volatility in Tapis oil price by means of taking the impact of such unexpected events on a global level into consideration may have some relevance for Malaysia as an emerging and developing country.

It is of utmost importance to assess how these unexpected shocks may potentially impact on the volatility over time, and more importantly, if the shocks are persistent or short-lived. In order to facilitate such an investigation, Autoregressive conditionally heteroscedastic (ARCH) models introduced by Engle \([6]\) which was later extended to generalized ARCH (GARCH) by Bollerslev \([7]\) have been developed. It has to be noted that persistence may occur in the first or higher order moments of a time series. In this regard, the persistence in the first moment, or levels, of a time series can be verified by means of either unit root tests or stationarity tests to the levels, while the persistence in the volatility of the time series is usually demonstrated by means of a highly persistent fitted GARCH model.

It is worth highlighting that the present study has three main objectives; 1) to detect the presence of additive outliers within Tapis return series 2) to examine the impact on the estimation of volatility and 3) to access the persistence of volatility within the Tapis. The plan of the paper is presented in the following sections. Section 2 describes the statistical procedures involved for detecting additive outliers and the measure of persistence of volatility. Moving on, the data set and the empirical results are discussed in section 3. Lastly, section 4 provides a summary and conclusion.

2. Data and methodology

2.1. Data

The dataset of the present study consists of daily closing spot prices for the Malaysian Tapis. The price is available from Thomson Financial Datasync and is expressed in US dollars per barrel. The data set spanning from December 1, 1987 to Dec 31, 2015 (a total of 7589 observations) and which has the advantage to cover various unexpected events such as 1990 Iraq–Kuwait war, 1997 Asian financial crisis and the 2008 global financial crisis, among others. The price series are converted into the logarithmic return series to obtain stationary series. Let \( p_t \) denotes daily spot prices of Tapis and \( r_t \) denotes corresponding daily return:

\[
\log p_t = \log p_t - \log p_{t-1} \quad \text{for } t = 1, 2, \ldots, T
\]

in which \( r_t \) is the return for each index at time \( t \), \( p_t \) is the current price, and \( p_{t-1} \) is the price from the previous day. The return series can be converted with the following conditional mean and variance dynamics equation:

\[
r_t = \delta + \varepsilon_t \quad \text{where } \varepsilon_t = \eta_1 \sqrt{h_t}
\]

where \( \delta \) is the conditional mean, \( \varepsilon_t \) is independent and identically distributed with \( N(0,1) \) and \( h_t \) is the conditional variance, which can be estimated with GARCH type models. Plot of original Tapis series and return series is given demonstrated in Figure 1. There is a strong evidence of volatility
clustering, in which, where large price movements on one day are followed by similarly large movements on successive days similarly large movements on successive days follow large price movements on one day.

**Figure 1.** Time series plot of Tapis crude oil market

2.2. Persistence of volatility

To estimate the volatility persistence in Tapis, the researchers employed the GARCH (1,1) model, which was developed by Bollerslev [7]. This model is the continuation of Engle's [6] autoregressive conditional heteroskedastic (ARCH) model for time-varying volatility in a time series. The conditional variance can be described as follows:

\[ h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 h_{t-1} \]  

in which, \( \alpha_0 \) denotes the constant term, \( \varepsilon_{t-1}^2 \) denotes the ARCH term, \( h_{t-1} \) denotes the GARCH parameter where \( \alpha_0 > 0, \alpha_1 > 0 \) and \( \beta_1 \geq 0 \). The process is stationary if \( \alpha_1 + \beta_1 < 1 \).

2.3. Outlier detection method

In the present study, the researchers employed Doornik & Ooms [8] procedure, which Chen & Liu [9] employed to detect the presence of additive outliers in GARCH (1,1) with errors following a Gaussian or a Student-t distribution. This likelihood-based procedure and associated test are asymptotically similar with the GARCH parameters \( \alpha_1 \) and \( \beta_1 \). The approach involves several steps as the following:

**Step 1.** Estimate the baseline model of GARCH model to obtain the log likelihood \( \hat{l}_g \), residuals \( \varepsilon \), and volatility \( h \) and the largest standardized residuals in absolute value, \( \max|\varepsilon_t/h_t| \). Denote this observation by \( t = s \).

