403 Final Project

CRUDE OIL PRICE FORECASTING

Yuxuan Zhao Tianjue Tao Gary Vartanian Ozgur Cetinok
004880206 305037681 804195815 105054619

Introduction

The main purpose the paper is to predict crude oil prices using time series techniques. This is a challenging task because empirically, oil prices seem to follow random walk (German 2017, Hamilton 2009). There have been many successful attempts to beat the random walk over long horizon (Kilian 2014, Benjamin 2015). In this paper, we focus on short term forecast.

The motivation for the paper is that oil is one of the most traded commodities, countries rise and fall with oil prices. For example, if the oil price were to fall, the US benefits as an oil importer. Foreign countries that export oil will lose. The recent appreciation in the US dollar will amplify the effect by reducing the costs for the US oil import. This will leave a profound impact on the global market. Although there are conflicting evidences on the effect of oil prices in financial markets (Faff and Brailsford 1999), it is generally agreed upon that that oil prices have asymmetric impact on stock markets (Arouri, 2011, Hamilton 1996). Real industry such as mining and natural gas also seem to be significantly influenced by oil prices (Bert & Cenk 2012). Therefore, the ability to forecast oil prices is crucial in understanding global economy.

For this paper, we use monthly data on oil prices from 1974 to 2017. We use both nominal and real prices. The nominal price is West Texas Intermediate spot price from quandl. The real price is the inflation adjusted price using 2016 dollar as a benchmark. The data can be downloaded from the U.S Energy Information Administration.

We use ARIMA, ARIMA+GARCH to make short term forecast on oil price changes. We first taking the log transformation then taking seasonal difference. In this way, we reduce the influence from trend and seasonality. Box test shows that the time series is stationary. We then fit the series to an ARMA model. ACF and PACF shows that there is no structure left in the residual, however, the variance of the residual seems to follow a certain pattern. Finally, we model the dynamics in variation by assuming the variance follows an ARMA process. Diagnostic statistics tells us that our model fully captures the dynamics in the mean and variance of the time series.

The general result is that our model has an in-sample forecast error at 6% and out of sample forecast error at 9%. When compared to random walk, our model outperforms for monthly and quarterly forecast, but underperforms for longer horizon. Back testing and robustness check show that our model is unbiased for short horizon but has tendency to underestimate for longer horizon. Although the model passes stability test, (CUSUM plot appears to be randomly distributed around 0) further investigation on MSPE reveals that our model does not fully capture the 2008 financial crisis and the 2014 oil price crush. Although our model was able to predict the price drop in 2014 we cannot capture the magnitude of the price drop. Baumeister and Kilian (2015) shows that their model is able to predict the 2014 financial crisis. However, their model requires ex-post information of varies time series. It is unknown whether the model will be able to predict the next financial crisis. From Baumeister and Kilian, we learned that we can improve the stability of long-term forecasting by adding macroeconomic global oil production. Alternatively, we can use the weighted average of different forecast results to make a more accurate forecast.

Data.

The nominal price is monthly data on West Texas Intermediate (WTI)spot price from 1974M01 to 2017M12. Although there are other oil prices such as Brent Crude (ICE), the price difference is only about 0.03% . Such a small difference will hardly change the conclusion from our forecast model. For the sake of simplicity, we use WTI as measurement of oil prices.

While attempting to improve forecast error, we notice that the variation in price level induces noise in oil prices. In order to eliminate the noise, we use real price instead of nominal price. The real price is the inflation adjusted price using 2016 dollar as a benchmark. It is calculated by apply the formula below to all nominal prices. All data can be downloaded from the U.S Energy Information Administration.

$$\frac{Nominal\ Price \times Base\ Year\ CPI}{Current\ CPI}$$

Finally, we use monthly US commercial gas price to build VAR model. The data is from 1974 to 2017. It can be accessed from Quandl.

