What Influences the Fluctuation of Global Oil Price?

Analytics for Applied Economics and Business

Group 30

Michael McDonnell, Die Tao, Manfei Cui, Jiaxuan Chen, Fengxu Gao, Floris ter Haar

Introduction

Crude Oil is and has been one of the most influential commodities in the 21st and 20th centuries. It not only accounts for the lion-share of many countries' energy mix, but is also crucial in the transporting of goods and people around the world. (Perkins, 2020) Globalisation and the ever-increasing interconnectedness of the world economies and markets is in part due to the exploitation of crude oil. As a result, Oil has a very significant impact on world economic growth. In the US, the country we will be focusing our analysis on, the oil and gas sector alone accounts for roughly 8% of GDP. (American Petroleum Institute, 2019) Fluctuations in crude oil prices are known to have profound effects on the economic well-being at the individual, corporate as well as national levels. (Magnussen, 2000) A study conducted by the IMF found that a 10% increase in oil prices leads to roughly a 0.2% decrease in global GDP. (Mussa, 2000) It is for these reasons that it is important and interesting to understand what affects the price of oil. As a result, in this paper we will aim to answer the following research question:

What influences the fluctuation of global oil price?

Movements in crude oil prices are influenced by various factors. Factors include, changes in oil production, economic growth, and the price of alternative energy resources to name a few. We will be conducting regression analyses of the WTI index, over a 10-year period, on our explanatory variables and provide analysis and interpretations on the data that is significant to answering our research question. We will be analysing variables that we have classified into three distinct groups. We will be looking at the supply of oil, the demand of oil as well as the relative competitiveness of substitutes. We will also employ a dummy variable in our analysis which will explore the effect of seasonality on our dependent variable.

Hypothesis

In terms of the supply variables, we expect increases in the supply of oil to be negatively correlated with oil price. We base this assertion on the simple economics mechanisms that illustrate that there is an inverse relationship between supply and price. We expect demand to be positively correlated with price, as the demand for oil increases so will the price. In terms of the competitiveness of substitutes we expect as substitutes get more competitive (lower price) then the price of oil will also decrease. Lastly, we expect our dummy variable will demonstrate that oil prices will increase during summer as there is increased travel and economic activity.

Methods and Data

i. Methods

We use OLS regression to test the effect of oil price on supply-side and demand-side data and add one variable in each step for more comprehensive interpretations. We also use VIF test to test the multicollinearity between independent variables.

ii. Data

The data are monthly from January 2010 to December 2019. The start end begins in 2010 because we want to avoid the strong influence of the 2008 financial crisis on our regression results. We use monthly data, instead of quarterly data, to overcome the limitation of few observations over a 10-year period. Oil-related data is obtained from EIA. Data of Industrial Production Index, disposable income and unemployment rate is obtained from Federal Reserve Bank of St. Louis. Data of natural gas price is obtained from IndexMundi.

Dependent variable: Oil Price - WTI spot price $(monthly)

We used the WTI oil price figures as it is one of the largest and most reliable indexes used in the US.

Explanatory variable 1: OPEC Crude oil production

As a commodity, the supply of crude oil has a direct and decisive impact on the price of crude oil. As the world's most influential oil exporting organisation, OPEC was set up with the aim of avoiding competition among members and thus reaping the benefits of a consistent oil policy to control oil production and exports. This is why we choose OPEC's crude oil production as the supply-side measure of crude oil.

Explanatory variable 2: Industrial Production Index (IPI)

IPI Measures real output in the manufacturing, mining, electric, and gas industries. These industries are highly energy-intensive and thus form the majority of corporates' demand. So, we use this index as a variable from the demand side to test the relationship between oil price and market demand from the company level.

Explanatory variable 3: Disposable Income Per Capita

Measures individual's consumption ability, thus reflecting market demand trend. We use this variable to test the relationship between oil price and market demand from the individual level.

Explanatory variable 4: Unemployment Rate

In the US measures the economy's ability to generate jobs for jobseekers and thus can reflect the macroeconomic environment. Therefore, we use the unemployment rate to see if there is correlation between labour market performance and crude oil price.