**Step 2.** Re-estimate a GARCH generalized additive outlier (i.e., gao) model with dummy variable, \( d_t \), (which corresponds to the date of the largest standardized residual, \( \max|\varepsilon_t/h_t| \) obtained in Step 1) in the mean equation, as well as adding a single-observation lagged dummy variable \( d_{t-1} \) in the variance equation. This will provide estimates for the added parameters, \( \gamma_{gao,t} \) and \( \tilde{\gamma}_{gao,t} \) and a new log-likelihood \( \tilde{l}_{gao,t} \). The GARCH generalized additive model maybe written as:

\[ r_t = \delta + \gamma d_t + \varepsilon_t \text{ where } \varepsilon_t = \eta_t \sqrt{h_t} \]
\[ h_t = \alpha_0 + \sum_{i=1}^p \alpha_i e_{t-i}^2 + \sum_{j=1}^q \beta_j h_{t-j} + \tau d_{t-s} \text{ for } t = 1, 2, K, T \] (5)

where \( d_t \) is defined as before, such that \( d_{t-s} \) equals one when \( t = s + 1 \) and zero otherwise.

**Step 3.** Detect the potential outliers in the series. The largest standardized residuals absolute value \( \max |e_t|/h_t \) is an outlier date if \( 2(\hat{t} - \hat{i}) > C^\alpha_\tau \). Doornik & Ooms [8] suggested that an approximation of the asymptotic distribution under the null-hypothesis of no outliers is \( C_\tau = 5.66 + 1.88 \log(T) \), where \( T \) indicates the sample size, at significance \( \alpha \) of 5%. If the outlier is confirmed, it will be corrected by the new value obtained by means of forecasting. Notably, the procedure is recursive until no further outlier is detected. Following the work of Behmiri & Manera [10], in which, the focus is only on additive outlier detection. Therefore, the above procedure is deemed sufficient without further classification between additive volatility outliers (AVO) and additive level outliers (ALO).

3. Results and discussion

3.1. Descriptive statistics

Some key statistics for the raw data is presented in Table 1. Based on Table 1, Tapis oil return series have high kurtosis, suggesting that the series are leptokurtic (the presence of fat-tails distribution) and negative skewness statistics (-0.051) indicate that the series has a slightly longer left tail (extreme losses) than right tail (extreme gain). Moreover, the Jarque–Bera (JB) statistics are statistically significant and thus, suggesting that the distribution of Tapis is not standard normal, which may be due to the presence of extreme observations. The residual diagnostics suggest that there is an ARCH effect in the series. The Ljung-Box serial correlation (Q-stats) shows the presence of serial correlation up to 20 lags. The Tapis return series is stationary as indicated by Augmented Dickey and Fuller (ADF) and the Phillips and Perron (P-P) unit root tests.

| Summary statistics | Mean (%) | SD | Min | Max | Skewness | Kurtosis | JB test |
|-------------------|---------|----|-----|-----|----------|----------|---------|
| 0.005             | 0.008   | -0.079 | 0.097 | -0.051 | 12.548   | 29007.70** |

| Diagnostics | ARCH F-test (lag 20) | ADF test | Phillips-Perron test |
|-------------|----------------------|----------|----------------------|
|             | Constant             | Constant and trend | Constant |
|             |                     |           | Constant and trend |
| 124.591**   | 70.797**             | -90.429** | -90.433** |
|             |                      | -91.083** | -91.076** |

** indicates that the null hypothesis is rejected at 5% level. ARCH LM test: Null hypothesis- No ARCH effect; Ljung Box serial correlation test (Q-statistics): Null hypothesis- No serial correlation. ADF and PP tests: Null hypothesis- Return has a unit root.

3.2. Outliers in Tapis oil market

Table 2 present the number of identified outliers in the GARCH models using Doornik & Ooms [8] approach for both normal and Student-t distribution. It should be highlighted that the impact of a known and unknown event on the time series was accessed in the present study. Hence, the potential outlier was identified through various unpredictable events such as Iraq-Kuwait conflict, political crises and financial crisis, which by and large may affect the volatility of Tapis.
The identified outliers are arranged in chronological order and the impact of these outliers, the percentage change is also presented. In addition, the events that correspond to each of the outliers were specified. Notably, the outliers were clustering due to several major events such as the invasion of Kuwait by Iraq, Operation Desert Storm, Asian financial crisis, Russian political crisis, Libya crisis and Israel-Palestine crisis.