Methodology

For the mean model, we assume that the log of oil price follows general ARMA process.

$$Y_t = \phi_1 Y_{t-1} + \cdots + \phi_p Y_{t-p} + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q} + \varepsilon_t$$
$$\varepsilon_t \sim WN(0, \sigma^2)$$

The model can be simplified to.

$$\Phi(L) Y_t = \Theta(L) \varepsilon_t$$

$$\Phi(L) = 1 - \phi_1 L - \phi_2 L^2 - \cdots - \phi_p L^p$$

$$\Theta(L) = 1 + \theta_1 L + \theta_2 L^2 + \cdots + \theta_q L^q$$

The intuition for the model is that the current oil price Y_t can be found from past values Y_{t-k} and random shocks ε_t. The core assumption is that our series has been detrended and seasonally adjusted such that the ARMA process captures the clean effect of cycle. It is also crucial that the time series is covariance stationary. The diagnostic statistics in the appendix shows that the log of the price is stationary.

For variance model, we use GARCH(1,1) process to model variance, specifically we define:

$$Y_t = \mu_{t|t-1} + \varepsilon_t = \mu_{t|t-1} + \sigma_{t|t-1} z_t,$$
$$\sigma^2_{t|t-1} = \omega + \alpha\varepsilon^2_{t-1} + \beta\sigma^2_{t-1|t-2}$$

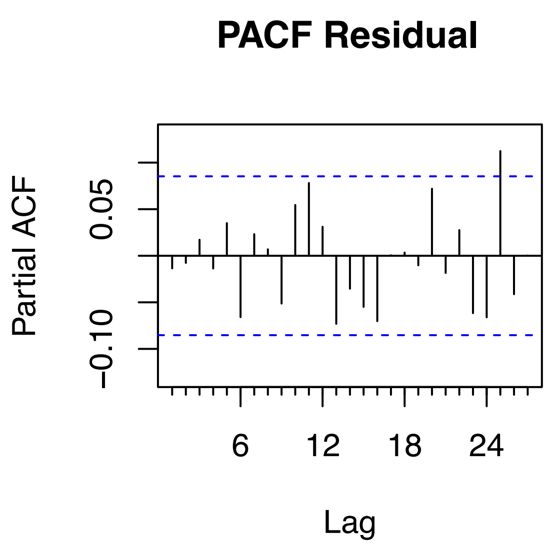
Here $\mu_{t|t-1}$ is the conditional mean and ε_t is a white noise process such that

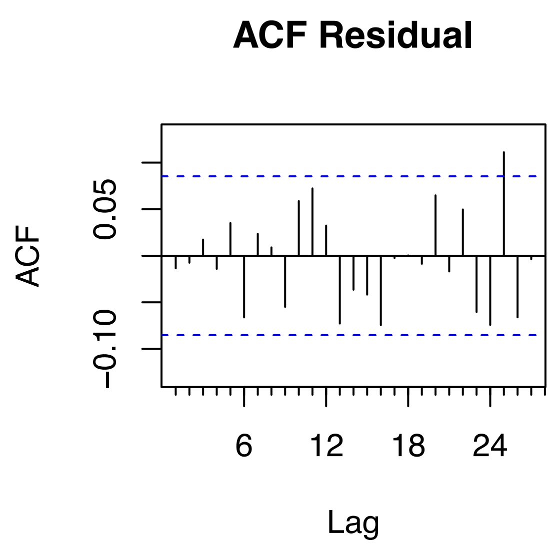
$$\varepsilon_t = \sigma_{t|t-1} z_t \; z_t \sim WN(0,1)$$

where $\omega>0$, $\alpha\geq 0$, and $\beta\geq 0$.

The intuition for GARCH process is that the variance today depends on the most recent volatility $\sigma^2_{t-1|t-2}$ and random shock ε^2_{t-1}. Similar to ARMA model, we need to make sure that the series is stationary before implementing the GARCH process. The diagnostic statistics in the appendix shows that our model captures all the dynamics in the variance and there are no apparent structure breaks.

Our identification strategy relies on ACF and PACF. The goal is to reduce the residual of the model to white noise. We first decompose the time series to remove all trend and seasonality factors. To account for cycles, we fit the detrended series to an ARMA process. We then model the volatility using GARCH. Finally, we visually check the ACF and PACF of residuals. check whether the residuals are white noise by performing Box-Ljung test.





