Group Assignment
MSc Finance & Big Data 2021-2022

Applied Data Science

Transitory and permanent shocks in the global market for crude oil, and the role of speculative trading.

Group members:

Arnaud GUIRAUD

Basma EL KHAMLICHI

Mehdi GHAOUTI

Satnam SINGH

Mohamad DAKKOURI

Thibault MOREL

Table of Contents

Introduction 2

i. Identify the determinants of crude oil price fluctuations, and in particular the potential role of speculative trading. 3

ii. Assess how shocks to these determinants affect the crude oil price. 9

.... 19

iii. Evaluate the extent to which the role of these determinant changes over time. 20

iv. Shed light on the evolution of the real price of crude oil since 1974 towards 2008 21

Introduction

All along this document we will focus to understand the mechanisms of oil price variation over the last 50 years. We will study the impacts of transitory and permanent shocks and the role of the different determinant.

To do so, we will answer 4 main topics:

- Identify the determinants of crude oil price fluctuations, and in particular the potential role of speculative trading.
- Assess how shocks to these determinants affect the crude oil price.
- Evaluate the extent to which the role of these determinant changes over time.
- Shed light on the evolution of the real price of crude oil since 1974 towards 2008
.

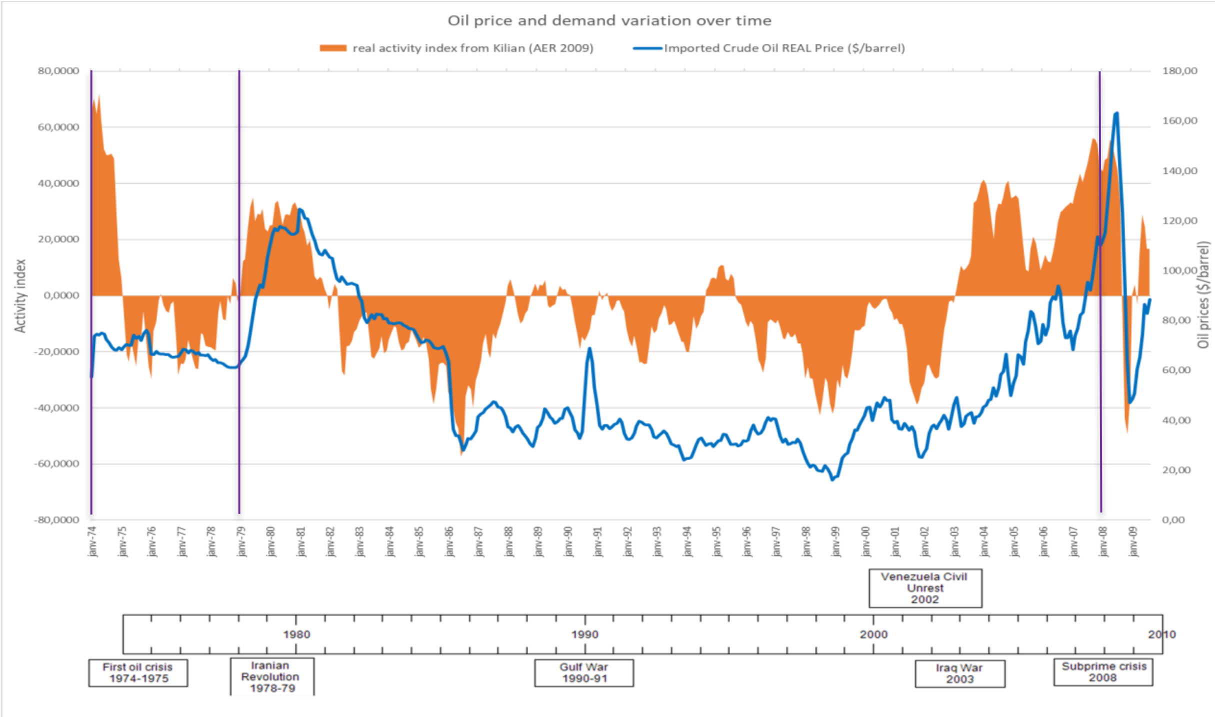
i. Identify the determinants of crude oil price fluctuations, and in particular the potential role of speculative trading.

Historically, the oil price has been driven by a global strong demand, following the global economic growth. Supply and demands are fundamental drivers of oil price.
However, over time it can be observe that sometimes other drivers took places among the fundamental ones.
This is the case with speculation.

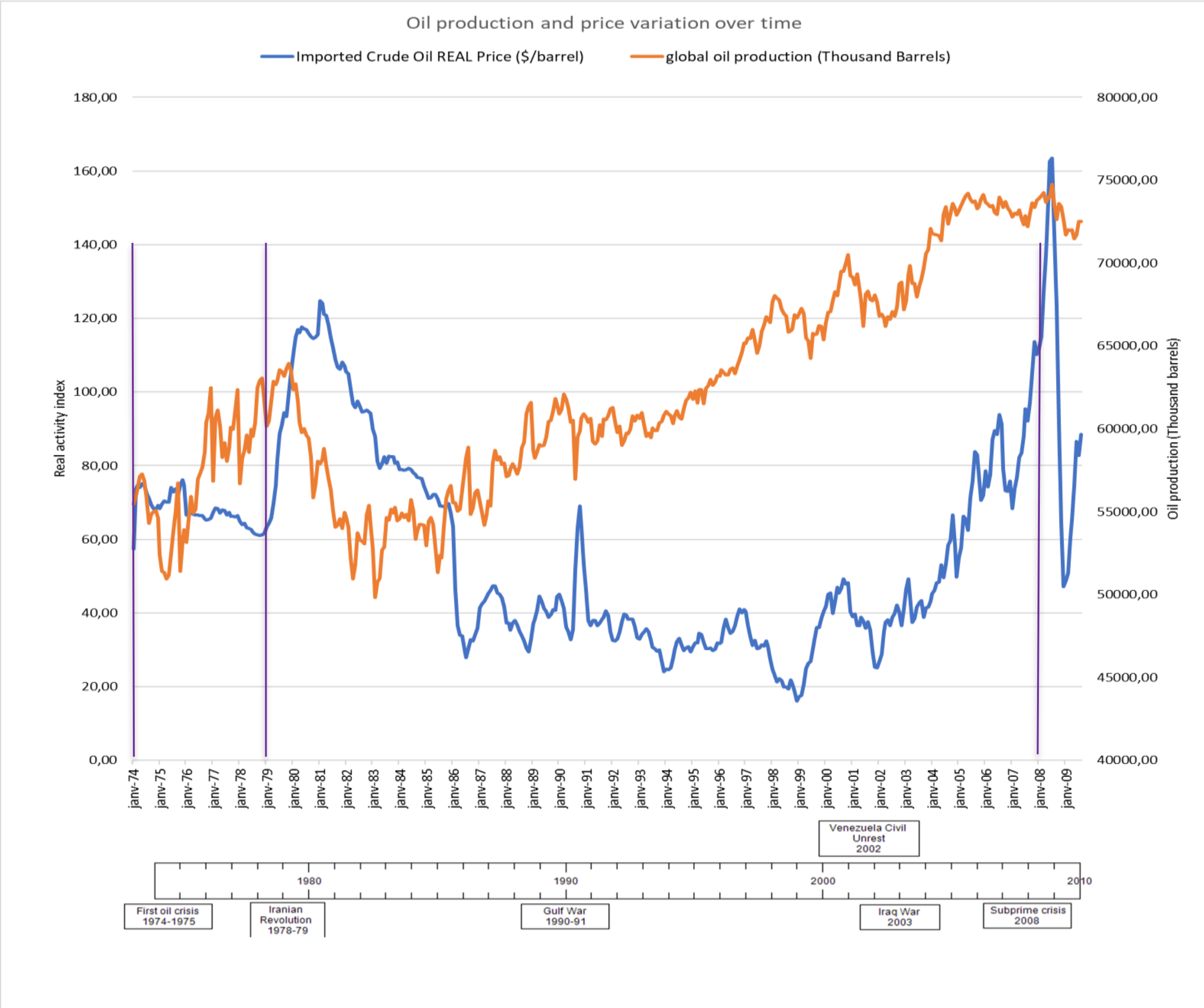
To answer this 1st question and illustrate our explanations or assumptions, we plotted 4 graphs that we present here under:

- Oil prices vs demand,
- Oil prices vs production,
- Oil prices vs inventories,
- Oil production vs demand,
- Oil production vs inventories.