Nonetheless, some outliers were identified as isolated outliers, i.e., unexpected oil spills and OPEC announcement on the production cut. It was expected that the oil price would increase following the announcement. Loutia, Mellios, and Andriosopoulos [11] posited that the cut decision has a stronger weightage and was much awaited when price fluctuations are lower. However, the OPEC members may disagree with the decision due to various interests. It is noteworthy that Venezuela, a country with highly subsidized fuel was most affected by the increase of oil price, in which, the petrol costed almost nothing in that country.

The present study reveals that the volatility of Tapis is more affected by major event such as war than other factors. In contrast to the West Texas Intermediate (WTI), Tapis was not affected by any of the U.S announcements on crude inventories (Charles & Darné [3]) suggesting that Tapis is less volatile.

3.3. Persistence of volatility
The present study examines the persistence of volatility of Malaysia Tapis by means of the “original data” and the “outlier corrected data” and compares the estimation of the models under normal and student-t distribution. In order to evaluate the most appropriate model, Akaiake Information Criteria (AIC) and Hannan-Quin (HQ) criteria were employed. The best model was made to appear in bold, representing the lowest value of AIC and HQ. Also, the residual tests demonstrated the appropriateness of the model.

The estimation results of GARCH with normal and Student-t distribution were shown in Table 3. It was found that the value of $\alpha_1$ decreases and the value of $\beta_1$ increases when the data is corrected for outliers. Carnero, Peña, & Ruiz [4,12,13] explicate that these results may be due to some patch of outliers, which affected the estimation of volatility parameter. Also, it was found that the sum of $\alpha_1$ and $\beta_1$ were high, which indicated the high degree of persistence in volatility. Notably, the results were consistent with other similar studies on crude oil return series [3,14], stock market [5] and metal price [9]. Moreover, based on the model criteria, removing outliers may improve the performance in capturing volatility of Tapis crude oil return series for both normal and Student-t distribution. The GARCH model, which employed outlier, corrected data-Student-t distribution showed superior results and was measured by the lowest value of AIC and HQ criteria to the rest.

| Date of outlier | Model | Tapis return (% $\Delta$) | Events |
|----------------|-------|---------------------------|--------|
| 23 Dec 1986    | A, B  | +6.376                    | OPEC agreement to cut production by seven percent for the first six months of 1987. |
| 2 Jan 1987     | A     | +8.125                    | OPEC announcement to cut production. |
| 6 Jan 1987     | A     | +2.312                    | OPEC announcement to cut production. |
| 21 Aug 1987    | A     | -5.500                    | Gulf war. |
| 17 Dec 1987    | A     | -4.420                    | OPEC meeting failed to agree on new production quotas. |
| 19 Dec 1989    | A     | +5.025                    | Iranian tanker KHARK 5 was damaged in a storm, exploded and caught fire in the north east of the Canary Islands. |
| 03 Dec 1990    | A     | -9.604                    | Iraq experiments the effectiveness of explosives on selected oil wells. |
| 17 Jan 1991    | A, B  | -16.667                   | Operation Desert Storm was initiated. Many oil wells in Kuwait were destroyed and set on fire. |
19 July 1991 A +0.990 The start of the failed revolution against Mikhail Gorbachev, and the beginning of the end for the Soviet Union.
24 July 1991 A -0.980 Greek tanker Kirki lost its bow off the coast of Western Australia.
6 Aug 1991 A +0.990 Removal of Russian President Mikhail Gorbachev.
9 Aug 1991 A -0.784 Boris Yeltsin was appointed as the president of Russia.
28 April 1997 A, B +8.378 Asian financial crisis.
18 Mar 1998 A +9.756 OPEC agreement to cut production.
19 May 1998 A -7.843 India announced that it has conducted five underground nuclear tests in the southwest of New Delhi, near Pakistani border.
24 Feb 2011 A +8.714 Political turmoil in Libya limited the oil production from the North African countries.
6 May 2011 A -10.903 Fear of US, economic slowed down.
27 Aug 2015 A +10.917 Venezuela requested OPEC to hold an emergency meeting.