Box-Ljung test is a test for correlation. The test statistic is $Q* = T(T+2)\sum(T-k)^{-1} r_k^2$
Where T is the length and R_k is the kth autocorrelation coefficient of the residuals. It can be compared against a χ^2 distribution. The P value for the residual is 0.756, therefore, we fail to reject the null hypothesis that the residual is white noise.

Statistic Test

Figure 1 The Original Data of Crude Oil Price

To extract useful information from time series, it is required to make sure it is covariance stationary, so we did ADF test before we ffitted the model. P-value bigger than 0.05 in the result shows that this series is not covariance stationary.

```
##
##  Augmented Dickey-Fuller Test
##
## data:  ts.oil
## Dickey-Fuller = -1.9254, Lag order = 6, p-value = 0.608
## alternative hypothesis: stationary
```

Figure 2 The Result of ADF Test

Then we took first difference to make it stationary. We did ADF test again and got stationary series. Figure 3 shows the result that P-value is pretty small that we can reject the null hypothesis.

```
## Warning in adf.test(difts.oil): p-value smaller than printed p-value
##
##  Augmented Dickey-Fuller Test
##
## data:  difts.oil
## Dickey-Fuller = -7.7298, Lag order = 6, p-value = 0.01
## alternative hypothesis: stationary
```

Figure 3 The Result of ADF Test on First-Difference Series

Although stationary, the series may be useless for prediction if it is a white noise series, so we used Box-Ljung test which shows it is not a white noise. This enabled us to deeply analyze the series and fit the model.

```
##
##  Box-Ljung test
##
## data:  difts.oil
## X-squared = 47.864, df = 1, p-value = 4.568e-12
```

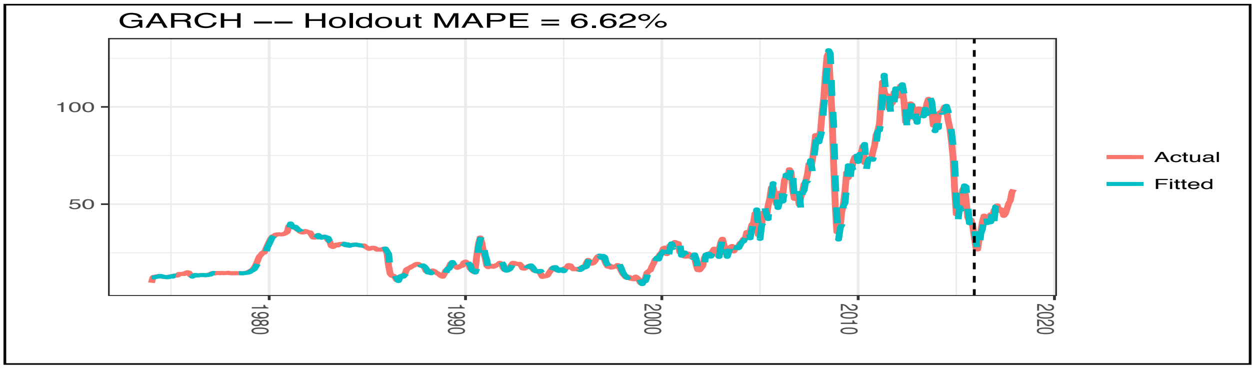
Figure 4 The Result of Box-Ljung Test

.