We also added a chronological timeline to our plots to put in perspective oil prices vs macroeconomic or geopolitics events. Then the 3 vertical lines represent the 3 major oil crises: 1973, 1979, 2008.



<u>Comments:</u>
This graph represents yearly oil prices and demand variations over time, from 1974 to 2009.
We can observe that usually, oil prices rise when demand increases and vice-versa.
The demand follows the needs of oil and raw materials during specific periods.
For example, on the period 2002 – 2007, one reason of this price increase was the important global growth which required massive oil consumption. As production was not able to satisfy such a demand, we ended with massive price run up.
The following graph illustrates this.

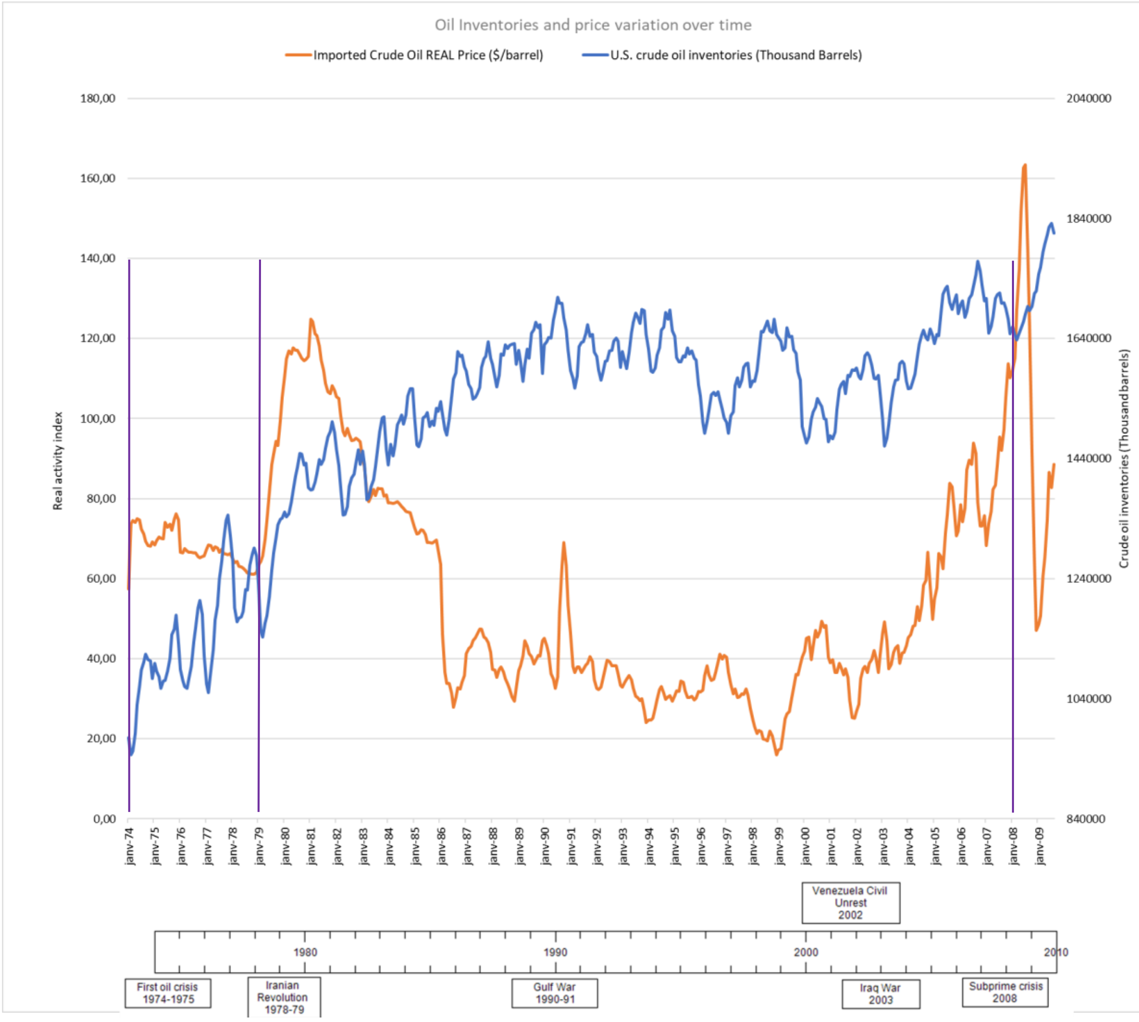


<u>Comments:</u>

This graph represents yearly oil production vs prices variations over time, from 1974 to 2009.

Except during abnormal periods, we can see that when production increases, prices decrease. Production increases steadily over time considering oil demand and investments into production facilities.

The abnormal periods are the 3 oil crisis plus the gulf war in the 90's where we can observe a punctual production drop and an immediate price peak.



<u>Comments:</u>

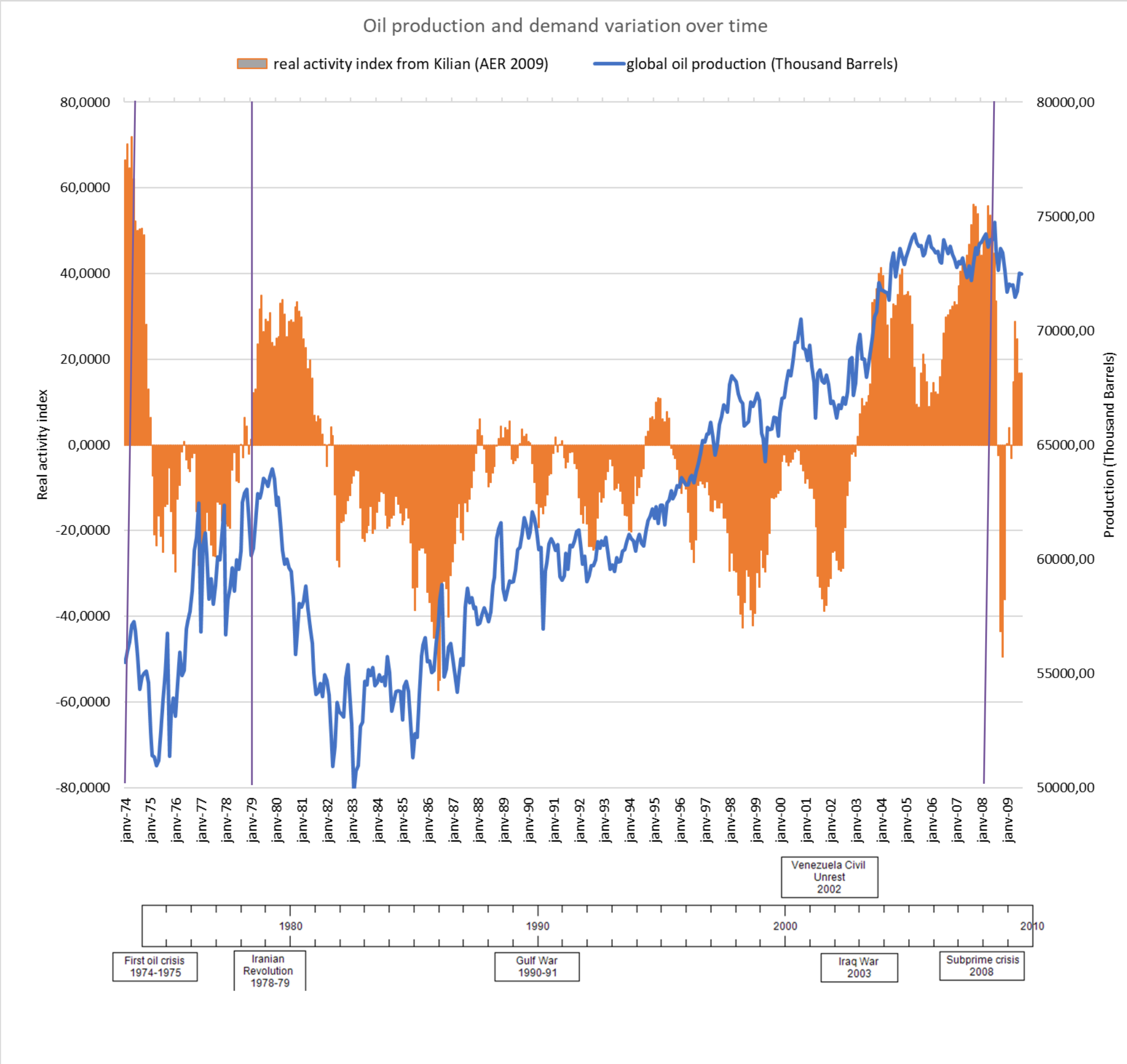
This graph represents yearly inventories vs oil prices variation over time, from 1974 to 2009.