Explanatory variable 5: Price of Natural Gas

Natural gas is a substitute of crude oil. We use this variable to assess the role of substitutes on oil price.

Explanatory variable 6: Seasonality

We also employed a dummy variable to explore the effect of seasonality of oil price. When the dummy variable is equal to 1 this relates to the summer months of June-September.

we are dealing with time series...how about some stationarity tests?

Results and Discussion

Step 1:

We first run a regression of WTI oil price on production, to analyse the relationship between oil supply and price:

$$WTI = \beta 0 + \beta 1\ Prod + \varepsilon$$

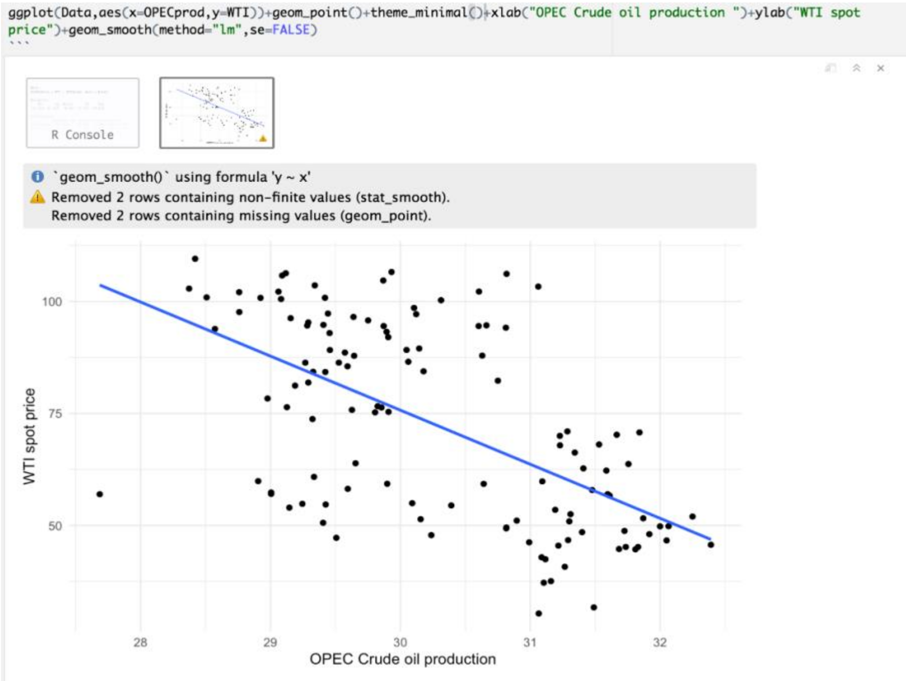
```
```{r}
reg1=lm(WTI~OPECprod, Data)
summary(reg1)
```
```

```
Call:
lm(formula = WTI ~ OPECprod, data = Data)

Residuals:
    Min      1Q  Median      3Q     Max 
-46.713 -11.359   0.486  12.493  40.420 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)  437.867     46.540   9.408 5.05e-16 ***
OPECprod     -12.072      1.536  -7.857 2.03e-12 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 17.82 on 118 degrees of freedom
  (2 observations deleted due to missingness)
Multiple R-squared:  0.3435,	Adjusted R-squared:  0.3379 
F-statistic: 61.73 on 1 and 118 DF,  p-value: 2.034e-12
```



The estimated coefficient of oil price on OPEC production (-12.072) is significant and negative. It means that if the OPEC production increases by one unit (million barrels per day), the oil price would lead to a $12 decrease, which is aligned with expectation. This confirms our assertion that oil price follows the normal economic rules of supply, namely, as supply increases the price will decrease. It is also important to note here that the supply of oil (production) amongst the OPEC nations is subject to strategic as well as geopolitical factors. The OPEC group has often been accused of working as a cartel to fix oil prices at certain levels by regulating the amount of oil production. (Lioudis, 2020) They would for example lower oil production well below maximum capacity in order to ensure higher oil prices. This behaviour is very much explained by our model.

The adjusted R-squared is equal to 0.3379, thus approximately 33% of the model is explained by OPEC production. The significance of the model is further reinforced by the F-Test value of 61.73, indicating that our explanatory variable is significantly different from zero.

