Erasmus University Rotterdam

MSc in Maritime Economics and Logistics

2015/2016

Explaining the oil price spread between WTI and Brent during U.S. shale oil revolution

by

Dimitrios Kontaxis

copyright © Dimitrios Kontaxis

Acknowledgements

The completion of this thesis signals the end of a milestone year for me. To start with, I would like to thank my supervisor Achim Vogt. The trust that you showed on my face, pushed me harder to reach to a new understanding of a subject that I am passionate about. Your advice and support when needed, gave me courage to continue searching for the most appropriate approach that I should research the oil market, without passing from my mind that I will give up from this area of research.

Furthermore, I would like to state that I feel lucky of being a member of MEL class 2015/2016. Thank you each and every of you guys! The moments that I passed with you during the year will always be on my mind. The academic year just ended but the friendships that we created will last forever.

Finally, I would like to express how grateful I am for having such a great family. Each of you, are an indispensable part of my life and the reason for who I am right now. Thank you…

Abstract

The oil price spread between WTI and Brent experienced changing patterns during the last six years. Related to this it is often referred to the growth in U.S. oil production due to the shift to unconventional shale oil wells. In the light of these developments this study aims to explain the relationship between the price spread of the two international crude oil benchmarks and empirically test, which market forces are responsible for the spread fluctuation.

Previous research has already focused on structural change in the oil price spread and found that on the 15 December 2010 the oil price spread moved from a stationary to a non stationary process. Coming from a relatively stable premium against Brent the WTI oil benchmark started trading at heavy discounts. This paper contributes by identifying a new structural change in the oil price spread on 13 March 2014 when WTI and Brent formed a new stationary relationship. From that point onwards WTI has been traded at a slight discount compared to Brent. In order to arrive at this conclusion, we used Chow F-tests for structural break and unit root tests for (non)stationarity.

Further, we build an Autoregressive Distributed Lag model to detect which factors drive the oil spread movement. We find that the spread has a positive long term relationship with the pipeline oil flows from the U.S. Midwest to the U.S Gulf Coast, with that implying that transportation infrastructure plays a key role for the WTI price. Furthermore, we surmise that the oil price spread has an inverse relationship with the U.S. oil production and the Canadian oil imports to the U.S. Midwest.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

WTI	West Texas Intermediate
U.S.	United States of America
API	American Petroleum Institute
USSR	Union of Soviet Socialist Republics
OPEC	Organization of the Petroleum Exporting Countries
y-o-y	Year over Year
bpd	Barrels per Day
PADD	Petroleum Administration for Defense Districts
UK	United Kingdom
BFOE	Brent, Forties, Oseberg and Ekofisk
LLS	Light Louisiana Sweet
MENA	Middle East North Africa
SSR	Sum of Squared Residuals
AR	Auto Regressive
AIC	Akaike's Information Criterion
ADF	Augmented Dickey Fuller
PP	Phillips & Perron
OLS	Ordinary Least Squares
GDP	Gross Domestic Product
BDI	Baltic Dry Index
ADS	Aruoba, Diebold, & Scotti
EIA	Energy Information Administration
ARDL	Auto Regressive Distributed Lag
NYMEX	New York Mercantile Exchange
ICE	Intercontinental Exchange

1. Introduction

Crude oil is the number one most traded commodity in the world and has been through the passage of decades, indispensable part of our everyday life. Its applications dictate the way global economy grows since oil is the major fuel to our means of transportation, the raw material for the production of plastic goods and a source of electricity power in factories and our home. As a raw material, crude oil has no use to our life since what we use are products of oil that are offered through refineries and chemicals which convert crude oil and oil intermediate feedstocks respectively to end products. This means that the price refiners pay to have access to crude oil quantities, largely reflects the price that consumers pay in order to use the end products. As a result, pricing of crude oil directly or indirectly affects many sectors of the global economy.

West Texas Intermediate (WTI) and Brent are the oil benchmarks that influence the most the international oil markets. Their prices are used as an indicator and define decisions and strategies in the refining business, financial trading and government policies (Chen, Huang, & Yi, 2015). Their similar physical characteristics are depicted to the development of their price, with WTI trading at a slight premium until the end of 2010 in order to capture small differences in quality. What is more, Brent, which according to Fattouh (2011) is the basis for pricing the majority of international crude oil trade, was priced at a small discount against WTI, the benchmark to price oil imports to U.S., because Brent price had to justify freight costs and insurance that were needed in order to carry Brent crude oil to the U.S. market due to the fact that the country is a net oil importer.

During the last six years, the oil market has experienced fluctuations in the spread between the two oil benchmarks that were not prevalent before. More in detail, WTI price started decreasing in level without Brent price corresponding to the same downward pressure. The price differential reached record negative levels with WTI trading – 29 dollars per barrel compared to Brent in late September 2011. As of now we observe a narrower gap between the two oil benchmarks but WTI is still traded at a discount compared to WTI.

Taking into consideration that oil exporters used WTI as a reference to price their oil towards U.S., these fluctuations had a negative impact to their price strategy. Fattouh (2011) mentions that WTI could no longer serve as an international benchmark as U.S. regional factors became dominant into setting WTI price. As such our thesis will focus on searching for the reasons that caused these unprecedented price movements in the oil market and affected the pricing strategy of oil exporting countries. With U.S. oil production increasing by one million barrels per day in both 2011 and 2012, unconventional light oil came as a supply shock to the oil market. Factors like these set strong candidacy on changing global oil dynamics.

1.1 Research Objectives

The purpose of this research is to explore the properties of the relationship between the two oil benchmarks by analyzing the structure and underlying drivers of WTI-Brent spread movement. A stable pattern in the spread signals a link between WTI and

Brent but great fluctuations as the one that the oil market observed the last years is a challenge for the oil industry to monitor a specific pattern, a fact that has economic implications. Thus, we will focus on finding the patterns of the oil price spread by searching for break points when the structure in the relationship between the two oil benchmarks may have changed. Further, our next step is to identify factors that caused the fluctuation and test if these factors are able to explain the spread movement. This will give an extra reason to monitor the movement in relevant explanatory variables in order to interpret how the oil price spread will behave.

Having stated the above, this paper aims to answer the following main research question:

"What is the relationship of the oil price spread from 2010 until 2016 and which factors can explain this relationship?"

The rationale behind this research question lies to the fact that the last six years we have observed strange patterns in the oil price spread. Analyzing the characteristics and underlying reasons of these changes gives us a better understanding of the oil market as well as enables us to identify events that are responsible for the spread fluctuation.

In order to give an in depth answer to our main research question the following sub-research questions must also be answered:

1. "What is the relationship of the oil price spread before 2010 and why?"

2. "Has the relationship between WTI and BRENT changed after 2010 and if yes, how many times during the period under research?"

3. "Which events attributed to the divergence of the oil price spread after 2010 and to the closing of the gap during 2014 till now?"

4. "Which fundamental variables can explain quantitatively the fluctuation of the oil price spread?"

1.2 Research Design and Methodology

This thesis will use time series analysis in order to answer the questions outlined above. To begin with, literature review will be used in order to give an answer to our first sub-research question. Further, Chow F-tests for examining structural changes in the spread relationship will be used while Unit Root tests will identify the pattern of the price relationship based on estimated breaks that Chow technique will provide us. The above techniques will help us in answering our second sub-research question. By having examined possible new formations in the oil price spread relationship, an Auto Regressive Distributed Lag model will be employed in order to test the explanatory power of fundamental variables in the movement of the spread and answer our third and fourth sub-research question.

1.3 Thesis Structure

The thesis is structured as follows: First, Chapter 2 gives an overview of the commodity that we examine, its different physical characteristics, the evolution of the oil pricing system as well as what is unique in U.S. shale oil. Chapter 3 focuses on the papers that have already examined the relationship between the two oil benchmarks while Chapter 4 analyzes the events that took place in the period under research and have affected independently WTI and Brent. On the basis of the literature review in the previous chapters, Chapter 5 serves to explain the methodology and the steps that we will follow in order to research questions 2 and 4, while Chapter 6 presents the analysis of our results. Chapter 7 tries to give an interpretation of these results and the possible implications for the refining industry. Last, Chapter 8 summarizes the main findings of our research and provides advice for further research.

2. Fundamentals of crude oil

Before proceeding to the details regarding the relation of the crude oil price spread, an analysis of the fundamentals of the crude oil market is required in order to shed light on the definition of the commodity that we research and its different characteristics. What is more, we would like to describe the phenomenon of shale oil as these developments are important to an understanding of oil price developments since 2010.

2.1 Physical characteristics of crude oil

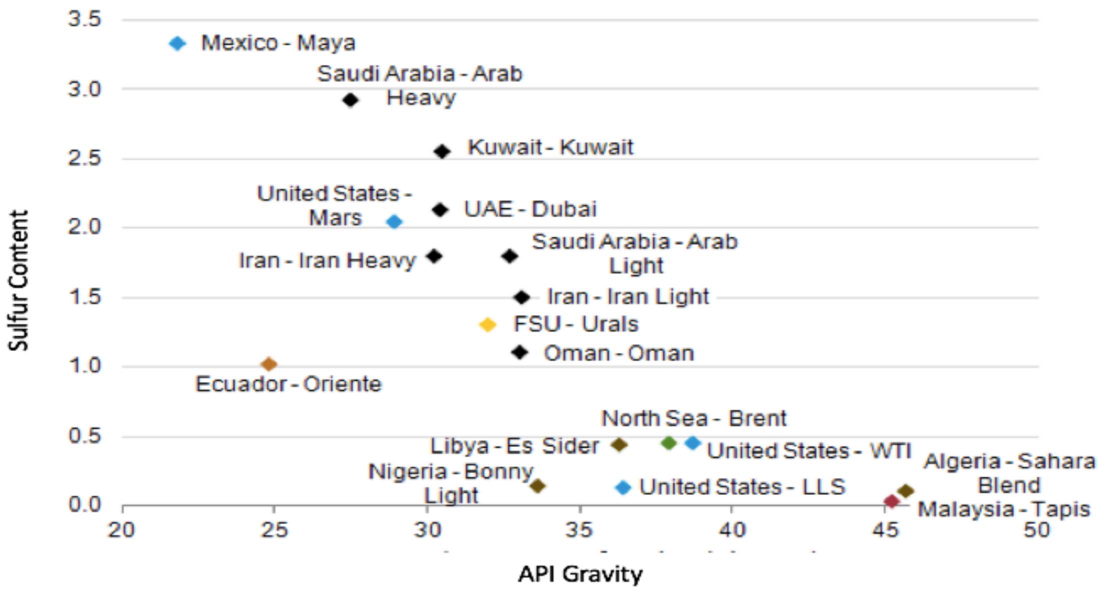
"Crude oil is a mixture of hydrocarbons that formed from plants and animals that lived millions of years ago. Crude oil is a fossil fuel, and it exists in liquid form in underground pools or reservoirs, in tiny spaces within sedimentary rocks, and near the surface in tar (or oil) sands." (EIA, 2016) Based on Deutsche Bank (2013), the remaining part of this section explains the different characteristics of crude oil since it is not a homogeneous product. More specifically, crude oil consists of different organic compounds and the number as well as the concentration of the compounds that each crude contains, defines the classification of the oil.

The most common way to distinguish different types of crude oil is to categorize them based on their density or API gravity and on their sulfur content. Firstly, density of crude oil is defined by the American Petroleum Institute where they classify crude oil into light, medium and heavy according to the API gravity. The higher the API gravity, the less dense the crude oil is and thus can be characterized as light oil. On the other hand, crudes with low API gravity are denser and are classified as heavy ones. Speaking in numbers, light crudes have an API gravity ranging from 35 to 40 degrees while heavy crudes range from 16 to 20 degrees at the API gravity scale. The main difference between light and heavy crudes is the viscosity level of each crude. The lower the viscosity of the crude the easier it is to extract it from the ground and to transport it to the refining location. This can be translated to lower operating costs for the refinery operator and hence to higher global demand for light crudes, which translates to a premium price that light crude is typically traded at compared to heavy grades.

Secondly, crude oils are classified as sweet or sour according to their sulfur content. The percentage of the sulfur that each crude contains, explains the concentration of impurities that have to be removed before the combustion, where crude oil after refining process, gives an end-product. Sweet crudes are those that contain less than 0.5% sulfur while sour crudes are identified when their sulfur is over 1%. As it is understood, sour oils are of lower quality crudes and need complex refining techniques which add cost to the refiner before transforming it to an oil product. This is a reason why sour grades are of lower demand and are priced at a discount compared to sweet oils.

Having already described the physical characteristics of crude oil, the most well-known crude oil streams are shown in Figure 1 accompanied by their specific physical characteristics.