We can notice that oil prices are down when inventories are important.

And on the other way around, when prices are increasing, we can observe that shortly after, stocks level are decreasing, and prices also decrease after.

This can be explained by the countries' willingness of to control the increasing prices, created generally by a surging demand, using the oil in the inventories to ease the pressure on production.

Using these inventories, countries can bring oil into the market until production can support the demand or that demand is reduced.

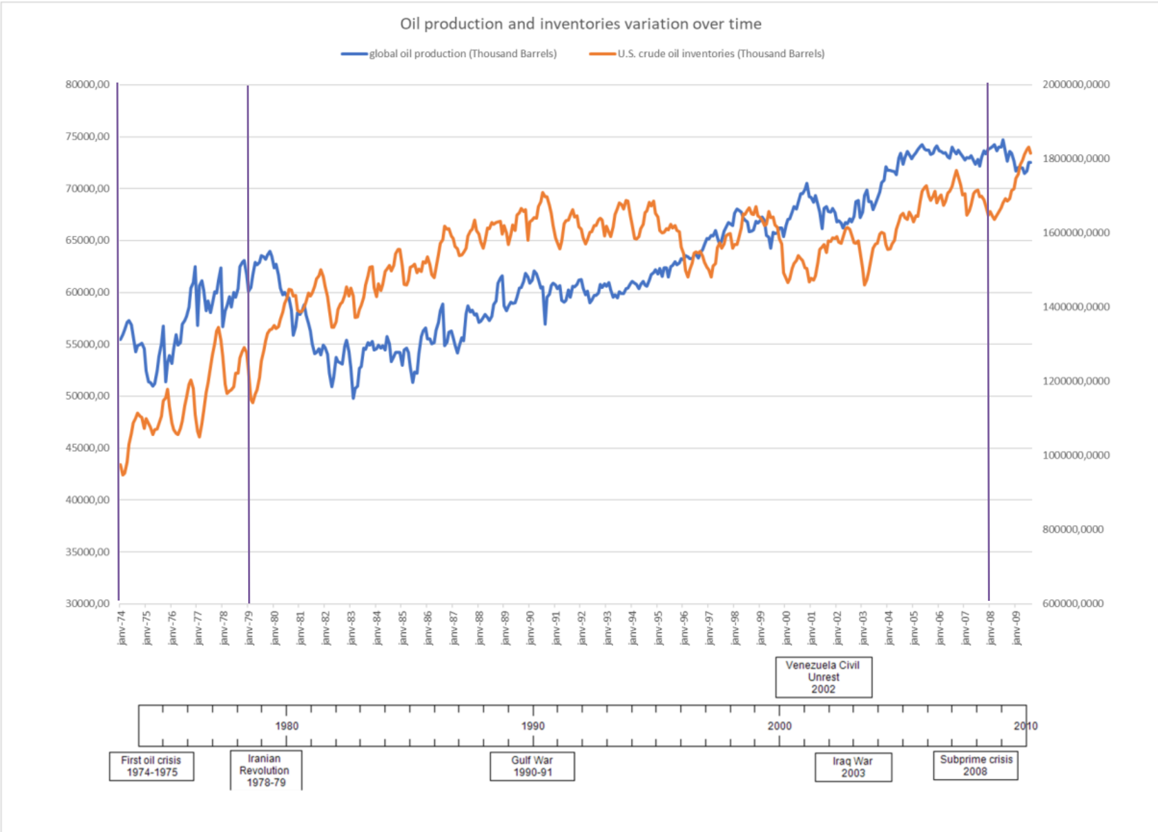


<u>Comments:</u>

This graph represents yearly oil production vs demand variations over time, from 1974 to 2009.

We observe that, globally, oil production increases, except during the periods between 1974 – 1975 and 1980 – 1983.

The first period because of a decrease of demand due to 1973's oil crisis. The second period also because of a demand decrease due to combination of Iranian revolution and Iranian/iraqi war. That led to global geopolitical and economic instability.



<u>Comments:</u>

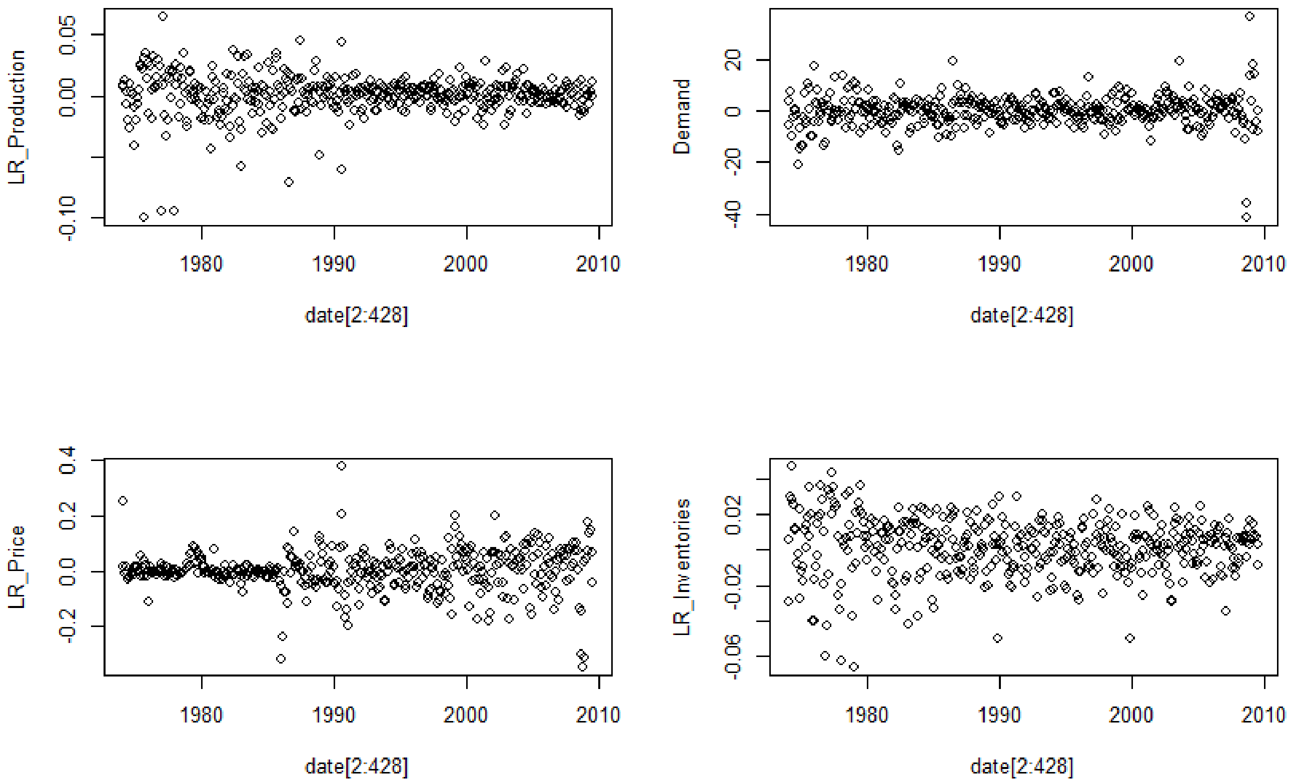
This graph represents yearly oil production vs inventories variation over time, from 1974 to 2009.
We can notice that when oil production is high, inventories are lower and vice-versa.
After the first and second oil crisis, we can observe an important decrease in the global oil production.
However, after these crises, we see that production increases again, at a high level with an increase of almost 50% of production capacity between 1980 and 2009.
On the other hand, inventories capacity has increased at a less important speed, around 28%.
Considering that production almost follows the demand, this shows that 2009 inventories are less susceptible to answer a demand crisis than before.
These decreasing inventories capacity can be explained by the important cost that this immobilization represents and also by the cost of building and maintaining the facilities.

ii. Assess how shocks to these determinants affect the crude oil price.