Step 2:

We then add the IPI variable to the regression to assess the effect of the demand of oil on its price.

$$WTI = \beta 0 + \beta 1 Prod + \beta 2\, IPI + \varepsilon$$

```{r}
reg2=lm(WTI~OPECprod+IPI, Data)
summary(reg2)

#The estimated coefficient on IPI is equal to -1.3586.

```

```
Call:
lm(formula = WTI ~ OPECprod + IPI, data = Data)

Residuals:
   Min     1Q Median     3Q    Max 
-34.36 -12.36  -0.96  14.07  35.44 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) 532.6778    49.9408  10.666  < 2e-16 ***
OPECprod    -10.5942     1.4953  -7.085 1.12e-10 ***
IPI          -1.3586     0.3419  -3.974 0.000123 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 16.8 on 117 degrees of freedom
  (2 observations deleted due to missingness)
Multiple R-squared:  0.4215,	Adjusted R-squared:  0.4117 
F-statistic: 42.63 on 2 and 117 DF,  p-value: 1.239e-14
```

The addition of the IPI variable leads to an increase in the estimated coefficient on the OPEC production variable, hence there was a downwards bias in the previous regression model. The IPI variable is negative and significant at the 1% level. The coefficient explains that a one unit increase in the IPI index leads to a fall in the WTI price by $1.36.

The scenario could be when IPI index increases, demand from companies rise, thus leading to increase in oil production, lowering the price. Another scenario could be when IPI index increases, those manufacturing companies have a higher output because they innovate their operation processes including improving the efficiency in producing oil, thus decreasing the cost of oil production and lowering the price.

However, there could be a reverse causality between oil price and IPI. If oil price increases, the raw material cost of manufacturing companies increases so they are more likely to scale down their operation, or return to alternative energy, such as natural gas (which will be examined in Step6), thus influencing their output.

Step 3:

We add yet another variable which is an indicator of the demand for oil:

$$WTI = \beta0 + \beta1Prod + \beta2\ IPI + \beta3Dinc + \varepsilon$$

```{r}
reg3=lm(WTI~OPECprod+IPI+Dinc, Data)
summary(reg3)

#Adding the disposable income variable has resulted in a postive IPI variable, while the OPECprod variable remains negative. The estimated coefficient on the disposable income is negative and significant.
```

```
Call:
lm(formula = WTI ~ OPECprod + IPI + Dinc, data = Data)

Residuals:
    Min      1Q  Median      3Q     Max 
-41.791  -6.019   0.289   9.281  30.153 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) 169.5285    61.2921   2.766  0.00661 ** 
OPECprod     -5.9915     1.3447  -4.456 1.94e-05 ***
IPI           3.3581     0.6590   5.096 1.36e-06 ***
Dinc         -6.1175     0.7755  -7.888 1.88e-12 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 13.61 on 116 degrees of freedom
  (2 observations deleted due to missingness)
Multiple R-squared:  0.6235,	Adjusted R-squared:  0.6138 
F-statistic: 64.03 on 3 and 116 DF,  p-value: < 2.2e-16
```

The addition of the Disposable income variable leads to a further increase in the estimated coefficient on the OPEC production variable. The addition also turns the estimate of IPI from negative to positive, indicating that there could be a downward bias in step 2. There could be a correlation between IPI and Dinc, which we will test in Step4 using an interaction between these two variables. The Dinc variable is negative and significant at the 1% level. The coefficient explains that a one unit increase in the Dinc index leads to a fall in the WTI price by $6.12.

The explanation could be an increase in disposable income indicates an upward trend in the macroeconomic environment and an increase in individual demand for fuel, thus leading to the decrease in oil price.