Forecast Evaluation

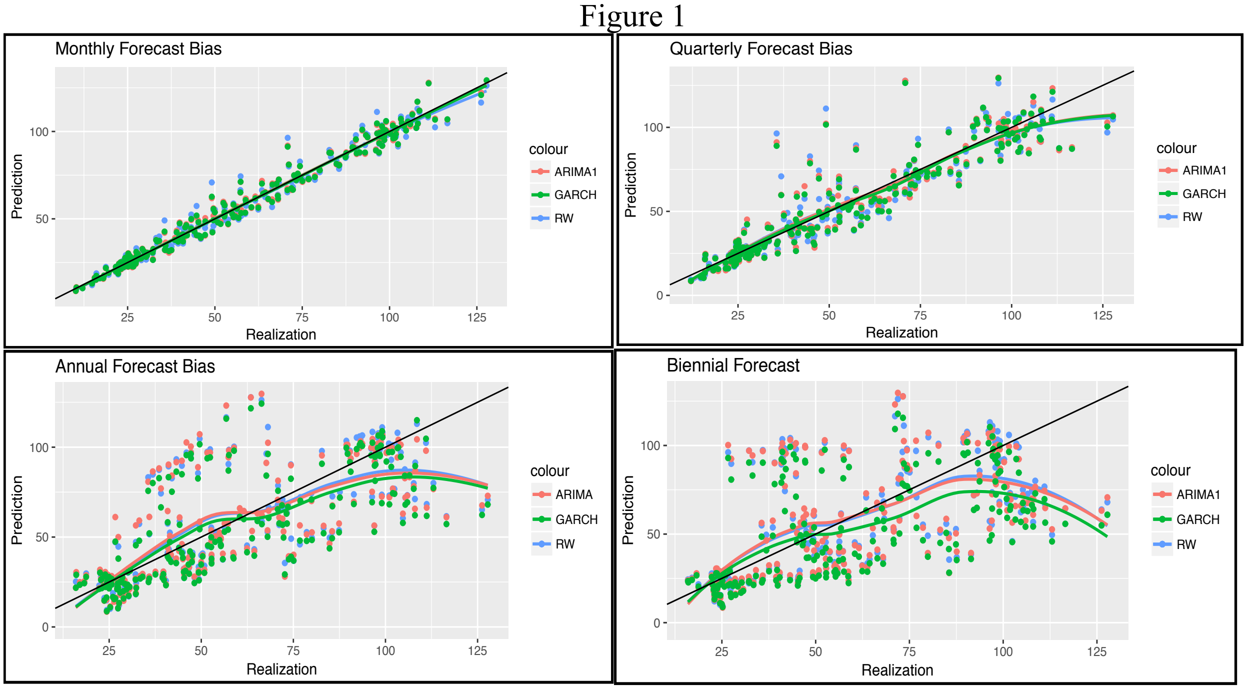
General Result

The mean absolute percentage error for the in sample monthly forecast is about 6%. This means that our forecast is only 6% away from the true value. The out of sample forecast predicts oil prices in 2018 Jan and Feb to be 57.80 and 57.70 respectively. The actual monthly oil prices are 63.70 and 62.21 respectively. The out of sample mean absolute percentage error is around 9.2%.