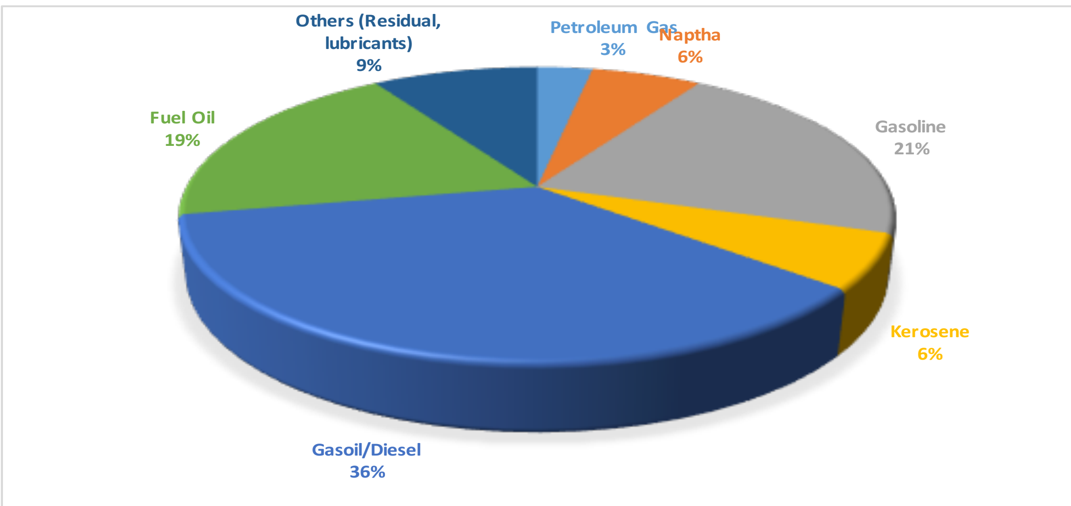
Figure 1: Grades of crude oil



Source: U.S. EIA

Furthermore, downstream producers invest to their refinery configuration based on the characteristics that nearby crude oils have, since the purpose of crude oil is to be transformed to an end-product and cover consumers' needs. Consequently, crude oil price is partly determined by the ability of crude oil to be transformed at the lowest possible cost to an end-product (IHS, 2014). Light-sweet oil is the most preferable from the majority of the refiners since it can easily be transformed to high quality products that consumers use such as gasoline, kerosene and high-quality diesel. A figure with the share of each oil product that light sweet crude oils yield is shown below.

Figure 2: Light Sweet Crude Yield



Source:(Deutsche Bank, 2013)

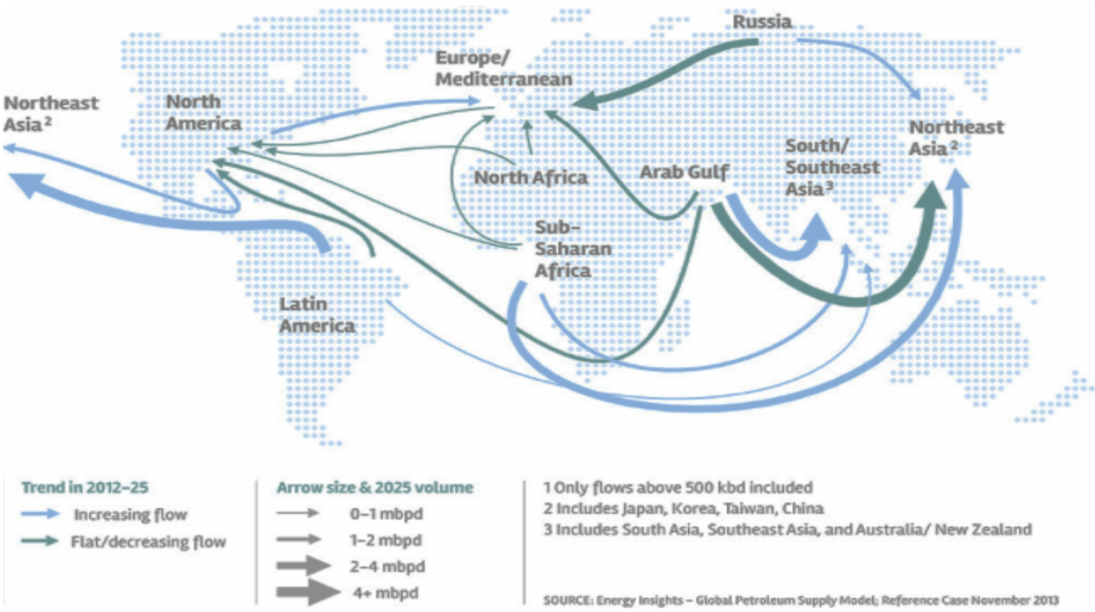
2.2 The world crude oil market

Before we describe more in detail our two oil benchmarks of interest, an overview of the crude oil trade flows will be given. Further, it is necessary to explain the development of the pricing system of crude oil through the years and how we reached to the period where Brent, WTI and other important crude streams are considered as markers for pricing oils around the world.

2.2.1 World trade

The figure below shows the trend of world crude oil trade in 2012-2025 provided by Energy Insights of McKinsey. As can be seen crude oil demand will be driven by Asia where major oil producers like the countries in the Arab Gulf will send their oil output. What is more, it is worth noticing that U.S. will start exporting oil quantiles through both the Atlantic and the Pacific Ocean while the foreign oil that U.S. will receive, will decrease. Last, we observe that Europe will not be a continent that will drive global crude oil demand.

Figure 3: World crude oil trade



Source: (Energy Insights, 2014)

2.2.2 Price formation and benchmarking

The structure of the pricing formula that the oil market uses nowadays cannot be explained without reference to previous pricing systems. Thus, the analysis of the precious pricing systems will be based on the work that Fattouh (2011) provided to the academic and research community.

The main characteristic of the oil industry until the late 1950s was the great power and market share that major oil companies possessed around the world excluding U.S., China, Canada and USSR. The interrelationship of the upstream business with the downstream one and between the multinational oil companies enabled them to control the amount of oil trade flows so as to prevent any large amounts of crude oil at the side of buyers that could drive prices down. The governments in each country that oil majors were operating, did not have any influence in the production and pricing of oil since their only relationship with the oil industry was to receive payments as landowners and income taxes. The oil market was characterized by imperfect competition which did not reflect supply and demand fundamentals but the way the oil majors were setting the oil prices in order to minimize their tax duties around the world to governments. The above technique was commonly known as ‘posted price’.

The supremacy of the oil majors and their pricing system was contested by independent oil producers who, although not operating in both upstream and downstream businesses, could extract oil not already explored from the oil majors and create a competition into the oil market. During 1950s, Venezuela, Libya, Iran and Saudi Arabia granted oil concessions to companies not belonging to the large multinational oil companies. However, the share that these companies captured from the total oil production was negligible compared to the one oil majors possessed and as a consequence the ‘posted price’ oil pricing system remained unchanged until the mid 1970s. What is more, a great role in maintaining the posted price system played the foundation of the Organization of the Petroleum Exporting Countries (OPEC) in 1960. The main purpose of this trade union during its first decade of existence was to deter any competition that could put downward pressure to posted prices and consequently lower collected income for the member countries.

A growing world oil demand from 1965 until 1973 was the reason behind the power changes that the oil market observed -- from oil majors to OPEC governments. During this period, the year over year (y-o-y) increase in oil demand surpassed 3 million bpd and OPEC was the major supplier since OPEC’s oil market share increased substantially from 44% in 1965 to 51% in 1973. The fundamentals of the market were in favor of the sellers’ side, a fact which made oil producing countries rethink the posted price and the income tax they were receiving from oil majors. Libya made the start when in 1970 signed an oil concession with an independent producer who had no access to oil reserves rather than Libya’s oil fields and agreed to pay income tax on an increasing scale. This was the trigger that drove OPEC to seek large increases in posted prices, a fact which oil companies refused and negotiations did not proceed. In response to the unsuccessful negotiations for the posted price, OPEC Arab Gulf members raised the posted price separately from \$3.65 to \$5.119 in 16 October 1973. After three days, production cuts made their appearance at the Arab Oil Producing Countries and the event which changed the dynamics of the market came in December 1973 as OPEC increased the posted price of Arabian Light oil to \$11.651. This incident signaled something unique since until then governments could only maintain the posted price but not set it.

In the meantime, the oil industry entered in a new era with OPEC governments demanding a share of the oil production and rejecting any new oil concessions, which led to a multiple pricing system at the start of 1970s. More specifically, two more ways of pricing the extracted oil introduced, the official selling price and the buyback price. Official selling price or otherwise government selling price was applied when governments had to sell their share of oil produced to third party companies while the buyback price was in force when the oil majors were taking back their share at a specific price due to governments’ inability to market the commodity to the downstream business. All this complexity led to different oil prices from system to system which could not last for long and in 1975 a new pricing system introduced that was the beginning of the formation of the crude oil market that we know today.

The new technique of pricing crude oil was derived from the complete power of OPEC to set the oil price instead of oil majors and was named as OPEC Administered Pricing System. The equity participation of governments to oil companies as well as the nationalization of few oil countries like Iraq in 1972, was the main determinant of shifting to the new pricing period were state oil and individual companies had to price their crude oil based on a marker price with Arabian Light oil being the common benchmark. The majority of the deals between parties were under long-term contracts that specified the amount of oil to be carried based on the price of the Arabian benchmark. However, in the late 1970s and especially after the 1979 Iranian crisis, the oil market observed a surge in the companies that participated in the oil trade due to the fact that Iran abandoned any long term deals with oil companies and started buying from the spot market. This had as a result the unequal increase in spot prices compared to the ones oil companies could negotiate in long term contracts that led oil producers to enter the spot market and fix agreements with buyers who could offer the highest mark-up above the oil benchmark. The implications of the emergence of the spot market were greater competition among oil companies and a shift from long term agreements to short term ones which previously were used for negligible quantities of oil carried under constrained rules.

The introduction of the formula pricing, which prices each and every crude stream around the world based on other crudes as reference, came in place through a sequence of events during 1980s when the OPEC lost its price setting power and supply and demand determined the market price for oil. One of the most significant events was the exploration of new oil basins outside OPEC countries that brought to the market increased amounts of oil. The market became more diversified since oil suppliers could price their stream at the spot market in discounted prices in order to gain greater market share from OPEC. Indeed, Saudi Arabia and other OPEC members could not cope with an increased supply in the spot market because every effort to keep their market share by increasing prices in the Saudi benchmark, was replaced by non OPEC oil suppliers at competitive prices in the spot market. The market at the late 1980s was ready to transit to the system we know nowadays with many non-OPEC suppliers and an increased number of buyers.

The pricing system evolved to the highly complex one we know today with oil markets tightly linked to one another and ways of transacting to the spot as well as to the paper market which includes futures, options and other derivative markets as trading instruments. Based on different properties of crude oil discussed in section 2.1, each crude is set at a premium or a discount to a reference price, i.e. today’s oil benchmarks. It is acceptable that light sweet oils are priced higher than heavy sour oil grades. With everything said above, the pricing formula can apply to any kind of contractual agreement which defines the absolute price of a crude stream with regards to worldwide accepted benchmarks, so that:

$$P_x = P_R \pm D \tag{1}$$

where:
P_x = the price of a crude steam
P_R = the price of an oil benchmark

D = the price differential

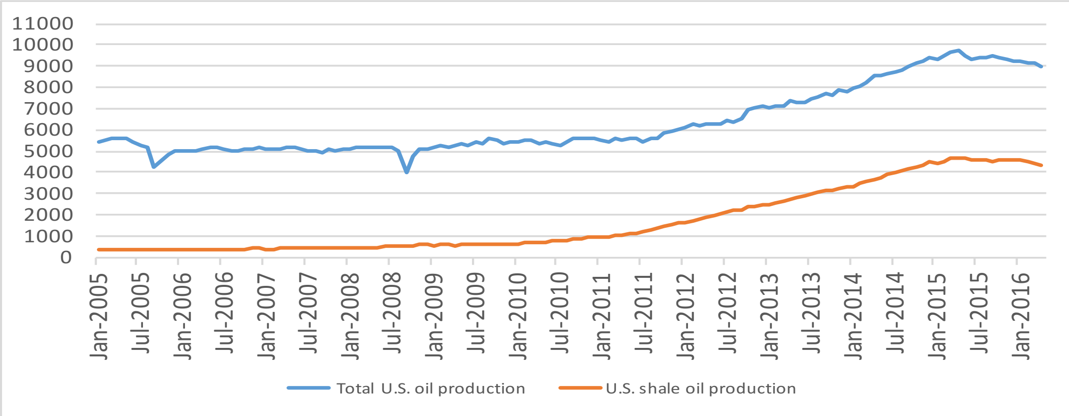
2.3 U.S. shale oil revolution

The following section serves to explain what makes extraction of shale oil unique compared to other conventional methods of extracting crude oil as well as what was the evolution that U.S. oil market experienced because of U.S. shale oil production. The in depth analysis by Kilian (2014) help us explain the below section.

“The production of shale oil (also referred to as tight (rock) oil) exploits technological advances in drilling. It involves horizontal drilling and the hydraulic fracturing (or fracking) of underground rock formations containing deposits of crude oil that are trapped within the rock. The hydraulic fracturing causes cracks and fissures in the rock formation that allow the crude oil to escape and to flow into the borehole, where it can be recovered. In some cases, advanced microseismic imaging is used to maximize the effects of hydraulic fracturing. This process is used to extract crude oil that would be impossible to release by conventional drilling methods designed for extracting oil from permeable rock formations.” (Kilian, The impact of the shale oil revolution on U.S. oil and gasoline prices, 2014, p. 1) Three important reasons were behind the surge in U.S. oil production. Firstly, high oil price levels after 2003 made this technology viable since the methods used at the start of U.S. shale story were capital intensive. Secondly, the high skills that U.S. labor possesses and thirdly the availability of suitable oil rigs compared to other countries, made shale oil production until now a feasible way of extracting crude oil for only one country (Kilian, The impact of the shale oil revolution on U.S. oil and gasoline prices, 2014).

As can be seen from Figure 4, the increase in U.S. oil production was mainly a result of the U.S. shale oil production. Production of shale oil basins increased exponentially through the first years after 2010 and then the trend of production started to flatten out. To account for the significance of the shale oil phenomenon, in March 2014 the daily demand for crude oil in U.S. was 15.5 million barrels of which U.S. economy could satisfy 8.2 million barrels. Furthermore, from the total U.S. oil production 3.6 million bpd came from U.S. shale oil basins while 7.3 million bpd had to be imported from oil exported countries. That is to say, that although shale oil managed to account for nearly half of total U.S. oil production, U.S. crude oil economy cannot depend on its own production and consequently it is considered as a net oil importer country.