Step 4:

To test the correlation between IPI and Dinc, we add an interaction between these two variables, IPI*Dinc.

$$WTI = \beta 0 + \beta 1 Prod + \beta 2\, IPI + \beta 3 Dinc + \beta 4\, IPI * Dinc + \varepsilon$$

```{r}
reg4=lm(WTI~OPECprod+IPI+Dinc+IPI*Dinc, Data)
summary(reg4)
```

```
Call:
lm(formula = WTI ~ OPECprod + IPI + Dinc + IPI * Dinc, data = Data)

Residuals:
    Min      1Q  Median      3Q     Max 
-40.786  -5.403   1.803   9.836  29.564 

Coefficients:
             Estimate Std. Error t value Pr(>|t|)   
(Intercept) 929.05756  311.09008   2.986  0.00345 **
OPECprod     -3.38145    1.68261  -2.010  0.04681 * 
IPI          -4.11933    3.07356  -1.340  0.18281   
Dinc        -29.15458    9.28970  -3.138  0.00216 **
IPI:Dinc      0.20735    0.08333   2.488  0.01427 * 
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 13.32 on 115 degrees of freedom
  (2 observations deleted due to missingness)
Multiple R-squared:  0.6427,	Adjusted R-squared:  0.6303 
F-statistic: 51.72 on 4 and 115 DF,  p-value: < 2.2e-16
```

According to the result, β4 = 0 can be rejected at the 5% level of significance (P-value = 0.014). This result shows that there is an interaction between disposable income and IPI. By introducing Dinc, the downward bias of IPI is reduced.

```{r}
summary(Data$Dinc)
```

```
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's 
  35.82   39.34   42.39   42.60   45.77   50.32       2 
```

Based on the summary, if Dinc is at its median value of 42.39, the slope of the regression function relating oil price and IPI is predicted to be -4.12+0.21*42.39=4.78. This means that increasing IPI by one unit is expected to increase oil price by $4.78. For the 75% quantile, the estimated change in price of a one-unit increase in IPI is estimated by -4.12+0.21*45.77=5.49, so the slope is somewhat steeper. The interpretation is that for a fixed disposable income level at $45770 (the unit of disposable income is 000'), a one-unit increase in IPI is expected to increase oil price by $5.49.

The scenario could be, as we mentioned in Step1, OPEC often regulates Oil production at certain levels to ensure higher prices. Thus, when the industrial output increases, companies' demand for oil rises, with a fixed production level set by OPEC, the oil price will increase.

Step 5:

$$WTI = \beta 0 + \beta 1 Prod + \beta 2\, IPI + \beta 3 Dinc + \beta 4\, URate + \varepsilon$$

Besides, we should also consider the macroeconomic situation in the US. In view of data availability, we choose the unemployment rate (shown as Urate) as the index to reflect economic development. Run the regression as follows:

```
```{r}
reg5=lm(WTI~OPECprod+IPI+Dinc+Urate, Data)
summary(reg5)
#There could be a reverse causality whereby an increase in oil price would lead to an increase in
unemployment, as GDP is known to decrease when an increase oil price occurs.
```

Call:
lm(formula = WTI ~ OPECprod + IPI + Dinc + Urate, data = Data)

Residuals:
    Min   1Q Median   3Q   Max
-35.934 -8.078   1.762   7.397 20.445

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) -573.8675   102.8912 -5.577 1.64e-07 ***
OPECprod   -0.2934   1.2773 -0.230   0.819
IPI   5.8418   0.6064   9.634 < 2e-16 ***
Dinc   -1.1851   0.8622 -1.375   0.172
Urate   16.9926   2.0696   8.211 3.63e-13 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 10.86 on 115 degrees of freedom
  (2 observations deleted due to missingness)
Multiple R-squared: 0.7626,	Adjusted R-squared: 0.7544
F-statistic: 92.37 on 4 and 115 DF, p-value: < 2.2e-16
```

The coefficient of Urate is 16.99 which is significant, indicating that one percentage point of increase in unemployment rate will lead to the increase of oil price by $16.99 per barrel. However, we will also find that the coefficient of disposable income per capita turns insignificant.

There could also be the problem of reverse causality, whereby an increase in oil price would lead to an increase in unemployment rate. Rising global oil price is known to have a direct positive effect on the operating costs for any business. So, some firms may have to lay-off employees to drive costs down.