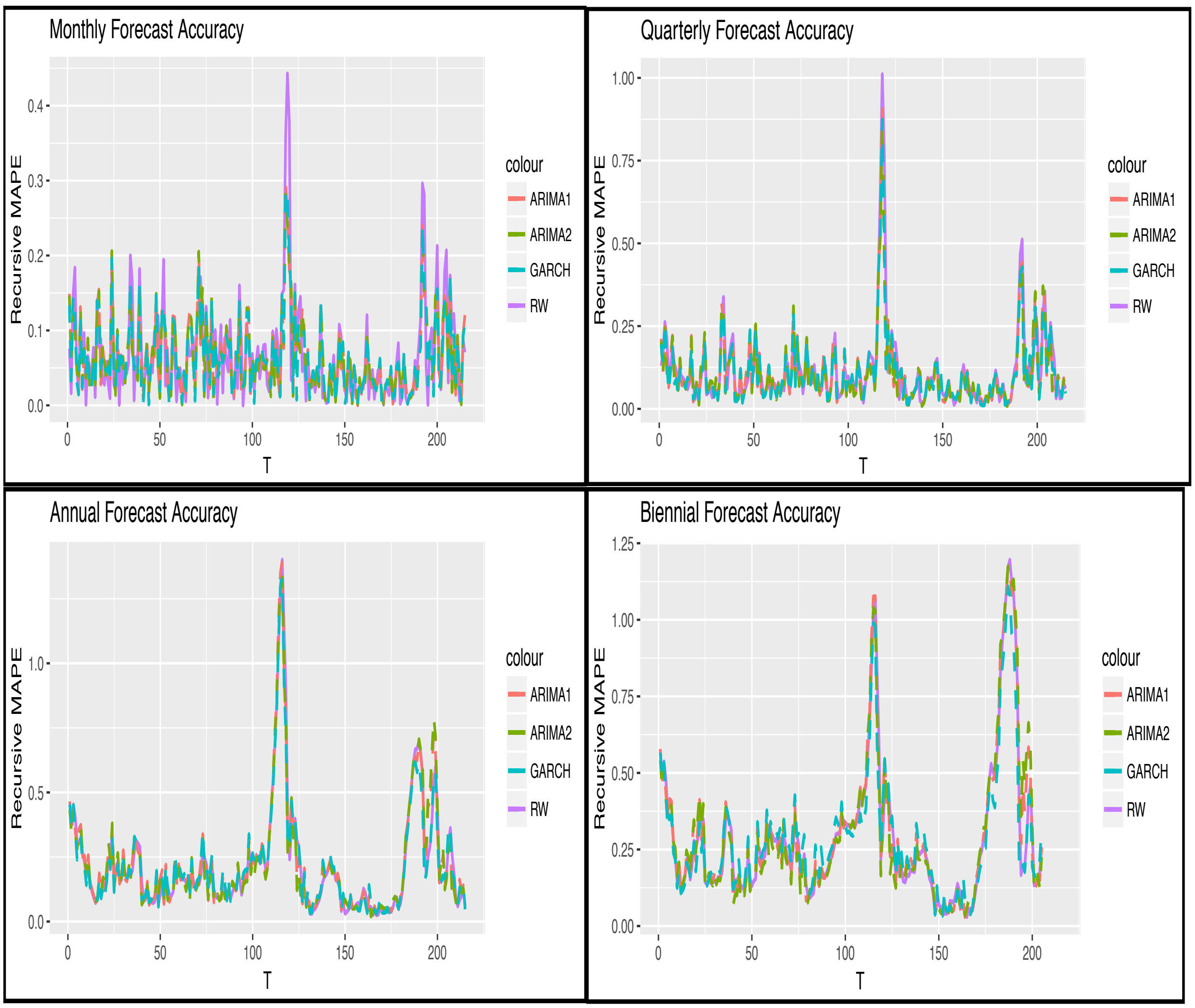


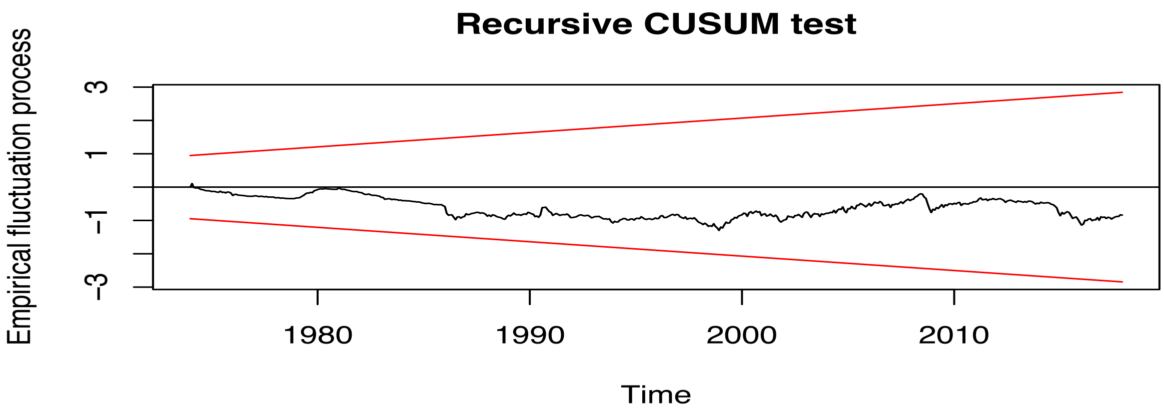
In this section, we evaluate the unbiasedness and accuracy of the forecast. For unbiasedness, we follow the standard approach by Mincer and Zarnowitz (1969). We use a recursive scheme to calculate out of sample forecast. Figure 1 shows prediction and corresponding realizations for ARIMA, ARIMA+GARCH, and the benchmark Random Walk. Although bias seem to be small (captured by points away from the 45-degree line) for monthly and quarterly forecast. Our forecast tends to under estimate real oil price. (Captured by the number of points below 45-degree line). For annual or biennial forecast, our model significantly underestimates oil prices. Compared to random walk, our model outperforms random walk on monthly and quarterly basis, but underperforms for longer forecasting horizon.