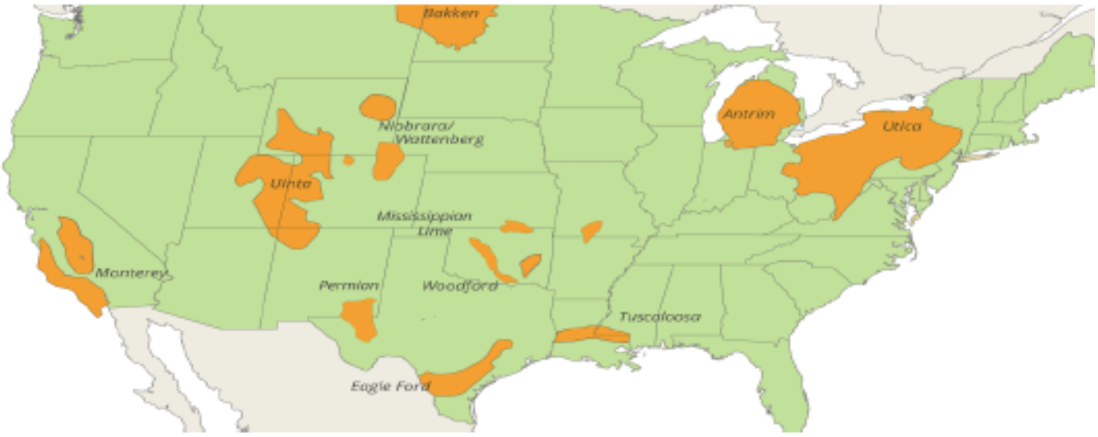
Figure 4: U.S. crude oil production vs U.S. shale oil production



Source: U.S. EIA

It is worth noticing that shale oil basins are located in specific geographic regions of U.S. and few of them possess the largest share of shale oil production. More in detail, Eagle Ford and Permnian Basin in Texas as well as Bakken shale area in North Dakota comprise more than half of the shale oil output. Figure 5 shows the location of all U.S. shale oil basins.

Figure 5: U.S. shale oil basins



Source: Energy Aspects

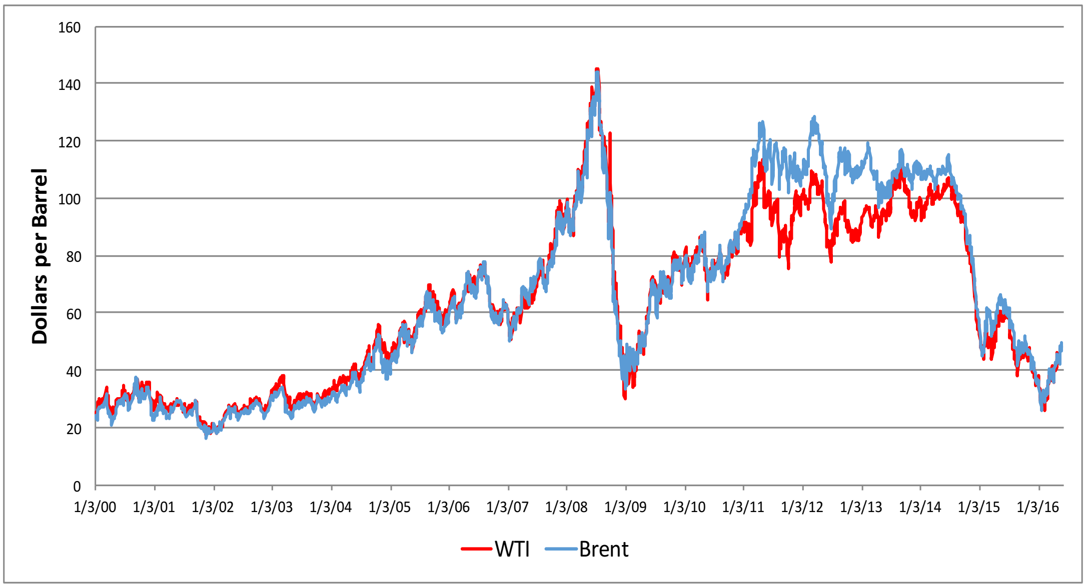
3. The WTI-Brent spread

Under this section, we present the two oil benchmarks under research as well as previous literature regarding the relationship of crude oil prices both as different time series and as a unified price spread. We will summarize studies that examine the relationship until 2010 and studies that the researchers were motivated by the unusual distortion of the oil price spread after 2010.

3.1 Brent and WTI as benchmark prices

This section serves to analyze the two oil benchmarks that our research is focused on. The majority of oil traded globally is priced on their pricing formula and the below figure depicts the evolution of their spot prices. Interestingly, both crude oil prices move in tandem until 2010 with WTI is being priced at a little premium. However, through the next four years, prices seem to diverge with WTI trading at a discount. At the last period of our sample oil benchmarks seem to move closely together again with Brent at this time being traded at a small premium compared to WTI.

Figure 6: Spot price movement between the two oil benchmarks



Source: (Thomson Reuters, 2016)

Getting into more detail, we present below the specific characteristics of each crude oil, which shows the slightly higher quality of WTI since its API gravity is higher than Brent while the percent of sulfur that WTI contains is lower. Due to the negligible differences between the two crudes, we will assume that both oil benchmarks are of similar quality.

Table 1: WTI & Brent Physical Characteristics

Crude oil	API Gravity	Sulfur Content
WTI	39.6	0.24%
Brent	38.3	0.37%

Source: (U.S. EIA, 2012b)

3.1.1 Brent oil benchmark

Brent crude oil with origins in UK is considered to be the3 most dominant oil benchmark, since 70% of international tradable crude oil is priced on its pricing formula directly or indirectly. Various reasons contribute to the fact of Brent oil as a benchmark. Located at the North Sea, Brent benchmark strategically serves two of the most important refining centers, Europe and U.S., rendering its physical location to an advantageous one compared to other basins. The means of transportation for Brent crude oil is via tanker vessels since it is a seaborne crude and its customers are either located in Europe or in U.S., if the arbitrage allows for transportation costs to the other side of the Atlantic Ocean (Fattouh, An Anatomy of the Crude Oil Pricing System, 2011). What it seems of no strategic importance for a market and a crude oil to be considered as a benchmark, is the output that can supply the markets. Although other locations contain equal or more crude oil output than Europe and more specifically the UK market, they have not emerged as an international benchmark. Other determinants for a particular oil to become a benchmark are legal, tax and operating regimes. For instance, Brent under the UK government regimes is considered to be a transparent benchmark for pricing (Fattouh, An Anatomy of the Crude Oil Pricing System, 2011). Another important determinant mentioned by Horsnell & Mabro, (1993) is the ownership diversification. In other words, the commodity under the specific benchmark should be offered by many sellers tackling monopoly effects that could manipulate the production and could deter buyers and traders to enter the market (Newbery, 1984). A striking example is constituted by the OPEC countries where each country is considered to be a unique seller and thus not satisfying the requirement of ownership diversification. Next to these countries, also Mexico could not create its own oil benchmark suffering from monopoly characteristics, a trend that is not observed at the Brent market. Historically, many oil companies were extracting crude oil entitled to the Brent market. Furthermore, the fact that more crude streams recently were included under the definition of the Brent market, enhanced even more the identity of Brent oil as a benchmark with ownership diversification.

In 2002, Platts allowed the definition of Brent to widen so that other crudes such as Forties (UK North Sea) and Oseberg (Norway) to be delivered under the Brent contracts. The new benchmark took the initials from the the three crude oils combined and formed the known BFO benchmark. In addition, Ekofisk crude oil was added under the existing crude streams in 2007 and together they created the current benchmark, also called as BFOE. It is worth mentioning that market practitioners still refer to this complex as Brent or North Sea and that the inclusion of other crude blends helped the benchmark to keep its oil output at a satisfactory level and enhance its diversification in terms of the number of sellers (Fattouh, An Anatomy of the Crude Oil Pricing System, 2011). The table below shows the quality characteristics of each crude oil that is included under the BFOE complex.

Table 2: BFOE physical characteristics

Crude Oil	API Gravity	Sulfur Content
Brent Blend	38.3	0.37%
Forties	40.3	0.56%
Oseberg	37.8	0.27%
Ekofisk	37.5	0.23%

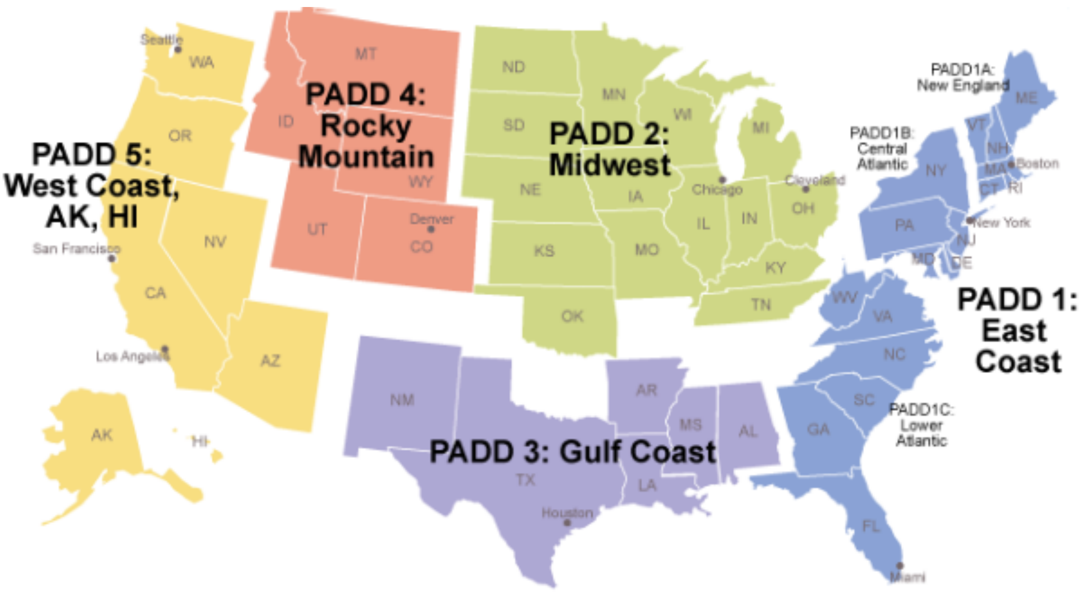
Source: (U.S. EIA, 2012b)

3.1.2 WTI oil benchmark

The following explanation of the WTI benchmark requires an understanding of the U.S. oil market and how the country is divided into different petroleum districts since WTI is not the unique U.S. crude oil but the leading one when it comes to pricing.

"The Petroleum Administration for Defense Districts (PADDs) are geographic aggregations of the 50 States and the District of Columbia into five districts: PADD 1 is the East Coast, PADD 2 the Midwest, PADD 3 the Gulf Coast, PADD 4 the Rocky Mountain Region, and PADD 5 the West Coast." (U.S. Energy Information Administration, 2012) The following figure depicts the 5 different PADDs and which cities each PADD contains.

Figure 7: U.S. PADDs



Source: (U.S. Energy Information Administration, 2012)

Based on Fattouh (2011), WTI has its oil fields located at Texas, New Mexico, Kansas and Oklahoma. WTI is a landlocked crude that has to be delivered via pipelines to the delivery hub of WTI at Cushing, Oklahoma, a state under PADD2. From this point, a pipeline network is distributing oil quantities to domestic refiners. Apart from serving domestic crude oil production, WTI was until 2011 an efficient way for pricing oil imports to U.S., the country with the largest daily demand in crude oil. The reason for that was the declining production of U.S. crude oil and the need to fill the storage facilities at Cushing and supply midcontinent refiners. This crude oil demand was satisfied with imported crude oil via pipelines from the Gulf Coast. This fact created the physical relationship between the Gulf Coast at PADD3, the location with the majority of U.S. refineries and PADD2 the physical delivery of WTI and the region of midcontinent refiners. The reference to WTI when it came to price foreign crudes, rendered WTI an international benchmark. Furthermore, WTI's significance lies to the fact that the sweet light oil futures contract, at NYMEX, is trading with a physical delivery point the WTI's physical location (Purvin & Gertz, 2010).

As stated above, there are also other significant crude streams that consist of the US crude oil market. One of them that will also help us at a later stage of our research is Light Louisiana Sweet (LLS), a sweet crude similar to the characteristics of WTI that is used as a benchmark in the US Gulf Coast, at PADD3 (Fattouh, An Anatomy of the Crude Oil Pricing System, 2011). LLS differs from WTI at the extraction process since it is a seaborne crude and does not suffer from pipeline problems and from on shore storage facilities that WTI might face during the years. Based on U.S. EIA (2012b), LLS has an API gravity of 35.6 degrees and sulfur content of 0.37%.

3.2 WTI-Brent relationship until 2010

One of the biggest arguments that researchers are trying to asses throughout the years is whether crude oil prices share a stable spread and move closely together, or whether regional factors affect crude oil prices and the movement of each benchmark or crude oil stream is independent. One of the explanatory approaches is the globalization-regionalization hypothesis firstly introduced by Adelman (1984), who thought about crude oil market as an 'one great pool'. Since then, many have tried to examine this price relationship using different techniques and finding sometimes contradicting results, too.