Step 6:

As an alternative energy sources for crude oil, we introduce the natural gas into the regression (shown as Ngprice) below.

$$WTI = \beta0 + \beta1 Prod + \beta2\, IPI + \beta3 Dinc + \beta4\, URate + \beta5\, Ngprice + \varepsilon$$

```{r}
reg6=lm(WTI~OPECprod+IPI+Dinc+Urate+Ngprice, Data)
summary(reg6)
```

```
Call:
lm(formula = WTI ~ OPECprod + IPI + Dinc + Urate + Ngprice, data = Data)

Residuals:
    Min      1Q  Median      3Q     Max
-33.808  -6.149   1.688   7.005  21.695

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) -604.2140   102.0859  -5.919 3.49e-08 ***
OPECprod       0.4394     1.2985   0.338   0.7357
IPI            5.6831     0.6005   9.464 4.92e-16 ***
Dinc          -0.8355     0.8622  -0.969   0.3346
Urate         16.7578     2.0377   8.224 3.55e-13 ***
Ngprice        3.3392     1.5014   2.224   0.0281 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 10.67 on 114 degrees of freedom
  (2 observations deleted due to missingness)
Multiple R-squared:  0.7725,    Adjusted R-squared:  0.7625
F-statistic: 77.42 on 5 and 114 DF,  p-value: < 2.2e-16
```

The results show that one extra increase of natural gas price will bring about \$3.34/mmbtu increase in the crude oil price, which is significant.

However, we found that the coefficient on OPEC production changes from negative to positive, which is not consistent with the initial assumptions. 5 variables may be multicollinear.

To test the degree of multicollinearity in the model, we conduct a variance inflation factor on the regression containing all explanatory variables, using the vif() command.

```{r}
reg6=lm(WTI~OPECprod+IPI+Dinc+Urate+Ngprice, Data)
vif(reg6)
```

```
 OPECprod       IPI      Dinc     Urate   Ngprice
 1.991385  8.145411 13.598567 18.476810  1.694883
```

Based on the results, it can be seen that the vif values for both disposable income and unemployment rate are greater than 10, and it can be determined that there is multicollinearity present in the model.

We also used the cor() function to check the correlation coefficients between the respective variables：

```
cor(Data %>% select(Urate,Dinc),use="complete.obs")
```

```
          Urate       Dinc
Urate  1.0000000 -0.9561977
Dinc  -0.9561977  1.0000000
```

The results show an almost perfect negative correlation between the unemployment rate and disposable income. To counter the problem of collinearity, we can drop the unemployment rate variable due to positive interaction effects between the IPI and Dinc variables, which helps to reduce the bias in our model. After dropping the Urate variable, we inputted a vif function on the regression, which returned vif values lower than 8 on all variables. The presence of multicollinearity has decreased in the model, as all values are less than 10.

```{r}
reg7=lm(WTI~OPECprod+IPI+Dinc+Ngprice, Data)
vif(reg7)
```

```
OPECprod      IPI     Dinc  Ngprice
1.470007 6.102245 7.407749 1.690335
```

The results of the regression model after excluding the unemployment rate show that the coefficient of OPEC production also becomes negative, which is in line with our hypothesis.

```{r}
reg7=lm(WTI~OPECprod+IPI+Dinc+Ngprice, Data)
summary(reg7)
```

```
Call:
lm(formula = WTI ~ OPECprod + IPI + Dinc + Ngprice, data = Data)

Residuals:
    Min      1Q  Median      3Q     Max
-39.162  -6.661  -0.011   8.354  35.132

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 121.1303    64.6014   1.875 0.063323 .
OPECprod     -5.0245     1.4020  -3.584 0.000498 ***
IPI           3.2099     0.6532   4.914 2.98e-06 ***
Dinc         -5.6197     0.7997  -7.027 1.59e-10 ***
Ngprice       3.9789     1.8844   2.112 0.036894 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 13.42 on 115 degrees of freedom
  (2 observations deleted due to missingness)
Multiple R-squared:  0.6376,	Adjusted R-squared:  0.6249
F-statistic: 50.57 on 4 and 115 DF,  p-value: < 2.2e-16
```

Step 7:

Lastly, we employed a dummy variable to test the potential seasonal fluctuations of oil price.