Figure 1

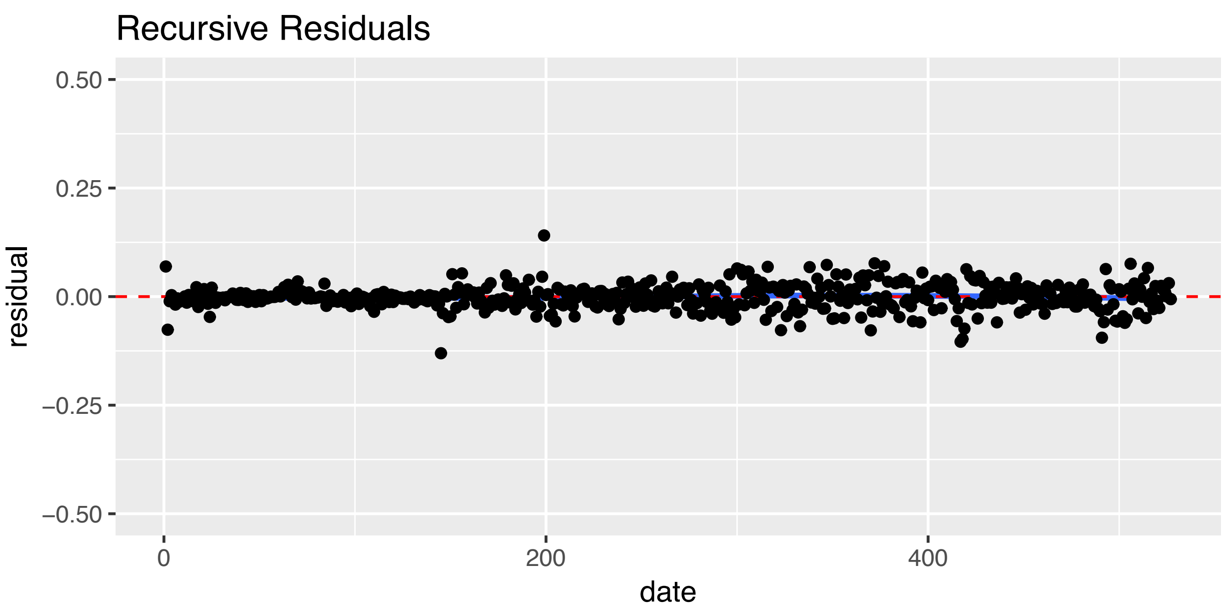


For accuracy, we calculate out of sample mean absolute percentage error recursively and plot them against each iteration. The forecast errors are mostly below 10% except 2 prominent spikes at 2008 and 2014. When compared to random walk, our model outperforms on short horizon and underperforms for long horizon. Overall, the model has outstanding performance over historical data, however, from an ex ante point of view, our model may be off by a considerable margin as new data arises. Although the model does not appear to have structural break. (Figure 2) (Figure3) the prominent spike in recursive MAPE plot (Figure 4) seem to suggest that we need multiple model for different time periods. Our next step is to build a dynamic model that can better accommodate structural break.





Since CUSUM plot is within the bounds and appear to be randomly moving around line 0, our model is within control.



Recursive residuals center around 0

Improve long-term Forecast
In this section, we want to improve long-term forecast by introducing the gas price in the U.S The reason why we chose American gas price is that the fluctuations and trend of two time series are similar.

Before further exploration, we accessed to the data of American gas price, took first difference to make it covariance stationary (it is non-stationary before taking first difference), and did Box-Ljung test showing the series can be used econometrically.