The first formal effort to explore and find an answer to this question, was made by Weiner (1991). He used simple correlation techniques and regression for price adjustment across the regions and found that the crude oil market is not unified as Adelman (1984) stated. A number of oil submarkets were functionally operational due to policies and sellers' power. However, Weiner just made the start to this hot topic and new studies came to support Adelman's statement of a unified world oil market. The hypothesis of a globalized oil market was supported by Gülen (1997,1999) who used co-integration methods to show that crude oil prices from different markets follow similar patterns. An arbitrage cost approach was used as an empirical method by Kleit (2001), who made estimations for transaction costs between oil regions. Kleit concluded that light oil markets are more unified during 1990s due to the continuous reduction of transaction costs. However, he mentioned that high transaction costs between oil markets were also existent. An extensive research to examine whether there is a long-run relationship between pairs of benchmarks was conducted by Hammoudeh et al. (2008). By using threshold co-integration techniques, they found that WTI and Brent prices have a long run and stable relationship. It is of great importance that when different pairs among benchmarks deviate from their long run relationship, they return to their equilibrium asymmetrically. That is to say, the time needed so that pairs of oil benchmarks return to their long run spread differs. The authors' reason to this realization stems from the fact that the arbitrage opportunities created vary according to the transaction costs of each pair of oil benchmarks as well as when the benchmarks examined, are highly liquid and are traded in futures markets. Lastly, a different approach to the oil prices relationship was applied by Reboredo (2011). Reboredo examined the dependence structure between WTI and Brent prices with copulas and found that oil prices are linked regardless the market situation, giving extra evidence for the hypothesis that the crude oil market is 'one great pool'.

The research for this argument was extended to the spread between two benchmarks, where different techniques could be applied and capturing the movement of the oil price spread as a determinant of the relationship between oil prices. The first study over the WTI-Brent price differential was conducted by Milonas & Henker (2001), who came up with the conclusion that the underlying markets are not fully integrated. The second and more recent study to the price differential of crude oils was a result of a two-regime threshold autoregressive process applied by Fattouh (2010). He concluded that the spreads between crude oil prices show a stationary pattern but the adjustment process when prices deviate from their usual spread differ according to crude oil qualities. Furthermore, he analyzed the WTI-Brent spread, which consists of quite similar crude oils as explained in Section 3.1, and attributed any deviations to temporary breakdowns of the WTI benchmark which is dependent on infrastructure logistics. On the contrary, as already stated, Brent is a seaborne crude which is not affected from pipeline bottlenecks. Although irregular movements from the stationary process of the price spread may arise, Fattouh (2010) mentions that they do not last for a long period and the markets return to their 'equilibrium', a relationship which is also given by a specific equation. More in detail, Fattouh was inspired by the cost of carry relationship, a linear equation that Alizadeh & Nomikos (2004) used to explain arbitrage opportunities when price spread between WTI and Brent was larger than carrying costs and quality discounts. The link between the two benchmarks could be given by the below equation:

$$P_{WTI,t} = P_{BR,t} + C_{BR} + D \quad (2)$$

where P_{WTI} and P_{BR} represent the prices of WTI and Brent, C_{BR} captures the carrying costs to transport the physical commodity Brent across the Atlantic to U.S. and include freight costs, insurance as well as pipeline tariffs to carry the cargo to Cushing, Oklahoma or a refiner at the Midwest. Finally, D captures the discount of Brent against WTI due to the slightly higher quality of WTI that yields to a larger share of valuable oil products and this discount is usually 0.30 cents (Alizadeh & Nomikos, 2004). If the price spread of WTI-Brent is greater than zero, the arbitrage window will open and U.S. refineries will import Brent crude oil instead of processing WTI until the 'window' closes again.

However, the distortion of the WTI-Brent price relationship that we observed after 2010, seems to have lasted for a long period and the above link between the crude oil prices cannot explain the heavy discount of WTI against Brent.

3.3 WTI-Brent relationship after 2010

Under this sub-section, we will refer to the literature that motivates us to pay special attention to the WTI-Brent relationship after 2010. Very few studies have examined the unusual distortion of the oil price spread after 2010.

An empirical analysis in identifying structural breaks in the long term relationship between WTI and Brent is given by Büyükşahin et al. (2012). By assuming break dates close to events that probably have affected the relationship of the spread, they build their hypothesis for structural breaks of the WTI-Brent spread and apply Chow F-tests to check whether their assumption can be statistically validated. They find structural breaks in 2008 and at the end of 2010 (15 December) but they did not research whether these breaks were temporary or whether they permanently changed the WTI-Brent spread relationship. What is more, they decomposed the oil price spread into benchmarks that are directly affected by the general spread in order to capture candidate explanatory variables for the divergence of the spread.

A different approach to test for a structural break date as well as for the end of the WTI-Brent relationship is offered by Heier & Skoglund (2014). They were inspired by the unusual premium of Brent against WTI and by using an Engle-Granger two step test for co-integration combined with a recursive analysis, they conclude that the long term relationship between WTI and Brent ended in 2010 and that a new co-integrated relationship was formed after 2014.

Finally, a technique which estimates the break date for a time series is performed by Chen et al. (2015). By using the CUSUM of squared based test and afterwards conventional unit root tests to check the stationarity of the spread, they find a persistence in change of the WTI-Brent spread at the end of 2010. The stationary process of the spread became non stationary after the break point at the end of 2010 and more specifically on 15th of December 2010. Moreover, they suggest researchers to further study the WTI-Brent spread since more recent events brought the two benchmarks at a closer price differential, a fact that might imply a new persistence in change from a non-stationary to a new stationary relationship.

The above mentioned studies used different techniques but they came up to the same conclusion that the relationship of WTI-Brent spread has already changed once and that the turning point was inside on the 15th of December 2010. Therefore, our research aims at finding a new structural change after 2010 that may have brought the WTI-Brent spread to a new stationary process.

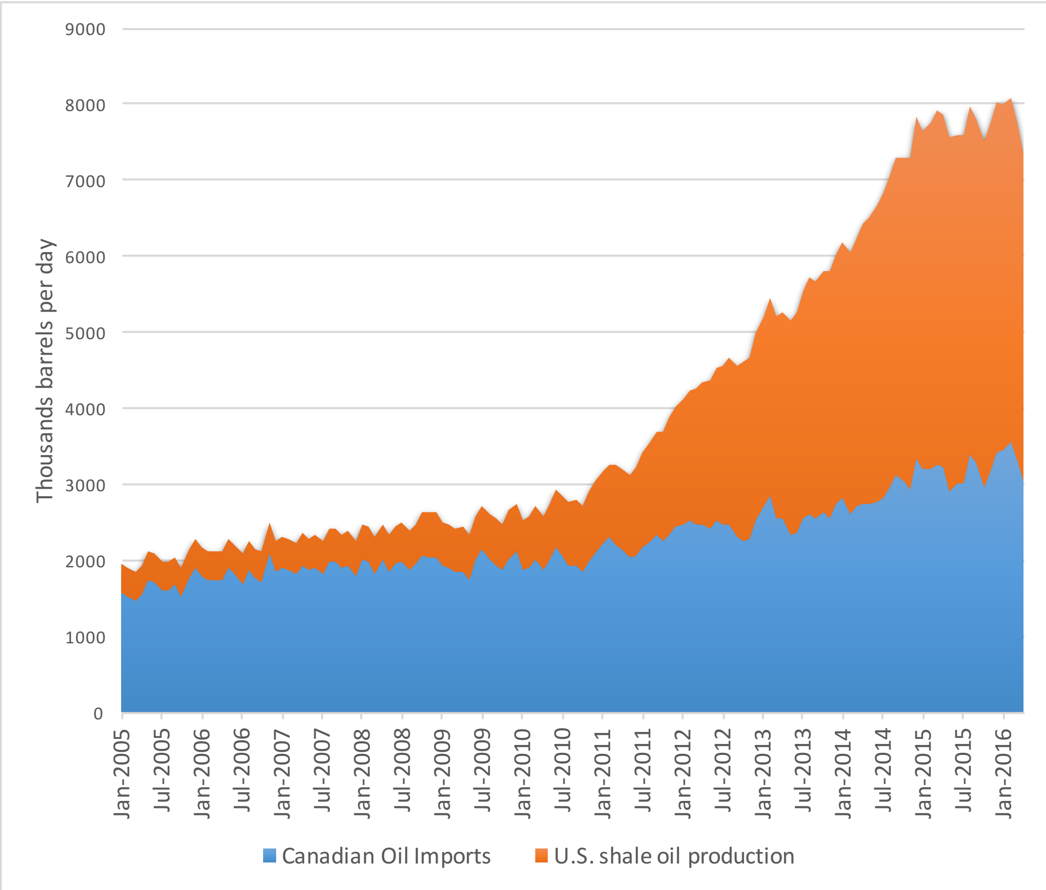
4. Theoretical effects on the WTI-Brent spread after 2010

As we have already stated, we would also like to search for the physical factors that indeed changed the relationship between WTI and Brent after 2010 as well as factors that brought the spread to a closer differential and might have changed once more the pattern of the spread after 2014. So we would like to search for events that caused the two benchmarks to move independently after 2010.

4.1 Events that affect WTI oil benchmark

Our analysis of the events that made the WTI to diverge from the international Brent benchmark after 2010, is mainly based on Kilian (2014). The results from hydraulic fracturing and horizontal drilling for extraction of unconventional light oil, were observed for the first time in the U.S. oil market in 2010 and were coincided with increasing Canadian oil imports. The figure below shows what we just mentioned with exact numbers of barrels of crude oil.

Figure 8: Increasing U.S. shale oil production and Canadian Oil imports



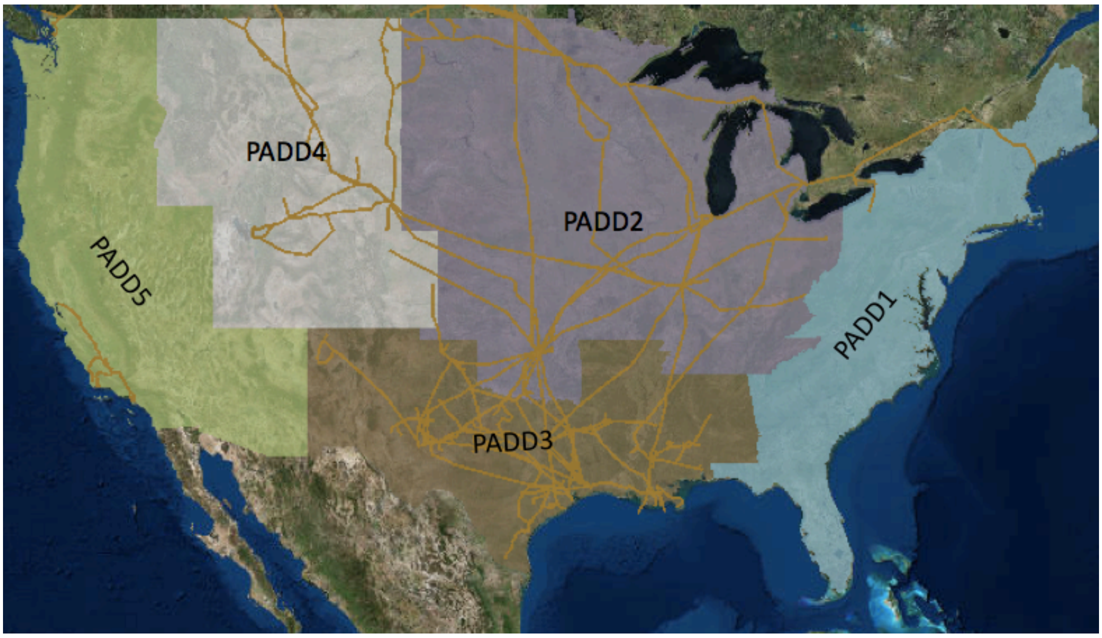
Source: (U.S. EIA, 2016)

Before getting into detail, three different constraints were responsible for the fragmentation that U.S. oil market faced during the increasing oil accumulation:

- Inability to export crude oil
- Capacity constraints in the transportation infrastructure
- Refining infrastructure

To start with, U.S. shale oil revolution came into place when U.S. refineries were making significant investments in order to change the technical configuration of their units and process heavier crude oil in the light of sweet light crude oil scarcity around the world. The majority of the refineries in the Gulf Coast, the region where at least 50% of the U.S. refineries are located, invested in this new technology in order to benefit from cheaper and heavier crude oils from Venezuela, Mexico, Saudi Arabia as well as other countries with heavy crudes. The only region that was able to benefit immediately from U.S. shale oil was PADD1 or U.S. East Coast. The reason why, lies to the fact that the refining configuration of PADD1 remained simple with no investments to complex technology that could enable refiners to process heavy sour crude oil. This particular petroleum district was traditionally importing light sweet crudes from the countries of West Africa and Europe and was supplying with oil products to the inland states of U.S.. However, there was no pipeline connection that could transfer U.S. crude oil from the Midcontinent to PADD1 and alternative means of transport like barges and rail seemed inadequate to satisfy the increasing oil production. The figure below, depicts the U.S. crude oil pipeline system among PADDs as of today.

Figure 9: U.S. PADDs and Crude Oil Pipeline System



Source: (U.S. EIA, 2016)

A similar situation proved to be the pipeline connection of PADD2, where Cushing oil hub is located, with PADD3 or otherwise, Gulf Coast. Although there is a pipeline system between the two petroleum districts, the flow of the pipelines was designed years ago to carry crude oil from the coast to the inland destinations and as a result crude oil that was accumulated at Cushing could not find a way towards the U.S. coasts. Reverting or building extra capacity of pipelines would be a solution to the unprecedented glut of oil but any operation like that would take time and would be costly.