Steps Walkthrough:

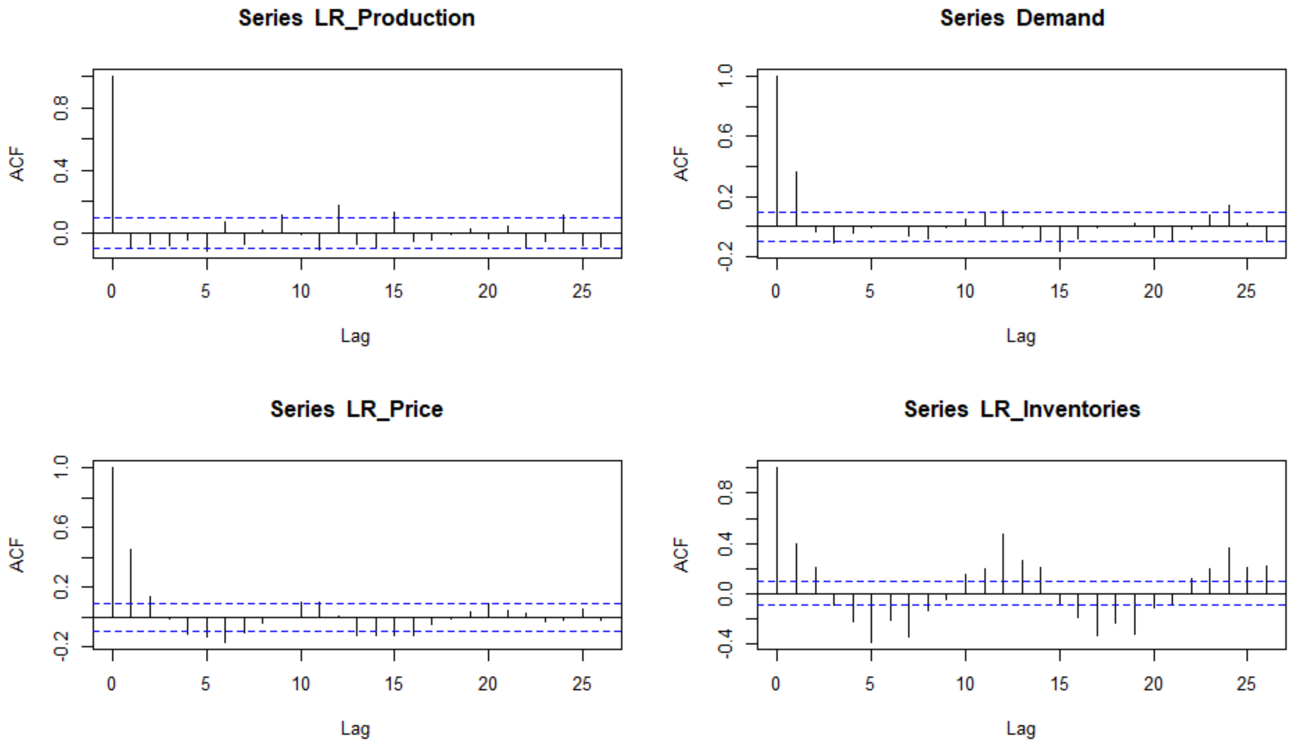
1) Plotting our diff log variables (Demand excluded)

After rendering our variables stationary through a first difference log transformation, we plotted the followings and used them in further calculations.



We visually notice a pronounced heteroskedasticity in production, inventories, and price (fan shape, although demand appears homoscedastic.
Eventually, we checked demand heteroskedasticity through an arch test to challenge our visually statement which displayed a contradictory result, meaning that in fact, demand series is heteroskedastic.

2) ACF plotting of our variables



Only inventories display a certain seasonality in its autocorrelation function that is significant at 95 %.

3) Seeking seasonality in our datasets

To determine our seasonality, we decided to use the dummy variable. To avoid multicollinearity, we committed the first season (1st month).

3.1) Inventories seasonality

```
> summary(fit.inventories)

Call:
tslm(formula = df_LInventories ~ trend + season)

Residuals:
      Min        1Q    Median        3Q       Max 
-0.054149 -0.007038 -0.000082  0.007613  0.036706 

Coefficients:
              Estimate Std. Error t value Pr(>|t|)    
(Intercept) -3.532e-03  2.445e-03  -1.445  0.14927    
trend       -7.971e-06  5.076e-06  -1.570  0.11712    
season2     -8.609e-03  3.069e-03  -2.805  0.00526 ** 
season3      1.922e-03  3.069e-03   0.626  0.53155    
season4      1.352e-02  3.069e-03   4.405 1.35e-05 ***
season5      2.226e-02  3.069e-03   7.253 2.02e-12 ***
season6      1.285e-02  3.069e-03   4.187 3.46e-05 ***
season7      1.685e-02  3.069e-03   5.492 6.97e-08 ***
season8      8.438e-03  3.069e-03   2.750  0.00623 ** 
season9      1.363e-02  3.091e-03   4.410 1.32e-05 ***
season10     6.027e-03  3.090e-03   1.950  0.05184 .  
season11     8.089e-03  3.090e-03   2.618  0.00918 ** 
season12    -1.523e-02  3.090e-03  -4.928 1.21e-06 ***
---
Signif. codes:  
0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.01293 on 414 degrees of freedom
Multiple R-squared:  0.3961,	Adjusted R-squared:  0.3786 
F-statistic: 22.63 on 12 and 414 DF,  p-value: < 2.2e-16
```

Comment:
We can observe a regular significant seasonality given a confidence interval of 99 %. We can see that it is even more true from spring to early fall and starts again the same pattern from fall enders to winter.

3.2) Demand seasonality

```
> summary(fit.demand)

Call:
tslm(formula = df_demand ~ trend + season)

Residuals:
    Min      1Q  Median      3Q     Max 
-42.321  -2.551  -0.168   2.541  35.499 

Coefficients:
             Estimate Std. Error t value Pr(>|t|)   
(Intercept) -0.236560   1.113925  -0.212  0.83193   
trend        0.002914   0.002313   1.260  0.20832   
season2      1.892071   1.398267   1.353  0.17674   
season3     -0.792269   1.398250  -0.567  0.57128   
season4      1.265508   1.398236   0.905  0.36595   
season5     -3.069398   1.398227  -2.195  0.02870 * 
season6     -3.758612   1.398221  -2.688  0.00748 **
season7     -1.047229   1.398219  -0.749  0.45430   
season8      1.957703   1.398221   1.400  0.16222   
season9      1.398621   1.408062   0.993  0.32115   
season10     0.171460   1.408048   0.122  0.90314   
season11    -2.361823   1.408039  -1.677  0.09422 . 
season12    -1.698400   1.408033  -1.206  0.22842   
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 5.89 on 414 degrees of freedom
Multiple R-squared:  0.0971,	Adjusted R-squared:  0.07093 
F-statistic:  3.71 on 12 and 414 DF,  p-value: 2.447e-05
```

Comment:

We observe no significant seasonality in demand except for the months of May and June. But combining the results with the previously mentioned ACF, we decided not to take it into consideration in our model.

3.3) Production seasonality

```
> summary(fit.production)

Call:
tslm(formula = df_production ~ trend + season)

Residuals:
      Min        1Q    Median        3Q       Max 
-0.103013 -0.006148  0.000760  0.007237  0.061769 

Coefficients:
              Estimate Std. Error t value Pr(>|t|)    
(Intercept) -1.314e-02  3.044e-03  -4.315 2.00e-05 ***
trend        1.663e-06  6.321e-06   0.263 0.792616    
season2      1.629e-02  3.822e-03   4.263 2.50e-05 ***
season3      1.346e-02  3.822e-03   3.523 0.000475 ***
season4      9.927e-03  3.821e-03   2.598 0.009719 ** 
season5      1.073e-02  3.821e-03   2.808 0.005219 ** 
season6      1.416e-02  3.821e-03   3.707 0.000239 ***
season7      1.985e-02  3.821e-03   5.194 3.24e-07 ***
season8      1.054e-02  3.821e-03   2.757 0.006093 ** 
season9      1.776e-02  3.848e-03   4.615 5.26e-06 ***
season10     1.704e-02  3.848e-03   4.428 1.22e-05 ***
season11     1.780e-02  3.848e-03   4.624 5.03e-06 ***
season12     1.329e-02  3.848e-03   3.455 0.000608 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.0161 on 414 degrees of freedom
Multiple R-squared:  0.09229,	Adjusted R-squared:  0.06598 
F-statistic: 3.508 on 12 and 414 DF,  p-value: 5.793e-05
```

Comments:

Production displays significant seasonality throughout the year given a confidence level of 95 %.

4) VAR Model

Based on literature, the granger causality test and previous seasonality results, we determined our VAR model variable order as follow :

→ Production, Inventories, Demand and Prices

Below capture represents our model results.