Considering the feasibility of VAR model, cointegration and granger test are required. We did cointegration test which shows a small P-value which indicates they are cointegrated. Due to the result that two time series are I(0) and cointegrated, there is a long-term equilibrium between variables. Plus, we assumed granger causality might be also involved.

```
adf.test(lm.oil.gas$residuals)

## Warning in adf.test(lm.oil.gas$residuals): p-value smaller than printed p-
## value

##
##  Augmented Dickey-Fuller Test
##
## data:  lm.oil.gas$residuals
## Dickey-Fuller = -4.4553, Lag order = 6, p-value = 0.01
## alternative hypothesis: stationary
```

Figure 4 The Result of Cointegration Test

We expected that the coefficients of granger test on two variables could be statistically significant, especially the regression of global crude oil price on American gas price, so that we could built a VAR model. Unfortunately, we found only crude oil price affects gas price, which means that the variable, American gas price, is not qualified to predict crude oil price.

```
grangertest(difts.gasdata~difts.oil, order=10)

## Granger causality test
##
## Model 1: difts.gasdata ~ Lags(difts.gasdata, 1:10) + Lags(difts.oil, 1:10)
## Model 2: difts.gasdata ~ Lags(difts.gasdata, 1:10)
##   Res.Df  Df     F   Pr(>F)
## 1    291
## 2    301 -10 5.528 1.54e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

grangertest(difts.oil~difts.gasdata, order=1)

## Granger causality test
##
## Model 1: difts.oil ~ Lags(difts.oil, 1:1) + Lags(difts.gasdata, 1:1)
## Model 2: difts.oil ~ Lags(difts.oil, 1:1)
##   Res.Df Df      F Pr(>F)
## 1    318
## 2    319 -1 1.7458 0.1874
```

Figure 5 The Result of Granger test on both variables

Therefore, the VAR model is not feasible. We found another variable, the total American vehicle sale. However, we got the same result as what crude oil price and gas price showed us. After trying to find some potential factors affecting crude oil price, we that more global and financial data was better, such as the gold price and political issues. This topic remains to be the future work for our project

Conclusion

We were able to forecast monthly oil prices with 94% accuracy. Back testing shows that our forecast is unbiased on short horizon. However, our model tends to underestimate prices for long horizon. Although our model is more accurate than random walk on short horizon, the accuracy decreases as time increase. The biggest challenge we have is predicting 2014 oil crush. This is because accommodating financial crisis needs ex-post information. The cause behind the 2008 financial crisis and 2014 crisis is drastically different. Adding time series such as CDS price will improve 2008 in sample prediction but provide little information on the next crisis. In the future, we will combine our model with State Space Model to better accommodate the financial crisis.

Reference:

Baumeister, C. and L. Kilian, “Real-Time Forecasts of the Real Price of Oil,” *IMF Economic Review, Palgrave Macmillan*, vol. 62(1), (2014). pp. 119-145

Baumeister, C. and L. Kilian, “Understanding the Decline in the price of Oil since June 2014”, forthcoming in the *Journal of the Association of Environmental and Resource Economists*.

Bert Scholtens, Cenk Yurtsever, “Oil price shocks and European industries” *Energy Economics*, vol. 34(4), (2012). pp. 1187-1195

James D. Hamilton, "This is what happened to the oil price-macroeconomy relationship," *Journal of Monetary Economics, Elsevier*, vol. 38(2), (1996). pp. 215-220,
James D. Hamilton, “Understanding Crude Oil Prices” *The Energy Journal* Vol. 30, No. 2 (2009), pp. 179-206

Mohamed El Hedi Arouri “Does crude oil move stock markets in Europe? A sector investigation.” *Econ. Modelling,* 28 (2011), pp. 1716-1725

Mincer, J. A., and V. Zarnowitz, “The Evaluation of Economic Forecasts,” NBER Chapters in *Economic Forecasts and Expectations: Analysis of Forecasting Behavior and Performance,* (1969), pp.1–46.

Robert W. Faff, Timothy J. Brailsford, “Oil price risk and the Australian stock market” *J. Energy Finance* Dev., (1999). pp. 69-87