Furthermore, we can observe at Figure 9 that Canadian crude oil was taking greater share at U.S. oil market. Canadian crude oil was not competing with U.S. shale oil since the first one is a heavy sour crude oil and its pricing formula was different from light sweet oils. However, the way to carry this specific crude, was similar to the transportation that the U.S. oil was carried among PADDs since Canadian oil was shipped via pipelines, barges or rail. PADD1 could not satisfy Canadian crude quality due to simpler configuration techniques that East Coast refineries were equipped with. As a result, Canadian oil was delivered to the U.S. oil trading hub at Cushing and from this point transportation capacity was not enough to carry the crude towards the Gulf Coast, where complex refineries that process heavy crude are located.

The increasing inflow of heavy and light crude oils at Cushing was a mismatch of local refinery demand and crude oil supply which, all else being equal, put downward pressure to the price of WTI against Brent. It is rational for anyone to think that since U.S. and Canadian crude oil is captivated at PADD2, oil exports would be a strategy that would relieve the built in oil storage at Cushing.

4.1.1 Inability to export crude oil

Under this section we present events that justify the increasing oil inventories as both U.S. and Canadian oil producers seemed to have no alternative markets to supply with their increased production. The analysis of the reasons is based on Kilian (2014).

Starting with the heavy Canadian oil, sellers found it difficult to carry their production towards ports since the pipeline system of Canada is well developed in the mainland but not at the coasts. Furthermore, the option to construct pipeline infrastructure at the East Coast of Canada made no sense since the European oil refineries that sellers would target across the Atlantic Ocean, are processing light sweet crude oils. The reason for that is that historically Europe was close to light crude basins like Brent and European refiners found no intention to upgrade their configuration technology in order to process heavier crude qualities like the Canadian oil. Exports from the Pacific side would seem more reasonable since Asian countries are able to handle such crude quality but disputes with land owners dragged the construction of any such pipeline network. Therefore, the bulk of Canadian oil production seemed to have only one direction, towards U.S. oil market. An alternative choice to export Canadian crude oil would be from U.S. ports at the Gulf Coast. However, this would create a disadvantage for Canadian oil producers since they would have to heavily discount their oil output in order to cover pipeline tariffs until U.S. ports and then shipping costs around the Cape of Good Hope, Africa. The increase in transit time in order to deliver crude oil to the Asian market would put them out of competition in comparison with Middle East oil producers. Last but not least, U.S. ports do not have adequate capacity to receive port calls from Very Large Crude Carriers, a fact that restrained any economies of scale for maritime transportation. The below figure depicts the competitive disadvantage that Canadian oil producers had in comparison with Middle East oil producers when it came to export crude oil from U.S. Gulf ports and final destination the Asian refineries.

Figure 10: Crude oil shipping routes

Source: (U.S. EIA, 2014)

With regards to U.S. crude oil, U.S. law had imposed a crude oil export ban that was dating back to 1975, where U.S. suffered from the oil crisis of 1973/74, rendering the nation highly energy dependent. However, the shale oil revolution changed the energy identity of U.S. with the increased domestic oil production. A lift of the ban would act as a relief for the domestic oil production since U.S. oil producers could find a higher price for their oil output abroad since European markets could match their oil configuration with U.S. crude oil characteristics.

4.1.2 U.S. pipeline system

Under this section, we would like to refer to the evolution of the U.S. oil pipeline system since a possible expansion of the U.S. oil pipeline network, reversion of strategically located pipes as well as construction of new ones after 2010 would have played a significant role to the transportation of the increased U.S. oil and to the domestic pricing system. The reason for that is that pipelines are the cheapest means of transport for U.S. crude oil but the most capital intensive when it comes to configuration of the existing or the construction of a new pipeline network.

In 2011 we observed a decoupling of WTI from the other important domestic oil benchmark, LLS. The reason behind this divergence, as already stated briefly above, lies to the fact that there was only one main pipeline connection between the two benchmark regions, the Seaway pipeline (McRaey, 2015). More in detail, McRaey (2015) mentions that this pipeline network was constructed and was carrying oil from the Gulf Coast import terminals as well as from Gulf oil producers to the Midwest refineries until the end of 2012. However, the oversupply at the oil hub of Cushing, Oklahoma brought down the utilization rate of the pipeline since there was negligible demand to transport crude oil from south, Gulf Coast to North, Midwest as domestic oil production could meet the demands of the region. Thus, WTI experienced a great discount against LLS and Brent so as other means of transport like rail and barges could justify their higher freight rate and carry the crude oil from PADD2 to PADD3 and PADD1 (U.S. EIA, 2015).

4.2 Events that affect the Brent oil benchmark

As it has already been stated, Brent oil benchmark is affected by global dynamics since 70% of global oil traded is priced based on Brent's pricing formula (Fattouh, An Anatomy of the Crude Oil Pricing System, 2011). That is to say, any geopolitical unrest that causes volatility to the market as well as any oil supply disruption puts, *ceteris paribus*, an upward pressure to the global benchmark. Under this section, events will be discussed that affected the price of Brent during our period under research.

To begin with, the political unrest that the Middle East North Africa (MENA) region experienced during the Arab Oil Spring in 2011, caused substantial disruptions to the oil market since Libya, the country with the biggest oil reserves in Africa, suffered from nationwide sanctions. The result was an oil supply disruption from a country that was a consistent oil exporter to Europe (Darbouche & Fattouh, 2011). More in detail, the loss of Libyan oil supply reached 1.6 million bpd, a quantity which affected the price of Brent with an upward pressure. This is due to the fact that Libyan oil production is light and sweet and OPEC spare oil capacity could not substitute this type of oil quality, since OPEC producers produce mainly heavy sour oil. European refineries, in order to fill the lost output, had to raise the demand of West African crude oils with similar quality characteristics as Libya's and priced at Brent price formula. Increased demand for light sweet oils had as an outcome a rise in the Brent price (Darbouche & Fattouh, 2011).

At the same year and close to the events that took place in the MENA region, one more event came as a shock to the oil market. An earthquake of large scale and a tsunami caused Fukushima's nuclear power station in Japan to shut down. The loss of electricity from nuclear power had to be substituted with other form of energy and the demand of fossil fuels and more specifically crude oil and natural gas, increased substantially. The y-o-y increase in crude oil demand reached 143.1% which had as a result an increase in the price of Brent (Hayashi & Hughes, 2012).

We will empirically test the quantitative significance of the events stated above in our regression model by selecting explanatory variables that can capture the changes that these events brought to the spread relationship.

5. Methodology

Taking into consideration the unusual behavior of the oil price spread and the events that took place during U.S. shale oil revolution, we would like to introduce the quantitative tools that will help us answer our sub research questions regarding the possible formation of a new relationship between WTI and Brent as well as which factors possibly have a quantitative impact on the movement of the spread.

In order to test for a new structural relationship of the price spread we need to test for structural break dates close to events that we consider to have the potential to cause such changes. For this we will be using Chow tests for structural break on daily oil prices. We will extend the research of Büyükşahin et al. (2012) for a new possible structural break after 2010 since their sample data ends in July 2012. If we find a structural break, we will test if the break was a temporary situation or indeed changed the relationship of the oil spread from a non-stationary to new stationary process. We will examine this step of our research by dividing our price spread sample on the structural break point and test with conventional unit root tests if terms like the mean and the variance are constant for the sub sample over time. Finally, we will use an autoregressive distributed lag model (ARDL) to test which variables can explain the price spread during U.S. shale oil revolution.

As such, this chapter will explain step by step the econometric tools and statistical tests needed in order to fulfill the purpose of our research. We will explain the rationale behind using Chow tests, Dickey Fuller and Phillip and Perron tests and the ARDL time series model. Lastly, a detailed description of the data used for each step of our methodology is given.

5.1 Chow test for structural change

Chow test was originally introduced by Chow (1960) and the aimed at determining whether two different samples of observations could be pooled to the same linear regression. What Chow examined at his original paper, was whether time series data from two different time periods could show the same relationship between dependent and independent variables. The rational behind this original idea could then be applied to test whether a major shock in the economy has changed the relationship among the variables, i.e. whether the regression coefficients before and after this shock are different. In economics and econometrics this incident is referred as a structural break and Chow test can only be used if someone has a priori idea of the date that the coefficients of a time series have changed between the two sub-samples defined by the break date (Verbeek, 2004).

The Chow test is an F test which starts with the assumption that coefficients between two sub samples are equal. The first step applies to the regression equations of the two sub samples and the sum of squared residuals for both subsamples is obtained as SSR unrestricted. Afterwards, a regression equation is created for the combined data and the sum of squared residuals is obtained as SSR restricted. “Chow shows that the ratio of the difference between these two sums to the latter sum, adjusted for the corresponding degrees of freedom, will be distributed as an F-ratio under the null hypothesis.” (Lee, 2008).

Following Wooldridge (2012) we construct our equation for the spread of WTI-Brent as an autoregressive process of order one [AR(1)] plus a constant and we regress it over the whole sample in order to obtain the sum of squared residuals for the restricted model:

$$Spread_t = \alpha + \beta Spread_{t-1} + u_t \tag{3}$$

Where:
α = a constant
$Spread_{t-1}$ = the value of the oil price spread the day before
u_t = an error term

We will explain the break point that we chose for our Chow test in the following section, a date that divides our period into two sub-samples. We run regression for both sub samples obtaining the sum of squared residuals for the unrestricted model so that we can calculate our F-test:

$$F = \frac{(SSRr - SSRu)/k}{SSRu/(T-2k)} \tag{4}$$

Where:

$SSRu$ = sum of squared residuals for the unrestricted model by adding up the residuals from subsample one and two
$SSRr$ = sum of squared residuals for the restricted model
k = number of restrictions by deducting the number of coefficients of the restricted model from the number of coefficients of the unrestricted model
T = size of sample
$T - 2k$ = degrees of freedom

Our hypothesis for our Chow test will be:

Ho: No difference between the coefficients of the sub samples, no break

$H1$: There is a break and coefficients differ between the sub samples

Our decision rule to reject or no the null hypothesis is that the F value should be higher than the 5 percent F critical value when our sample experiences a structural change.

In order to capture the highest F statistic, we will run our Chow test for three consecutive months around the hypothesized break period. This way, we can search for the highest Chow value over a period of possible break dates and then we will decide which will be the unique structural break date. This is a technique quite similar of what Quandt (1960) introduced and is known as Quandt statistic.

5.1.1 Break date

We base our findings to previous research and benchmark dates of completed transportation projects and changes in the trade flow of U.S. crude oil that will help us to set our structural break period. Chow test as explained above will test the credibility of our assumption.

Starting with Bakken oil producers at North Dakota, instead of sending their oil output at Cushing, Oklahoma and take as a reward a heavy discount, they took advantage of rail transport and started delivering their oil output at St. James, where oil is priced at the domestic benchmark, Light Louisiana Sweet (LLS) crude (Fielden, RBN Energy, 2013). In Figure 8, we observe that rail transport came as an alternative means of carrying the increasing production of crude oil from the Midwest to the Gulf Coast since the operations started after 2010 and reached their peak during 2013. That is to say, *ceteris paribus*, less crude oil is heading to the oil hub of Cushing, where WTI is priced and is finding its way towards the market with the refining demand, giving a relief to the oil trade hub at PADD2 and an upward pressure to WTI.

An extra solution to the transportation infrastructure that blocked the way of crude oil towards Gulf Coast market, came with the reversal of Seaway Pipeline on May 2012, a network system that is connecting Cushing, Oklahoma and the Freeport, in Texas area. Initially, the reversal of the pipeline could carry 150,000 bpd but with an expansion early in 2013, the capacity of the pipeline increased close to 400,000 bpd. As a last upgrade of the the Seaway project, a parallel pipeline which was constructed next to the original pipeline network and started operations in July 2014, increased Seaway's capacity more than twofold to 850,000 bpd (Seaway Pipeline, 2016). This reversal worked as a response to pipeline bottlenecks from Cushing to Gulf Coast and, *ceteris paribus*, enabled increased crude oil flows of both sweet light oil (U.S. oil) and heavy sour oil (Canadian oil) to Gulf Coast refineries, decreasing oil stocks at Cushing and giving an uptick to WTI price.