Note: A denotes GARCH (1,1) normal dist, B denotes GARCH (1,1) Student-t dist

Table 3. Estimation result of GARCH (1,1)

| Parameter       | Gaussian Original data | Outlier-corrected | Student-t Original data | Outlier-corrected |
|-----------------|------------------------|-------------------|-------------------------|-------------------|
| $\alpha_1$      | 0.0526                 | 0.0435            | 0.0226                  | 0.0771            |
| $\beta_1$       | 0.9220                 | 0.9219            | 0.9167                  | 0.8147            |
| $\alpha_1 + \beta_1$ | 0.9746                 | 0.9654            | 0.9393                  | 0.8918            |
| $AIC$           | -7.2367                | -7.3485           | -7.5105                 | -7.5264           |
| $HQ$            | -7.2354                | -7.3473           | -7.5090                 | -7.5249           |
| $ARCH F$-stat   | 1.3402                 | 3.8850**          | 0.0078                  | 0.0267            |
| $Q$-stat (lag 20) | 92.82**               | 99.467**          | 109.33**                | 131.58**          |

** indicates that the null hypothesis is rejected at 5% level. ARCH LM test: Null hypothesis-No ARCH effect; Ljung Box serial correlation test (Q-statistics): Null hypothesis-No serial correlation.

4. Conclusion
The present study assessed the impact of outliers in modelling the volatility of Tapis, a type of Malaysian oil. Doornik & Ooms [8] procedure was employed to identify both unknown and known date of outliers. Interestingly, it is found that Tapis is less volatile and highly affected during the war period such as the invasion of Kuwait by Iraq and Operation Desert Storm. The other outliers were considered isolated.

It is noteworthy that the main findings of the present study are: (i) outliers can bias the estimation of parameters of the GARCH model; (ii) removing outliers improves the performance of the models; (iii) the model using the original data–Student-t distribution outperforms the model using the outlier corrected data–normal distribution; (iv) Tapis return show a high degree of persistence in volatility (i.e., might suffer from structural change). Therefore, the volatility modelling may improve the accuracy of the model, considering the presence of the outliers. Nevertheless, further study on the performance of the model using out-of-sample criteria, which takes into account the outlier and structural change is deemed necessary.

Acknowledgments
The authors would like to extend their sincere gratitude to the Ministry of Higher Education Malaysia (MOHE) for the financial supports received for this work under FRGS grant (203/PMATHS/6711604).

References

[1] van Dijk, D., Franses, P. H., & Lucas, A. (1999). Testing for ARCH in the Presence of Additive Outliers. Journal of Applied Econometrics, 14(October 1996), 539–562.

[2] Charles, A. (2008). Forecasting volatility with outliers in GARCH models. Journal of Forecasting, 27(7), 551–565.

[3] Charles, A., & Darné, O. (2014). Volatility persistence in crude oil markets. Energy Policy, 65, 729–742. http://doi.org/10.1016/j.enpol.2013.10.042

[4] Carnero, M. A., Peña, D., & Ruiz, E. (2007). Effects of outliers on the identification and estimation of GARCH models. Journal of Time Series Analysis, 28(4), 471–497.

[5] Franses, P. H., & Ghijsels, H. (1999). Additive outliers, GARCH and forecasting volatility. International Journal of Forecasting.

[6] Engle, R. F. (1982). Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. Econometrica, 50(4), 987–1007. http://doi.org/10.2307/1912773

[7] Bollerslev, T. (1986). Generalized Autoregressive Conditional Heteroskedasticity. Journal of Econometrics, 31(1), 307–327.

[8] Doornik, J. A., & Ooms, M. (2005). Outlier Detection in GARCH Models Outlier Detection in GARCH Models.

[9] Chen, C., & Liu, L. (1993). Joint Estimation of Model Parameters and in Time Series Outlier Effects. Journal of the American Statistical Association, 88(421), 284–297.

[10] Behmiri, N. B., & Manera, M. (2015). The role of outliers and oil price shocks on volatility of metal prices. Resources Policy, 46, 1–12.

[11] Louitia, A., Mellios, C., & Andriosopoulos, K. (2016). Do OPEC announcements influence oil prices? Energy Policy, 90, 262–272. http://doi.org/10.1016/j.enpol.2015.11.025

[12] Carnero, M. A., Peña, D., & Ruiz, E. (2001). Outliers and conditional autoregressive heteroscedasticity in time series.

[13] Carnero, M. A., Peña, D., & Ruiz, E. (2012). Estimating GARCH volatility in the presence of outliers. Economics Letters, 114(1), 86–90.

[14] Salisu, A. A., & Fasanya, I. O. (2013). Modelling oil price volatility with structural breaks. Energy Policy, 52, 554–562. http://doi.org/10.1016/j.enpol.2012.10.003