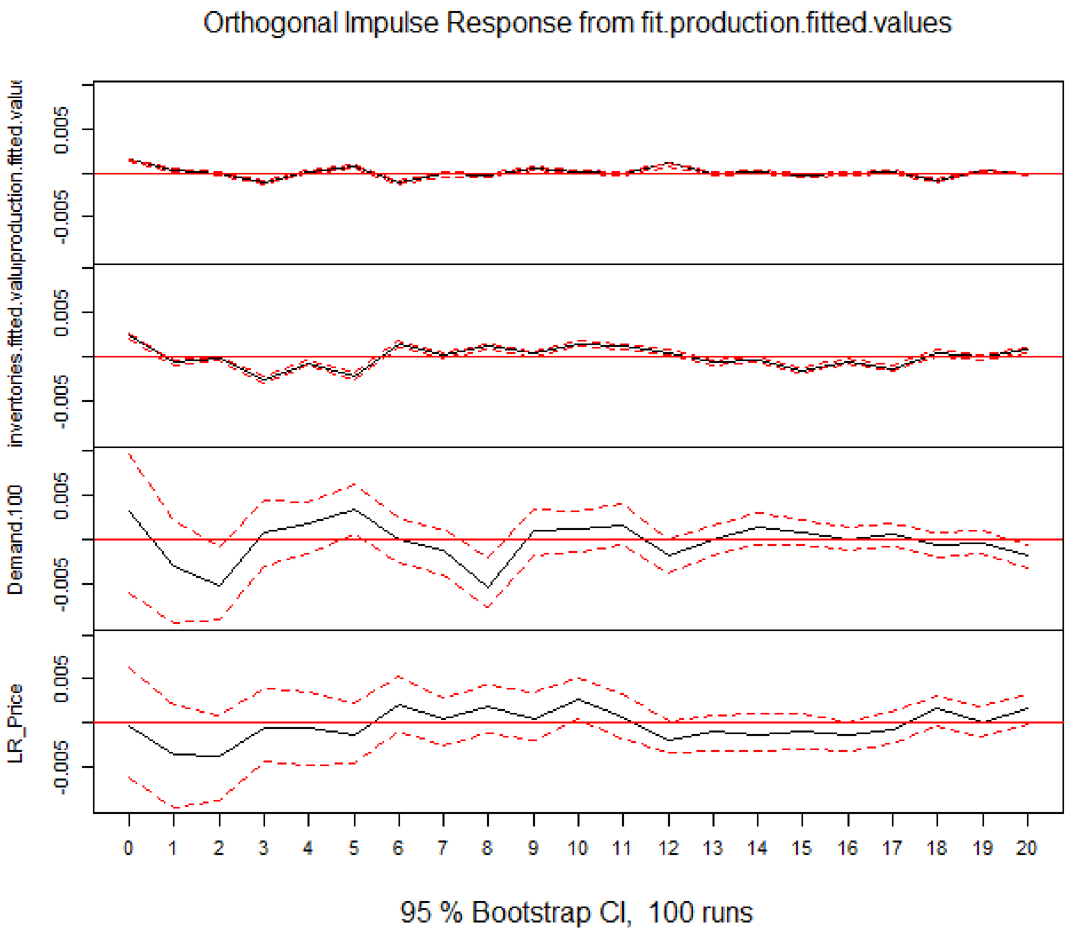
```
                                 Estimate Std. Error t value Pr(>|t|)    
fit.production.fitted.values.l1  -6.631e-01  2.418e-02 -27.427  < 2e-16 ***
fit.inventories.fitted.values.l1  5.773e-01  1.521e-02  37.964  < 2e-16 ***
Demand.100.l1                     2.856e-03  1.404e-03   2.034  0.04256 *  
LR_Price.l1                       8.254e-04  1.248e-03   0.661  0.50893    
fit.production.fitted.values.l2  -1.932e-01  2.531e-02  -7.634 1.68e-13 ***
fit.inventories.fitted.values.l2  3.513e-01  1.761e-02  19.945  < 2e-16 ***
Demand.100.l2                     1.616e-04  1.555e-03   0.104  0.91731    
LR_Price.l2                       1.730e-03  1.364e-03   1.269  0.20534    
fit.production.fitted.values.l3  -1.088e-01  2.367e-02  -4.599 5.71e-06 ***
fit.inventories.fitted.values.l3 -2.744e-01  1.436e-02 -19.107  < 2e-16 ***
Demand.100.l3                     1.449e-03  1.560e-03   0.929  0.35352    
LR_Price.l3                       5.469e-04  1.374e-03   0.398  0.69086    
fit.production.fitted.values.l4   3.721e-01  2.979e-02  12.491  < 2e-16 ***
fit.inventories.fitted.values.l4  1.125e-01  1.426e-02   7.885 3.00e-14 ***
Demand.100.l4                    -5.034e-03  1.523e-03  -3.306  0.00103 ** 
LR_Price.l4                       2.984e-04  1.362e-03   0.219  0.82667    
fit.production.fitted.values.l5   8.259e-01  2.302e-02  35.883  < 2e-16 ***
fit.inventories.fitted.values.l5  2.496e-01  1.520e-02  16.428  < 2e-16 ***
Demand.100.l5                     8.724e-04  1.460e-03   0.597  0.55056    
LR_Price.l5                       4.304e-04  1.192e-03   0.361  0.71823    
const                            -9.399e-04  8.835e-05 -10.638  < 2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1


Residual standard error: 0.001516 on 401 degrees of freedom
Multiple R-Squared: 0.9153,	Adjusted R-squared: 0.9111 
F-statistic: 216.6 on 20 and 401 DF,  p-value: < 2.2e-16
```

Our model displayed the following results, for each variable goodness of fit, meaning the model explains x % of the variable:

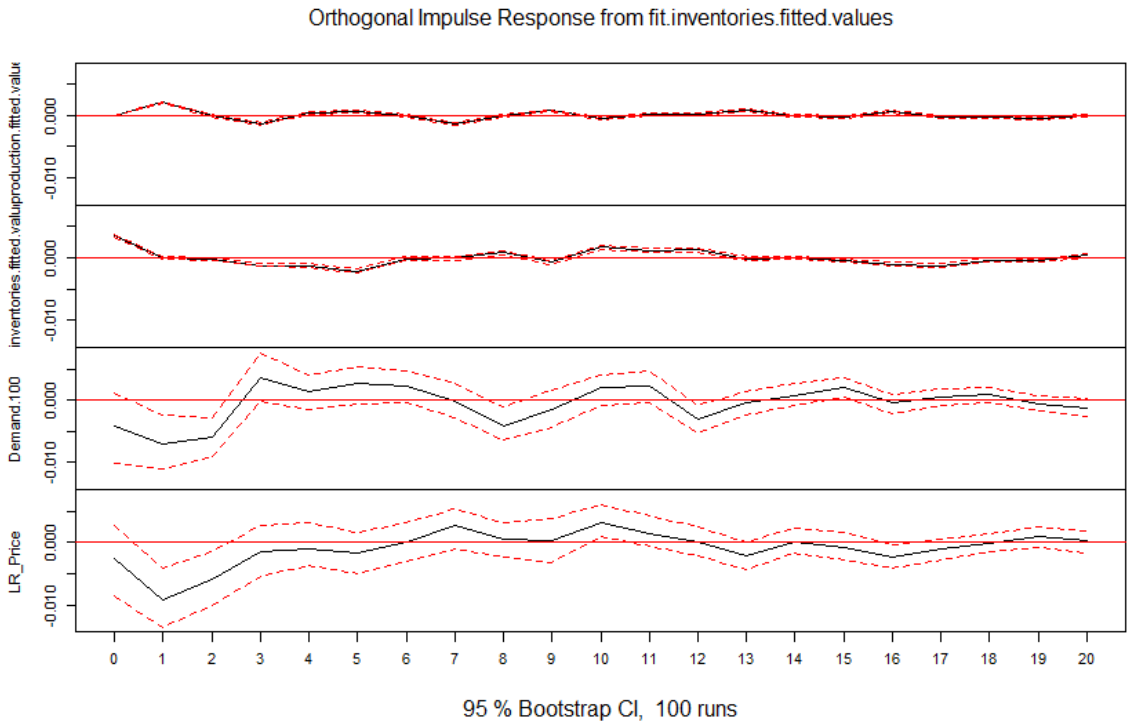
Production	0,91
Inventories	0,82
Demand	0,21
Price	0,28