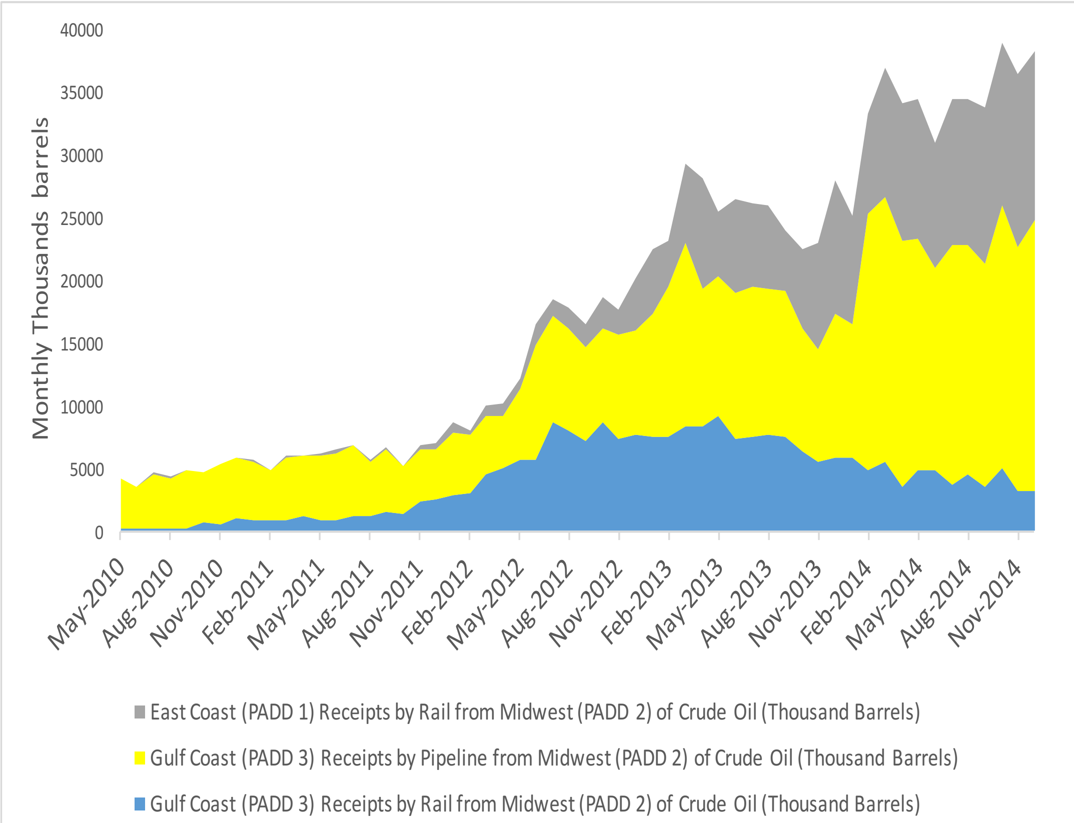
The project that impacted quite importantly the microstructure of U.S. crude oil market, is Keystone XL project, a pipeline network that connects many different oil regions. The most significant part of the pipeline is the one that connects Cushing with oil refineries in Texas and came into place in early 2014 (Kilian, The impact of the shale oil revolution on U.S. oil and gasoline prices, 2014). The completion of this part of the pipeline enabled a total capacity of 830,000 bpd to flow at the Gulf Coast refineries (Keystone XL, 2016). As with the Seaway Pipeline, Keystone XL project, *ceteris paribus*, provided a storage relief at Cushing, Oklahoma and smoothed the delivery of crude oil at the Gulf, increasing the price of the restrained till then WTI benchmark. The movement of crude oil by pipeline from PADD2 to PADD3 as shown at Figure 8, shows the sudden increased use of pipelines as a means of crude oil transport after the completion of Keystone XL, Gulf Project and the expansion of Seaway pipe.

Last but not least, at PADD1, although having limited refining capacity (1.3 million bpd), the origin of their feedstock was coming from abroad, priced at the international Brent benchmark, due to the domestic transportation problems mentioned in section 4.1. However, after 2011, where there were negligible quantities of crude by rail delivered at East Coast, we observed a surge of rail transport that was a result of the heavy discount of WTI to Brent. That is to say, East Coast refiners could secure domestic crude oil coming via rail from PADD2 and Bakken producers, replacing the more expensive Brent crude oil and covering rail freight costs (Fielden, RBN Energy, 2013). Taking a look at Figure 11, we observed that crude by rail from PADD2 to PADD1 reached its peak during 2014. We can interpret the specific event as, Bakken producers at PADD2 found new alternatives to supply their oil output and, *ceteris paribus*, less oil is heading to Cushing, Oklahoma giving a rise to WTI price.

The observations stated above, are the source of setting our structural break date.

Hypothesis: The WTI-Brent spread experienced a structural break in the first quarter of 2014.

Figure 11: Monthly movements of crude oil from PADD2 to PADD1 and PADD3



Source: (U.S. EIA, 2016)

5.2 Persistence in change?

As already stated in previous research, the historical stationary process of WTI-Brent spread stopped to exist at the end of 2010 and a non-stationary process came in place driven by the U.S. shale oil revolution and the inadequate U.S. pipeline system (Chen, Huang, & Yi, 2015). Under this section, we will explain the importance of stationarity in a time series, how to test for it and why we should apply unit root tests to our sub sample after the structural break, if found, based on the previous methodological step.

Following Wooldridge (2012), we refer to a stationary time series process when the probability distribution of our variables of interest is independent upon time. Most commonly, we focus on the mean, variance and covariance of the time series to check if they are independent of time instead of the whole distribution. Nonstationarity may have different reasons of existence but the one that we will examine is, if our spread time series has a unit root. In order to explain the unit root test, we treat our time series as an autoregressive process of order 1.

$$Spread_t = \theta Spread_{t-1} + \varepsilon_t \tag{5}$$

By setting θ=1, then our equation above is a first order autoregressive process with a unit root. By setting |θ|>=1, the unconditional variance of $Spread_t$ is infinite and the process is nonstationary (Verbeek, 2004). The difference between stationary and nonstationary time series lies to the fact that when random shocks or innovations happen, the first one is mean reverting, meaning that there is a tendency to return to its long-term mean, while the shocks in the second one have a persistent effect causing the time series to be in the so called random walk (Wooldridge, 2012).

5.2.1 ADF and PP test

The influential paper of Dickey & Fuller (1979) introduced the Dickey-Fuller test to examine whether a time series is nonstationary, that it to say if a time series has a unit root. Based on a first order autoregressive model, the Dickey Fuller test subtracts the lagged independent variable from both sides of the equation in order to check for a unit root so that:

$$\Delta Spread_t = \alpha + (\theta - 1)Spread_{t-1} + \varepsilon_t \tag{6}$$

Therefore, our null hypothesis for a unit root becomes:

H_0: $(\theta - 1)$=0

H_1: $(\theta - 1)$<0

If the null hypothesis holds and we cannot reject it, our time series has a unit root and is nonstationary. As explained by Dickey & Fuller (1979), the test statistic that we obtain follows an asymptotic distribution and we can reject our null hypothesis if the value of the test statistic is less than the 5% critical value of the distribution.

In this thesis, we will use the Augmented Dickey-Fuller test (ADF) which is simply a similar test to the one described above with more lagged differences. The purpose of this inclusion based on Verbeek (2004), is to convert the error term to a white noise in order to have valid distributional results. The selection of the number of the lags will be based on Akaike's Information Criterion (AIC) explained in Appendix I, so that we do not suffer from too many or too little laggs.

For robustness to the results that we will find from ADF tests, we will perform another conventional unit root test introduced by Phillips & Perron (1988). It is usually reffered as a nonparametric test for unit root and there is no clear indication of which of the two that we will apply at our paper is more powerful (Verbeek, 2004).Also for the PP test, we structure our null hypothesis on the assumption that the WTI-Brent spread has a unit root.

Our time series data will start after the structural break in 15 December 2010, a break date that has been empirically found by Büyükşahin et al. (2012) and Chen et al. (2015).

5.3 Auto Regressive Distributed Lag model (ARDL)

Now that we have explained the techniques that this research will utilize regarding the oil price spread relationship, it is of our interest to find out which factors explain the movement of the spread over time. A common problem with dynamic time series models is that they are built under the assumption that all variables in both sides of equation are stationary. However, time series data is frequently integrated in the first difference, meaning that few variables are non stationary at level. We will apply Augmented Dickey Fuller unit root tests as explained in section 5.2 to all our variables to check their order of integration.

More specifically, Verbeek (2004) explains that stationary variables are necessary in order for our empirical model to comply with the Gauss-Markov assumptions and provide an unbiased ordinary least squares (OLS) estimator. We refer to these assumptions in more detail in Appendix II. From section 5.2 we know that non stationary variables are not mean reverting so that if two of them are regressed together, no mechanism will show a relationship between them. What Verbeek (2004) mentions is that this type of regression will possibly lead to a high R^2 statistic, residuals with serial correlation and a statistically significant value for the coefficient of the explanatory variable. The misleading results from non stationarity in our time series is also called as spurious regression.

In order to solve problems that arise during regression of non stationary time series we will use the Autoregressive Distributed Lag Model assessed by ordinary least squares. More in detail, Pesaran et. al (1999) prove that the use of the ARDL model helps to test for longterm relationships between core variables without ending to spurious regressions since the long run coefficients can be consistent when you associate stationary with non stationary time series data or even when you only regress non stationary variables. A critical step of the model is the correct selection of lags for the dependent variable and the lags that will be used for the explanatory variables. The reason why lies to the fact that the ARDL model corrects at the same time for autocorrelation in our standard errors and for endogeneity between our explanatory variables and the error term. In order to select the appropriate number of lags for each variable, we will follow again Akaike's Information Criterion (AIC) explained in Appendix I. The time series data will be on a weekly basis and we will let lags of up to 2 months (8 lags) before.

Since we have mentioned the rationale behind using the ARDL model, we will introduce the equation of our model by using one explanatory variable on the right hand side so as to show its functionality no matter the variables we will empirically use.

$$Spread_t = a + \beta Spread_{t-1} + \varphi_0 X_t + \varphi_1 X_{t-1} + \epsilon_t \qquad (7)$$

Where:

$Spread_t$ = the difference between WTI and Brent at time t

a = a constant
$Spread_{t-1}$ = the value of the spread one period before
X_t = an independent variable that possibly explains the values of the spread
X_{t-1} = the value of the independent variable one period before
ϵ_t = the error term at time t

Based on Verbeek (2004), the above model interprets the direct impact of the explanatory variable by taking the partial derivatives of both variables so that:

$$\frac{\partial Spread_t}{\partial X_t} = \varphi_0 \tag{8}$$

It is common to refer to φ_0 as the impact multiplier since an increase of one unit at the explanatory variable has an instant impact on our spread of φ_0 units.

The long run coefficients will be obtained if we see how the partial derivatives interact when we check for the effects over period of the dependent variable:

For the next period:

$$\frac{\partial Spread_{t+1}}{\partial X_t} = \frac{\beta \partial Spread_t}{\partial X_t} + \varphi_1 = \beta\varphi_0 + \varphi_1 \tag{9}$$

After two periods:

$$\frac{\partial Spread_{t+2}}{\partial X_t} = \frac{\beta \partial Spread_{t+1}}{\partial X_t} = \beta(\beta\varphi_0 + \varphi_1) \tag{10}$$

By continuing the same process over the period of our sample based on the AIC given lags, we can determine the long run effect of a unit change in X_t which is given by the general function:

$$\varphi_0 + (\beta\varphi_0 + \varphi_1) + \beta(\beta\varphi_0 + \varphi_1) + \cdots$$
$$= \varphi_0 + (1 + \beta + \beta^2 + \cdots)(\beta\varphi_0 + \varphi_1) = \frac{\varphi_0 + \varphi_1}{1-\beta} \tag{11}$$

Therefore, if our explanatory variable increases by one unit, the expected cumulative increase in our spread is given by $\frac{\varphi_0+\varphi_1}{1-\beta}$ (Verbeek, 2004, p. 311).

5.4 Data

It is necessary for our research and especially for the execution of our methodological steps to collect data. We need to use data that corresponds to the events which were explained under chapter 4, as well as to the construction of our hypothesis for the structural break and possibly has an explanatory power regarding the oil price spread. The frequency of the data differs from one step to another since the availability of highly frequent variables, in other words daily, is rare. For our Chow tests and the Unit Root tests daily data is used since previous empirical research on commodities has shown that the frequency of the data plays an important role in order to capture more information (Narayan et al., 2013). For our regression analysis, that will help us find if there is a statistical significance between the events explained and the movement of the oil price spread, we utilize weekly data since all our explanatory variables are either in a weekly basis or in a monthly basis. In order to fit the same time interval to all variables, we interpolate our monthly data into weekly with a cubic spline interpolation. However, the weekly frequency of our data will not lead our empirical model to weak results since based on Baumeister et al. (2014), weekly frequency is able to capture a great share of the short term deviations. All our data is publicly available and the source of each variable will be described below:

5.4.1 Crude oil price data

Historical spot prices of WTI and BFOE or Brent are obtained through Thomson Reuters (2016). We construct our oil spread by deducting Brent price from WTI, so that *spread* equals $WTI - Brent$.

5.4.2 Demand variables

Since we have already stated through previous research that the U.S. oil market and consequently WTI suffers from infrastructure bottlenecks that led WTI to decouple from Brent and as a result to end their long term relationship at the end of 2010, we would like to capture what drives oil market demand in a segregated world. For this reason, we will introduce two different variables that can capture the economic health of the world and that of the U.S., separately.

Instead of using broad variables like world Gross Domestic Product (GDP) that their frequency is annual and the range of indicators that they reflect is too general, we would like to use a proxy that can capture the oil seaborne trade since the majority of oil is carried via sea. What is more, Klovland (2002) empirically found that shipping freight rates' behaviour are crucially determined by the activity of economic cycles and therefore a shipping index can be used as a proxy of the real economic activity and consequently for the world oil demand.

Furthermore, Kilian (2009) builds an index that captures global economic activity based on single-voyage freight rates for dry bulk commodities. The intuition behind using this proxy would be that increases in freight rates can explain higher demand for industrial commodities like iron ore and coal and consequently higher crude oil demand. However, the freight index cannot satisfy our sample since it is monthly available until 2009. Alternativelly, Sorensen (2009) uses the Baltic Dry Index (BDI) as a proxy for real economic activity and when he tested the coefficient correlation between Killian's Index and the BDI, he found a value of 0.96 showing great positive relationship.