$$WTI = \beta 0 + \beta 1 Prod + \beta 2\, IPI + \beta 3 Dinc + \beta 4 Ngprice + \ \beta 5\ Summer + \varepsilon$$

```
```{r}
reg8=lm(WTI~OPECprod+IPI+Dinc+Ngprice+Summer, Data)
summary(reg8)
```
```

```
Call:
lm(formula = WTI ~ OPECprod + IPI + Dinc + Ngprice + Summer,
    data = Data)

Residuals:
    Min      1Q  Median      3Q     Max
-38.560  -7.353   0.190   7.667  35.771

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) 125.8384    65.0749   1.934 0.055623 .
OPECprod     -5.1498     1.4160  -3.637 0.000416 ***
IPI           3.1839     0.6556   4.857 3.83e-06 ***
Dinc         -5.5894     0.8026  -6.964 2.24e-10 ***
Ngprice       3.9309     1.8896   2.080 0.039742 *
Summer        1.8733     2.6248   0.714 0.476889
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 13.44 on 114 degrees of freedom
  (2 observations deleted due to missingness)
Multiple R-squared:  0.6392,    Adjusted R-squared:  0.6233
F-statistic: 40.39 on 5 and 114 DF,  p-value: < 2.2e-16
```

We used a dummy variable whereby the months of June-September were equal to 1 to see if oil prices fluctuated during summer. We found that during the summer months, oil prices were roughly 1.87 dollars higher than the rest of the year. This is likely due to an increased demand for oil as a result of increased travel, as we stipulated in our hypothesis. Interestingly, this increased demand for crude oil in transport must thus outweigh the increased demand for power generation for households during winter. This could be due to the fact that power generation can come from multiple sources whereas the majority of transport is powered through crude oil.

Conclusion

Crude oil has played an extremely important role in the economic development and people’s well-being. In this case, we first analyse the relationship between crude oil production and price in the US. And then, considering the confounding factors, we add the variables, run the regressions, assess the significance and decide whether or not to drop the variables. Finally, we get the regression below:

$$WTI = 125.84 - 5.15OPECprod + 3.18IPI - 5.59Dinc + 3.93Ngprice + 1.87Summer$$

The result is quite meaningful for us to understand the underlying factors that lead to the changes in the crude oil price and even the extent of impacts.

However, the oil price mechanism is a complex system and contains too many factors to analyse them all. Perhaps the most important confounding factors we were unable to analyze are those short-term influences. An emergent political crisis, a sudden announcement of a government energy policy, a market intervention by the government or international institution, short-term flows of global capital, etc. We cannot put these into consideration as it's hard in nature to illustrate them with consistent data. Even for the data we used, the complexity lies in the fact that many of them are correlated. Our recommendation for future analysis is to add more observations for a more reliable result, to find appropriate instrument variables to solve endogeneity problem, and also to add more variables in the regression (i.e., U.S. dollar index, U.S. shale oil production) for a more comprehensive model.
(2425 words)

Bibliography

Amedeo, K., 2020. *What affects oil prices?* Available at: https://www.thebalance.com/how-are-oil-prices-determined-3305650#citation-3 [Accessed 20 November 2020].

American Petroleum Institute, 2019. Oil & Natural Gas Contribution To U.S. Economy Fact Sheet. [Online]
Available at: https://www.api.org/news-policy-and-issues/taxes/oil-and-natural-gas-contribution-to-us-economy-fact-sheet#:~:text=America's%20oil%20and%20natural%20gas,here%20at%20home%20every%20year. [Accessed 30 November 2020].

Lioudis,N., 2020. *OPEC's Influence on Global Oil Prices.* Available at: https://www.investopedia.com/ask/answers/060415/how-much-influence-does-opec-have-global-price-oil.asp [Accessed 20 November 2020].

Mussa, M., 2000. The Impact of Higher Oil Prices on the Global Economy, s.l.: IMF.

Perkins, R., 2020. Oil's share of global energy mix continues to retreat as renewables surge: BP. [Online]
Available at: https://www.spglobal.com/platts/en/market-insights/latest-news/electric-power/061720-oils-share-of-global-energy-mix-continues-to-retreat-as-renewables-surge-bp [Accessed 22 November 2020].

Smith,T., 2020. *Do Oil and Natural Gas Prices Rise And Fall Together?* Available at: https://www.investopedia.com/articles/active-trading/032515/do-oil-and-natural-gas-prices-rise-and-fall-together.asp [Accessed 20 November 2020].