5) Impulse response

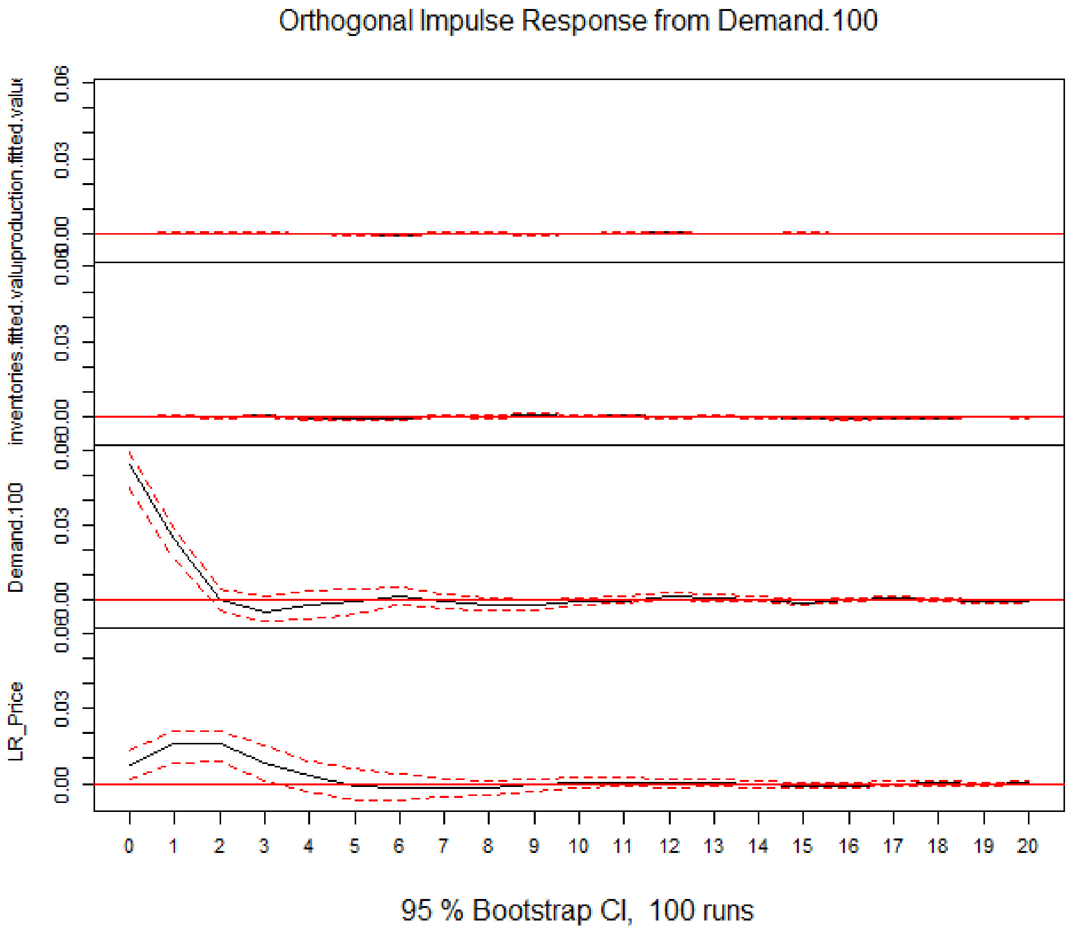
Underneath are the plots of the impulse response:



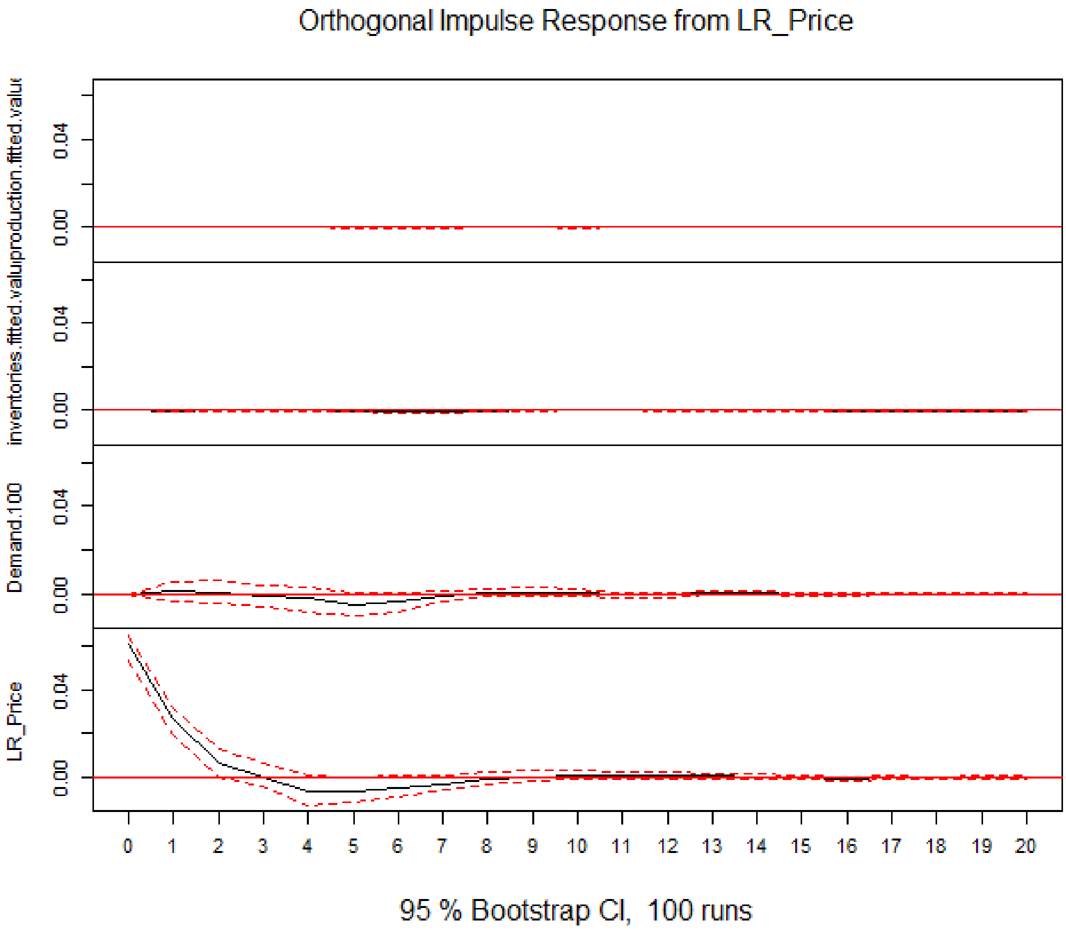
We can observe that a shock in production is spread across all variables.



Just as production, a shock in inventories affect all other variables.



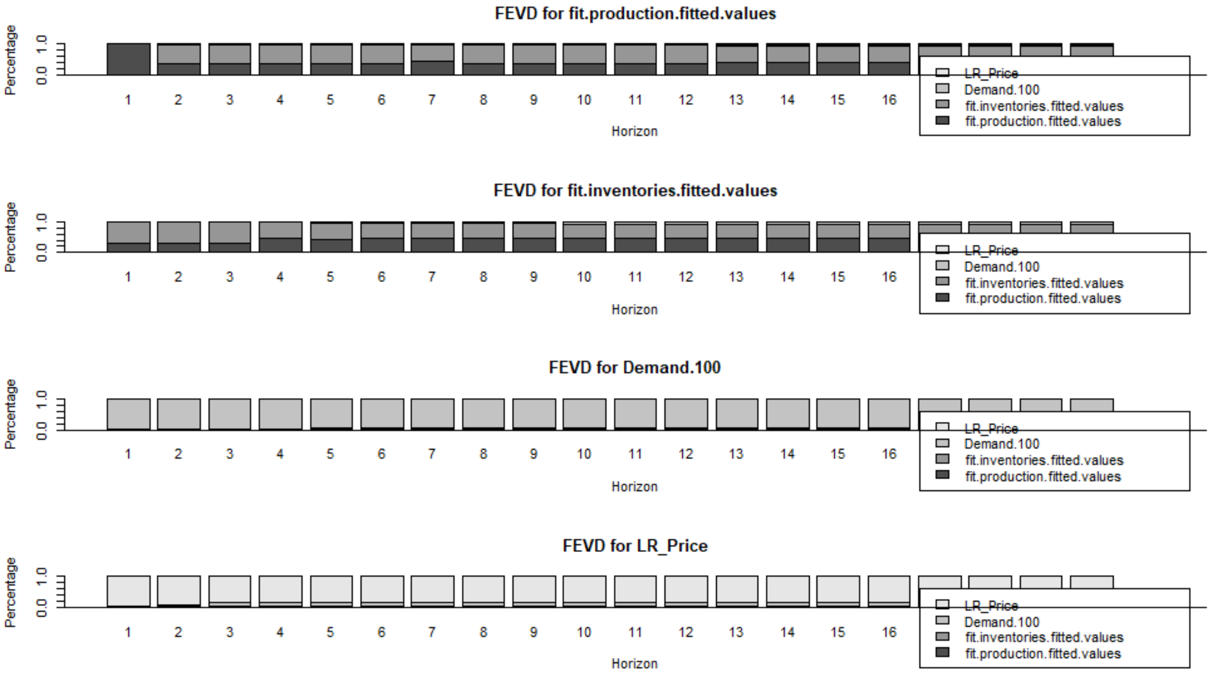
Interestingly, a shock in demand only affects the price. So far, this is a timely accordance with the construction of our model.