We will use weekly data from Baltic Dry Index that captures indicators for demand of raw materials from countries all over the world, a fact that would be difficult to construct with time series aggregation from each country. The data is obtained from Bloomberg L.P. (2016). One should expect that a high value of BDI reflects high demand for dry commodities which consequently indicates high demand for crude oil and an upward pressure on the global oil benchmark, Brent. All else being equal, this is likely to decrease our spread. We name our variable as *BDI.*

In order to isolate U.S. macroeconomic fundamentals that possibly affect the price of WTI, we make use of the ADS daily index of Aruoba, Diebold, & Scotti (2008), who captured indicators for U.S. economy like industrial production, quarterly real GDP, monthly payroll employment and others. The ADS index takes the value of zero to reflect the average U.S. economic situation and any positive or negative fluctuation captures upturns or downturns respectively of the U.S. economy. The publicly available data, which is retrieved from the Federal Reserve Bank of Philadelphia (2016), is converted on a weekly basis as the average of seven days. One should expect that a booming U.S. economy will lead to higher demand for domestic crude oil and will drive upwards the price of WTI. We refer to this variable in our model as *ADS.* With regards to our spread, we expect that there is a positive relationship between ADS index and the oil price spread WTI-Brent

In Appendix IV, we plot together the ADS and the Dry Bulk index. After the severe fluctuations that both indexes experienced during the global financial crisis in 2008, a period which is not of our interest, we observe two important trends. Regarding the ADS index, it is fluctuating around the average value without any extreme values. Paying closer attention to the BDI, we see that the period after financial crisis is relatively stable and weak. That is to say, global economy did not show strength for importing raw materials and freight rates remained low for the whole period under discussion.

5.4.3 Physical variables: Supply and Storage

We have seen from previous literature in section 4 that supply fundamentals can make oil benchmarks move independently. In order to quantify the impact of such market conditions we will use supply proxies that capture possible movements of WTI and Brent oil separately.

It is necessary to test the explanatory power of the increasing U.S. crude oil production from the shale oil basins as well as the increasing Canadian oil imports that were headed to PADD2, Cushing oil hub. By retrieving data of operating crude oil rigs from U.S. EIA (2016), which is a gathering information from Baker Hughes, Inc. and Weatherford International, Ltd., we can control for the activity of U.S. oil upstream business. We denote *rigs* as a reference to this variable. The series is on a monthly basis dating back to 1987. U.S. EIA (2016) also provides us monthly figures regarding the thousand barrels of oil per day that are reaching PADD2 from Canada, a variable which we utilize with the name *Canada*.

Appendix V depicts the evolution of the two series mentioned above from January 2005 until May 2016. Both variables experienced exponential growth after the financial crisis and all else equal, increasing supplies of oil in the U.S. oil market will put a downward pressure to our spread especially when regional demand in PADD2 cannot satisfy the accumulation of oil quantities and transportation system is unable to allocate efficiently the refinery demand. What is more, U.S. oil rigs reach their peak during fall of 2014 and from this period and afterwards, the number of operational U.S. oil rigs experiences a steep decline. The inverse relationship that we have observed implies that the reduction in U.S. oil production might be an extra reason of the closing of the spread gap after 2014.

We hypothesized that pipeline reversals and expansions could alleviate the fragmented Cushing oil hub and help the region recover from transportation bottlenecks. In order to monitor the oil flow from PADD2 to PADD3 where the Gulf Coast refineries are located, we use monthly pipeline flows between the two petroleum districts obtained from U.S. EIA (2016). We refer to this variable as *pipe*. An increase in flows from one district to the other, *ceteris paribus*, decreases the oil quantities to PADD2 where WTI is priced, giving an increase to our oil price spread.

Appendix VI shows the evolution of the oil flow by pipe between the two PADDs.

Any disruption of oil supply or decline in production from countries that produce light sweet oil and base their crude price on Brent price formula, should push the Brent price up, and consequently decrease the WTI-Brent spread. We account for North Sea crude oil production and for crude streams that are directly substitutable with North Sea crude oil since refineries cannot easily process heavier crudes amid scarcity of light sweet oil in the global market without investing in their refinery infrastructure. For this reason, a proxy *lightoil* will be used and will represent the production of countries like UK, Norway, Nigeria, Angola, Libya and Egypt. These countries are the major producers of light sweet oil excluding U.S., a fact which we can observe when we analyze the physical characteristics of few core crude streams in Table 4. We have already mentioned the quality characteristics of U.K. and Norway crude streams under the BFOE physical characteristics (Table 2). In order to compile oil production from the above countries we aggregate monthly oil production from each country obtained from Bloomberg L.P. (2016).

By including the light oil producing countries like Libya, we also account for geopolitical events that took out of the place great quantities of production. Decreasing light oil production outside U.S., *ceteris paribus*, increases the price of Brent which consequently decreases the spread. So we expect a positive relationship between light oil production and our spread. In Appendix VII, we plot the aggregate oil production of the above mentioned countries in thousand bpd.

Table 3: Physical Characteristics of African crudes

Crude oil	Country	API gravity	Sulfur Content
Bonny Light	Nigeria	33.6	0.14
Qua Iboe	Nigeria	35.2	0.12
Escravos	Nigeria	34.4	0.17
Cabinda	Angola	32.5	0.13
Brega	Libya	39.8	0.2
Es Sider	Libya	36.3	0.44
Zarzaitine	Algeria	42.8	0.06

Source: (Kaufmann & Banerjee, 2014)

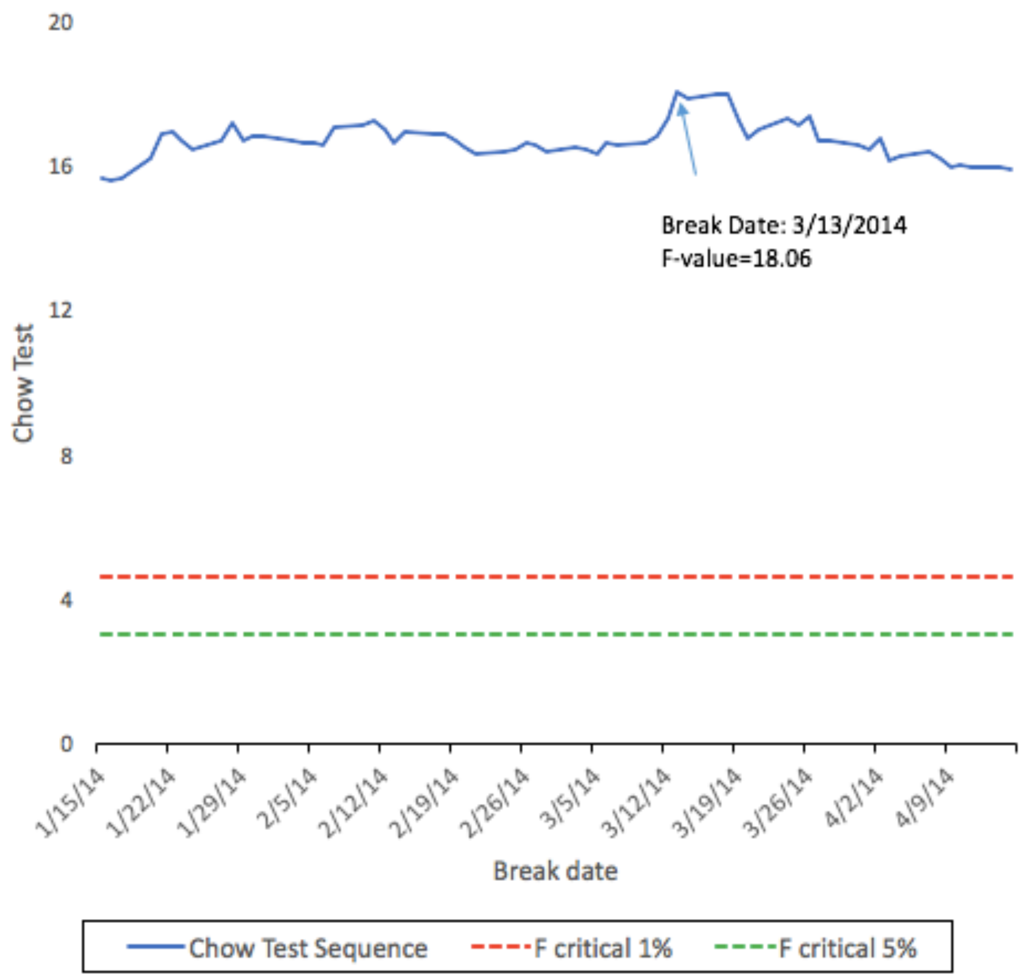
6. Results and Analysis

Following the methodological steps explained above, this section serves to present the results that we obtained after running Chow tests for structural break points, Unit root tests for a changing pattern in the spread relationship and ARDL models in order to find any long run relationship between the spread and explanatory variables as introduced in Data section.

6.1 Chow test results

In the figure below, we present the results after performing Chow tests during the period that we hypothesized that the oil spread of WTI-Brent may have experienced a structural break. A table which provides in numerical form all the F-values as well as the p-values is provided under Appendix VIII.

Figure 12: Chow test values



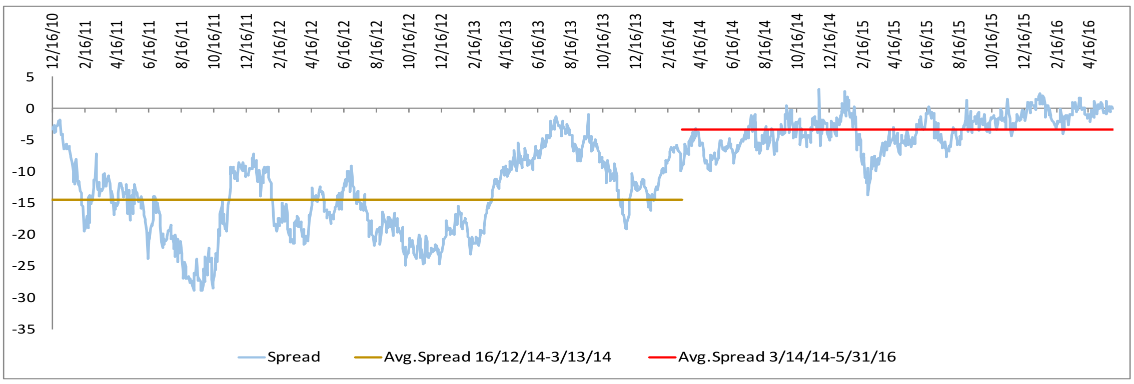
Interestingly, we observe that the sample chosen to be tested for structural break, rejects the null hypothesis of no structural break and we can see from Appendix IX that the intercepts as well as the slopes of our autoregressive equations change between the sub samples divided by the above dates. Following our argumentation above we select the date with the highest F-value as our structural break point that is significant at the 5% and 1% critical level, i.e. we obtain the 13th of March 2014. P-values are all significant under our sample since this is the period that the relationship between the two oil benchmarks faces a structural break. The Table below, summarizes the values of the test when the break date is set at 13/03/2014 while Appendix VIII also shows the mechanics of computing the F-value under each break date when you have obtained the SSRs.

Table 4: Break test for WTI-Brent spread

$Spread_t = a + \beta Spread_{t-1} + u_t$	H_0: No break in a and β
Chow Test	18.06
1% Critical Value	4.61
5% Critical Value	3
Null of no break	Reject
Break Date	13/03/2014

Based on our results, we can come up with the conclusion that the WTI-Brent spread dynamics have changed in early 2014 and the events that we presented at our hypothesis seem to have played a major role in the movement of the spread. The microstructure of the U.S. oil market changed in order to adapt to the new fundamentals of the oil industry and this drove the WTI-Brent spread to narrow down. The figure below shows the crude oil price spread for the period before our structural break and the period after the structural break. Chow test considers the average change statistically significant since the oil spread approaches $11.01. From 16/12/2010 till the break date, the mean value of the spread was -$14.43 and from the break date till 31/05/2016 the mean value went down to -$3.42. It remains to be seen if this structural change is persistent and drove our oil spread to a new stationary process. We should remind again that the previous stationary process of the WTI-Brent spread, had the U.S. oil benchmark trading at a premium against Brent.

Figure 13: Crude oil price spread based on the difference of WTI-Brent



Source: Author via Thomson Reuters (2016)

6.2 Analysis of Unit root tests results

Having rejected the hypothesis of no structural break date for our sample period, we divide the time series of WTI-Brent to two subsamples to check if the events that took place under our hypothesis of a structural change, formed a new WTI-Brent stationary process. The tables below exhibit the results for the sub-sample defined by our structural break till the end of May of 2016 but also the non stationary time series process that previous papers have examined based on the structural break of 15/12/2010.

Table 5: Augmented Dickey Fuller test

Sample periods	16/12/2010-13/03/2014	14/03/2014-31/05/2016
Number of lags (AIC)	3	3
ADF test (p-value)	-2.68 (0.0771)*	-2.97 (0.0381)**
10% critical value	-2.56	-2.56
5% critical value	-2.86	-2.86
1% critical value	-3.43	-3.43

Table 6: Phllips and Perron test

Sample periods	16/12/2010-13/03/2014	14/03/2014-31/05/2016
PP test (p-value)	-2.85 (0.515)*	-5.53 (0.0000)***
10% critical value	-2.56	-2.56
5% critical value	-2.86	-2.86
1% critical value	-3.44	-3.44

Both unit root tests for the first subsample cannot reject the null hypothesis of no stationarity at the 5% confidence level, which has as a result to confirm the non stationary relationship that previous papers have found. The spread indeed changed its relationship after 2010 with profound discounts of WTI price against Brent.

Our results for the second sub sample, which covers the period after our structural break date, are striking. Both ADF and PP tests reject the null hypothesis of a unit root since our subsample is statistically significant at the 5% level when using ADF test and statistically significant at the 1% level when using the PP test. Hence, it is valid to believe that the WTI-Brent relationship has changed one more time and the new stationary process is governed by a small discount of WTI against Brent. Appendix X shows both ADF and PP tests in detail.

The above result should be interpreted from the oil industry since market fundamentals have changed and the tight relationship between the two oil benchmarks signals new patterns compared to the ones that they experienced during the non stationary period.