Eventually, a shock in price will barely impact the demand.

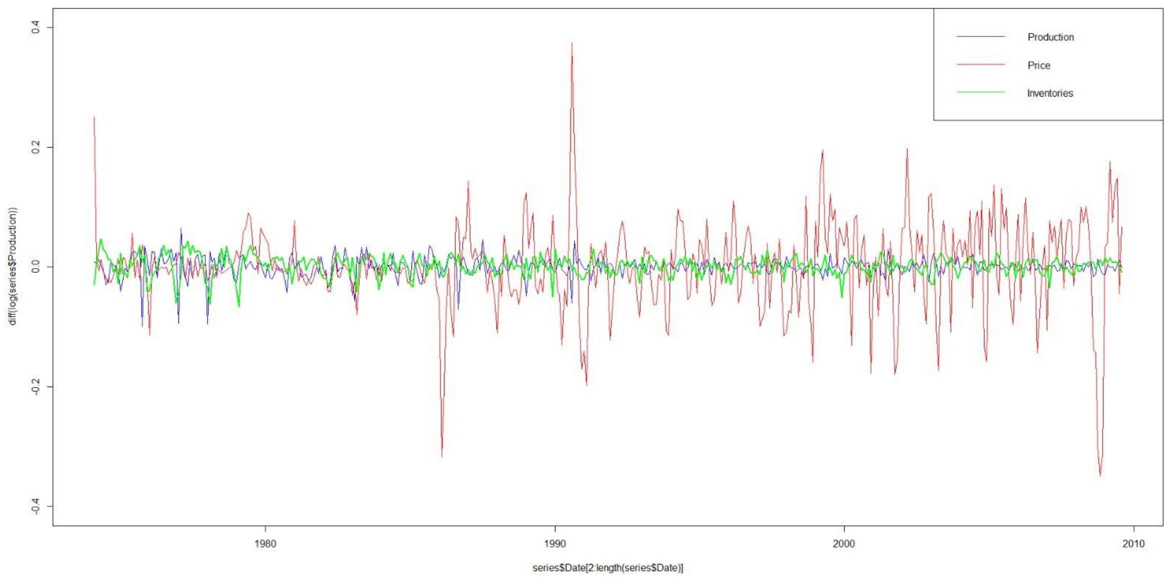
To conclude this part, production and inventories both affect all variables whereas demand and price seem to act together in pair.

6) Variance decomposition



With an Econometrics approach, it appears to be challenging to predict future prices and reconciliate with the approach used by traders which is mostly speculative and doesn't always have fundamental explanations.

Similarly, to our previous results, we can see on the below graph picturing production price and inventories, that prices began by following the overall trend and started to deviate from production and inventories right after notable crisis and worsened through what we can assume be the technology boom, allowing more people to access oil markets and speculate on it without a real interest in owning the physical underlying but benefiting price fluctuations based on their market anticipations. They usually enter cash settlement futures positions.



iii. Evaluate the extent to which the role of these determinant changes over time.

As we said before, historical determinant for oil price is supply/demand ratio.
However, other causes can affect the price.
If we study the reasons of the 3 oil crisis, we observe that the determinants are not exactly the same.
1974's crisis was essentially due to a lack of production capacity.
1979's crisis was essentially due to geopolitical issues (Iranian revolution), that resulted in supply difficulties, fears of production and/or supply issues and thus speculation.
2008's crisis was the results of combined phenomena:

- Strong economic and oil demand growth (particularly in China),
- Slow production,
- Low inventories,
- Speculation.

Here we can highlight that even if speculation has always been present on commodity markets, it became a more and more important market player in the last decades due to the increase of participants in the market. Consequence of the apogee of internet and numerical solutions for telecommunications.

Speculation can take different faces:
From speculators, that assume price direction and place orders to make a profit, without any wish to physically own the product.
From hedgers, who take futures or options to limit their risk exposure (price fluctuations) and cover their production for example.

However, it has been proven that even if speculation has its responsibility in prices movements, it is quite never the main ignition factor. It is just a catalyzer of the ongoing tendency.

Related to speculation, the market sentiment also affects and accentuate oil price direction.
As an example, a belief that oil price will increase in the future can make oil spot price sharply rise at present time.

iv. Shed light on the evolution of the real price of crude oil since 1974 towards 2008

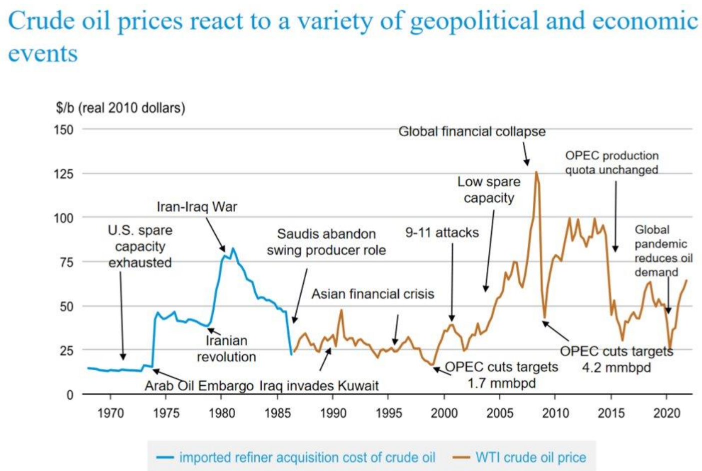
Most of the aspects of crude oil price evolution have been already addressed, so the idea in this part is to sum up all the observations drawn in this paper and explore deeply how the different factors are moving and have their own determinants.

In the previous task of this work, we have seen that there are many factors that impact the crude oil price. The supply, which is itself driven by a geopolitical context is one of the most important factors. But inventories and demand level also have a significant role in the evolution of crude oil price. The link between these different factors suggest that production was quite independent of shocks from other variables, which means there's quite a "lag" for oil producer to adapt to a demand shock for instance. But the reciprocal is not true, a shock on oil supply impact all the others factors and, of course, the price. This relation is also observed with the level of inventories.
Moreover, our analysis shows that a shock on the oil price have a slight impact on demand, which means, from an economical point of view, that oil has a low price elasticity of demand.
But at the end, supply and demand shocks can both have an important role in the variation of crude oil price. Even if this observation seems quite obvious, given that oil is not an easy substitutable asset in a globalized world, this is confirmed in some academic paper[1].
What we have also observed is that, over different period, these factors can have a different "weight" on the increase or decrease of crude oil price[2].

Let's now zoom on the different determinant for each factor:

A. Production:

In the below graph, we notice that there were some significant price spikes due to geopolitical event.



Source: U.S. Energy Information Administration, Refinitiv An LSEG Business

[1] Caldara, D., Cavallo, M., and Iacoviello, M. (2019). Oil price elasticities and oil price fluctuations. Journal of Monetary Economics, 103:1–20.
[2] Kilian, L. (2009b). Not all oil price shocks are alike: Disentangling demand and supply shocks in the crude oil market. American Economic Review, 99(3):1053–69.

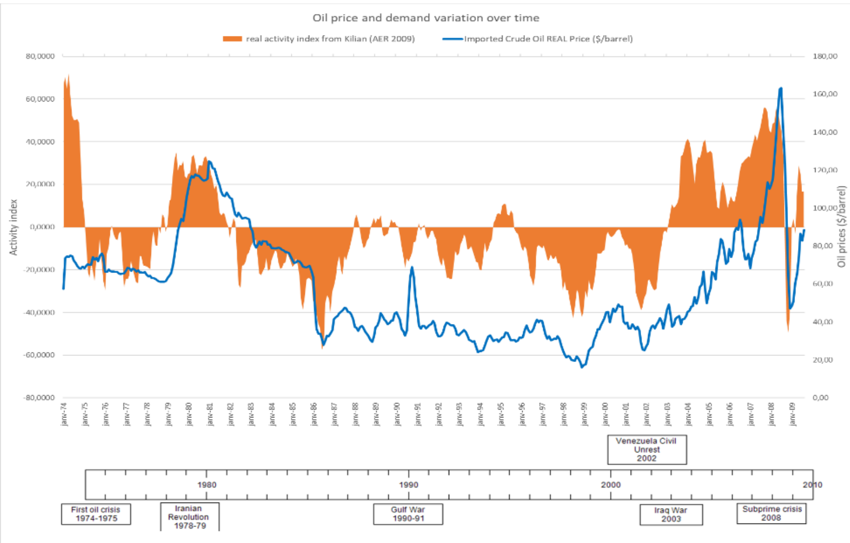
And as we know, crude oil is a very political asset due to his importance on the economic growth. The supply structure is also very specific to the crude oil. Indeed, crude oil producer are organized as a "cartel" trough the so-called OPEC[3].
The Organization of the Petroleum Exporting Countries (OPEC) created in 1960 by Saudi Arabia, Kuwait, Iraq, Iran and Venezuela which aims to control the increase in the price and quantities of oil and includes 13 member countries. It currently accounts for 39% of the world's oil supply. The main mission of OPEC members is to standardize the oil policies of member countries while considering the individual and collective interests of each. Also, OPEC negotiates with companies in the sector to exclude unnecessary fluctuations and stabilize prices. This organization has a role in the price of oil in the world.

Having said that, we must understand that it is difficult for oil producers to adapt quickly to the demand and that countries in OPEC organization -and all the others that are not in the OPEC – can let the price growing to a certain point because they have an economic interest for it, and they know that consumer will still buy oil.

We also must be careful about the qualification of supply shock. Some of them are strictly limited to crude oil while some others are more global, which means they affect all the other commodities and the market in general.

B. Demand

The crude oil demand is mostly determined by the level of economic growth in the world. As we know, when a country develops his economy, the need of oil is also increasing. Again, here the link between demand and macroeconomic situation is obvious but some academic paper show us how crude oil demand follow economic cycle.
And as a recall in the part 1, we have drawn the conclusion that, prices were highly linked to the demand:

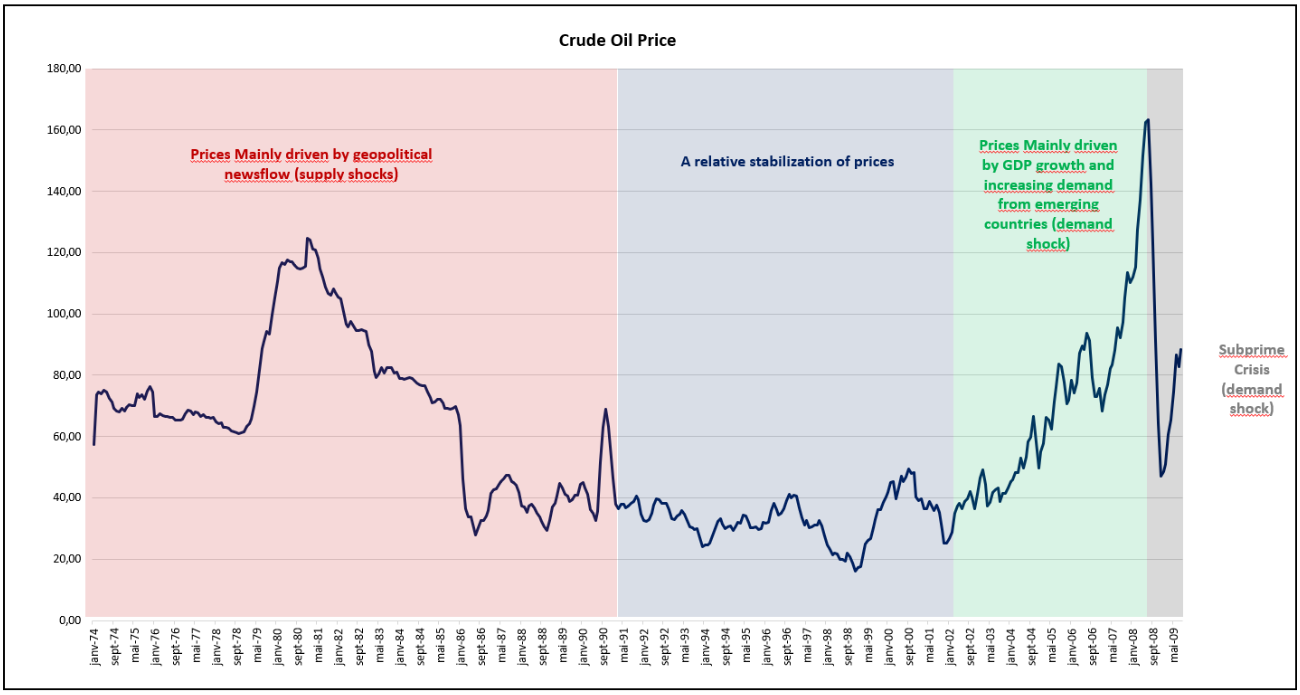


[3] Blanchard, O. J. and Gali, J. (2007). The macroeconomic effects of oil shocks: Why are the 2000s so different from the 1970s?

C. Inventories

Most of the time, a shock in inventories level is due to demand or supply shock. Inventories are used as an adjustment variable. For instance, if a negative demand shock occurs (i.e., decrease of demand), stock will be growing because of the inability of supply to adapt to the immediate drop of oil consumption. Oppositely, a positive shock in demand will have a negative impact on inventories due to inability of supply to adapt to the immediate increase of oil consumption. But generally, it is easier for a producer to increase his production than to reduce it. According to some expert, once an oil well is closed, it is very difficult to reopen it after that.
Somehow, the level of inventories shows the capacity to absorb a demand or supply shock.

D. Evolution of crude oil price between 1974 and 2008



Between 1974 and 2008, we can break down the evolution of oil crude prices between 3 main periods. The first one, going from 1974 to the beginning of the 90's are characterized by a high number of geopolitical events, which affected the supply.
The second period, going from the beginning of the 90's to the beginning of the 20's, are more "calm" period. At this time, Asian countries have a low impact on demand, so the Asian financial crisis have a lower impact than what happened a decade before. The third one, going from the beginning of the 20's (after the bursting of tech bubble) to 2008, is characterized by a growing demand in emerging countries, especially in China. Some academic paper shown that China has a very important role in the bullish rally of crude oil price after 2001[4].
During these different periods, one factor that we didn't mention yet, is the role of speculation in evolution of crude oil price. To understand the impact of speculation on prices, we need to look at

[4] Knittel, C. R. and Pindyck, R. S. (2016). The simple economics of commodity price speculation. American Economic Journal: Macroeconomics, 8(2):85–110.
Kilian, L. (2009a). Comment on 'causes and consequences of the oil shock of 2007-08'by james d. hamilton. Brookings Papers on Economic Activity, 1(2009):267–278.

academic review on this topic. In their study[5], Knittel, C. R. and Pindyck, R. S. show through their model that speculation is not responsible of the oil crude price increase during the beginning of the 20s. This increase being mainly due to the rise of demand in Asian countries. Going in this direction, Juvenal, Luciana and Petrella, Ivan[6], explain that *“oil prices have been historically driven by the strength of global demand”.*

Having said that, we cannot exclude speculation as not being an influencing factor of crude oil price. Indeed, speculation is by nature, a short-term stress factor for oil crude prices. To support this, Kilian, L. and Murphy, D. P.[7] show that speculation was an amplifying element in major period of high volatility. So, what we can learn from academic paper is that the role of speculation is often over evaluate.

---

[5] Knittel, C. R. and Pindyck, R. S. (2016). The simple economics of commodity price speculation. American Economic Journal: Macroeconomics, 8(2):85–110.

[6] Juvenal, Luciana and Petrella, Ivan. (2015) Speculation in the oil market. Journal of Applied Econometrics, 30 (4). pp. 621-649.

[7] Kilian, L. and Murphy, D. P. (2014). The role of inventories and speculative trading in the global market for crude oil. Journal of Applied econometrics, 29(3):454–478