6.3 Regression results

In Appendix XI, apart from the general statistics of our candidate variables, we show that none of them is stationary at level which means that the variables that we will analyze at our ARDL model are integrated at the first difference. What is more, Appendix XII depicts the coefficients of correlation among the core variables of our empirical model.

Under Table 8, a summary of the long term coefficients after running our regression is presented. Each model specification controls for different combination of variables and the lags that are used for the dependent and independent variables are shown under the *Notes*. The first model accounts for all the variables explained in section 5.4, the second monitors the fluctuation of the spread based on supply fundamentals, the third takes into consideration only U.S. conditions while the last one accounts for only supply and demand factors excluding pipeline oil flows.

Table 7: Summary of ARDL model results

	ARDL1_all	ARDL2_supply	ARDL3_U.S.	ARDL4_s&d
Intercept	-25.337272 (31.647113)	-21.787553 (33.320683)	**21.550051**** **(9.550223)**	-24438422 (38.167766)
ADS	-1.993293 (1.646205)	-----	-1.264018 (1.895713)	-1.045476 (1.998563)
BDI	**0.0000855*** **(0.000487)**	-----	-----	0.001093 (0.000664)
Rigs	-0.003410 (0.004218)	-004926 (0.004120)	**-0.007162**** **(0.003526)**	**-0.010733**** **(0.004339)**
Canada	-0.012380 (0.008034)	-0.013801 (0.008890)	**-0.019891**** **(0.009747)**	0.010856 (0.008987)
Pipe	**0.001147***** **(0.000407)**	**0.001108***** **(0.000418)**	**0.000932**** **(0.000418)**	-----
Lightoil	0.003104 (0.002487)	0.003424 (0.002639)	-----	0.000894 (0.002655)
N	589	589	589	589

Notes of the table: Sample period: January 7, 2005 to May 27, 2016
Standard errors in parentheses, * $p < 0.10$, ** $p<0.05$ ***$p<0.01$
# of lags used for each model (spread, ADS, BDI, Rigs, Canada, Pipe, Lightoil)
ARDL1_all (6,0,0,1,0,2,1)
ARDL2_supply (6,-,-,1,0,2,1)
ARDL3_U.S. (6,0,-,0,0,2,-)
ARDL4_s&d (6,0,0,0,0,-,1)

6.3.1 Demand Variables

ADS for U.S. economy and Baltic Dry Index (BDI) for global economic activity prove to be insignificant and unable to explain the WTI-Brent spread movement in the long run. Even the signs in front of the coefficients of the two independent variables are counterintuitive to what we expected when we introduced our data.

6.3.2 Supply Variables

Unlike demand variables, in all four ARDL models we find statistical significance for one or more of the supply variables. To start with, whenever we control for the pipeline oil flows from PADD2, Midwest to PADD3, Gulf Coast, the variable *pipe* is positively and statistically significant with the *spread.* More in detail, *pipe* is positively related and significant at 1% confidence level when we control for all the variables in our model as well as when we only account for supply fundamentals. The level of significance decreases to 5% when we associate *spread* with U.S. conditions. That is to say, pipeline infrastructure is a crucial factor for the WTI benchmark since, as we have already mentioned, WTI's physical location renders it dependent to inland transportation. An increased pipeline flow from reversals and expansion of pipes between the two petroleum districts not only alleviates the storage levels at Cushing oil hub but also brings transportation costs of crude oil down compared to other methods of transport (barge, rail) from PADD2. All these events give an upward pressure to WTI leaving Brent price unchanged.

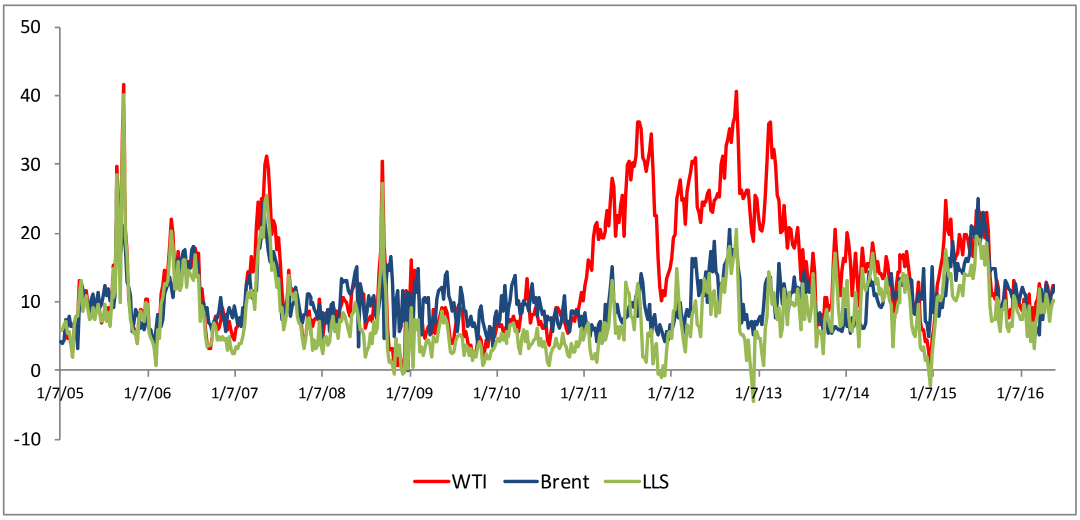
Regarding U.S. oil supplies, we find statistical significance of the variables when we only control for U.S. conditions and when we leave out of our model pipeline oil flows. Nonetheless, U.S. supply variables maintain the sign of their long term relationship with the spread even when their statistical significance is below the 10% confidence level. Notably, the independent variable *rigs*, which is used as a proxy of U.S. oil production, has a negative relationship with our spread, a fact which we predicted from our data introduction. The inverse long term relationship becomes stronger at the 5% confidence level when we do not account for the variable *pipe*. Greater number of U.S. operational rigs decreases our spread, a fact which we observed with the increased U.S. shale oil production. Canadian imports are also negatively associated with the movement of our spread and we find statistical significance for the proxy *Canada* when we only control for U.S. market conditions. These findings support our conjecture that indeed U.S. supply fundamentals have a significant role in explaining the movement of WTI-Brent spread in the long run.

Last, we observe that the production of sweet light production out of U.S. has no explanatory power regarding the movement of the spread. However, we should mention that through all our model specifications, the sign of *lightoil* implies a positive long run relationship with the spread, a fact which we surmised on our data section. Decreasing values of light oil production outside U.S. will increase the price of Brent and put a downward pressure to the spread leaving WTI unchanged.

7. Interpretation of results

Under this section, we would like to explain the implications of the WTI-Brent relationship from the moment that the two crude oil prices decoupled until today, where they have formed a new close relationship. In order to do that, we provide in Figure 14 the 3:2:1 crack spread for WTI, Brent and LLS. Based on U.S. Energy Information Administration (2013), a crack spread reflects the difference between the cost of purchasing crude oil and the price of selling finished oil products. In other words, it acounts for how much is the profit margin for a refinery when it purchases crude oil priced at a certain benchmark and sells gasoline or distillate fuel at a certain price. The rationale behind 3:2:1 is that 3 barrels of crude oil at a typical U.S. refinery will yield 2 barrels of gasoline and 1 barrel of distillate fuel.

Figure 14: Crack spreads 3:2:1 for WTI, Brent and LLS



Source: (Bloomberg L.P., 2016)

We observe that, at the end of 2010, WTI crack spread decouples from the other crack spreads. The period coincides with the significant discounts of WTI price due to the increase in U.S. crude oil production and Canadian imports, in combination with transportation infrastructure bottlenecks as we explained in above sections. The key difference is that WTI crack spread is found at a premium compared to other benchmarks which means that U.S. Midwest refineries, of which their feedstock was priced at WTI benchmark, were making excess profits compared to their competitors worldwide.

In order to explain the competitive advantage that U.S. Midwest refiners obtained during U.S. shale oil revolution, we follow Kilian (2014) to present few key differences between downstream and upstream business at U.S. oil market. To start with, although there was a ban in U.S. crude oil exports until December 2015 (The Wall Street Journal, 2015), the ban had no effect to oil product exports. Furthermore, the pipeline network of refined products is not associated with the crude oil pileline system which means that refiners could export their finished oil products with no difficulty. Bearing these two in mind, U.S. refiners in the Midwest reached global markets by charging the same fuel price with the refineries all over the world less transportation costs. This fact did not let U.S. gasoline prices to decouple from the other world since the extra oil products were exported due to the low cost that Midwest U.S. refiners had to cover in order to convert crude oil into finished oil products. Based on U.S. EIA (2016) U.S. oil product exports increased during the period under research and captured a greater share of world product market.

On the other hand, the period of great discounts for the WTI price was not good news for the U.S. upstream business since they had to sell their output at a lower price compared to the price they could find at global markets. However, based on our research and results, the new stationary process that WTI-Brent spread formed after 13 March 2014, worked in favor of the U.S. oil producers since pipeline expansions helped crude oil flows to reach the markets with oil demand and increase the price of WTI leaving Brent unchanged. The development of the crude oil transportation infrastructure took back the benefits that U.S. refiners were reaping from the increased U.S. oil production. This fact becomes more obvious when we see that at Figure 14, WTI crack spread decreases chronologically close with the period where new pipeline infrastructure came in place and pipe oil flows from PADD2 to PADD3 increased substantially as shown in Figure 10. From this period and after, WTI crack spread stays close with Brent as well as with LLS crack spread.

8. Conclusions

Our research focused on the oil price spread between WTI and Brent. The spread was in a long term stationary relationship until the end of 2010 where WTI was trading at a small premium against Brent, reflecting slight quality differences and freight costs so that crude oil streams priced at Brent could reach U.S. oil market. Large discounts of the WTI price caused the spread to significantly decrease in early 2011 with researchers finding a structural change in the relationship between WTI and Brent on 15 December 2010. The oil spread formed a non stationary relationship. The events that caused the two oil prices to diverge were: U.S. shale oil revolution along with Canadian oil imports to PADD2 substantially increased the storage levels at Cushing, Oklahoma where WTI oil benchmark is priced. Increasing supplies were coincided with inadequate transportation infrastructure, inability to export crude oil and unsuitable refinery configurations. This caused the mismatch of the new oil supply and markets with the oil demand, driving WTI price down. On the Brent side, the Arab Oil Spring and the Fukushima accident in 2011 had an upward impact on the price of Brent.

We hypothesized that spread experienced a new structural break in early 2014 and we used Chow F-tests to verify our assumptions. The events that we built our hypothesis had to do with the microstructure of the U.S. oi market and especially the evolution of the oil transportation system in response to the oil glut caused by the U.S. shale oil boom. The oil price spread experienced a new structural break with an estimated date on 13 March 2014. From this period till today the spread is in a new stationary process with Brent trading at a small premium against WTI. We tested the above statement with unit root tests over the spread time series data.

Last but not least, our aim of research was to test the explanatory power of variables with regards to the fluctuation of the oil price spread. We used the Autoregressive Distributed Lag Model in order to verify whether there is a long term relationship between the spread and the regressors. After controlling for different set of variables in four ARDL specifications, we find that the spread has a positive long term relationship with the pipeline oil flows from PADD2 to PADD3. That is to say expansions and reversals of the pipeline system are associated with an upward pressure of the price spread. On the other hand, U.S. oil production and Canadian oil imports to PADD2 show that they are in an inverse long run relationship with the spread. We interpret the above result as increasing oil quantities at Cushing will always put a downward pressure to the spread as long as oil quantities exceed transportation capacities.

8.1 Limitations of the Research

Although our results regarding the relationship between the spread and the physical side of crude oil market are influential, we should mention few limitations that will help the reader to better understand the relationships shown above as well as will give him/her few thoughts for further research. First of all, when we introduced our data selection, we did not refer to the financial side of WTI and Brent. NYMEX in U.S. and ICE Futures Europe are the futures markets where both commodities are highly liquid since crude oil is considered to be the most traded commodity in the world. Therefore, our research does not test the explanatory power of the paper-market and the speculative effect that traders have on the level price for both crude oil benchmarks.

Furthermore, modelling and testing for explanatory power for all oil market fundamentals is a work that needs great investment in time and access to more datasets since there are numerous factors affecting the price spread between WTI and Brent. An empirical model which also accounts for undisclosed data will help our research to give a greater picture of the oil market and the relationship between the two oil benchmarks. What is more, we run our regression on a weekly publicly available data. As explained under Data section, few of our variables could be found under weekly frequency. More in detail, Canadian oil imports, Pipeline Oil flows from PADD2 to PADD3, U.S. oil rigs and Light oil production are monthly time series data which we had to interpolate in order to fit in our regression. This fact may have led to minor measurement errors.

Last, we should state our reservations with regards to the Baltic Dry Index as a macroeconomic indicator which captures the global demand for transportation of raw materials and hence a direct link with real economic activity. Although, demand for shipping raw materials is reflected via this index, other factors also play an important role that may decrease the power of indicating the level of demand around the world. More specifically, freight rates also account for the supply side of the shipping market which means that overcapacity issues that the industry is experiencing the last years will weigh on the level of the index (BIMCO, 2016). Whatsoever, past papers that we made a reference during the introduction of the BDI proxy, did use the freight rates as an indicator for real economic activity since demand from countries like India or China are difficult to be captured in such a high frequency otherwise.

8.2 Suggestions for Further Research

Our findings show that U.S. refiners, during U.S. shale oil revolution, captured greater market share in the global oil product market. As explained, this was a result of cheap feedstock into their refineries that gave them a competitive advantage compared to other refineries around the world. The new tight stationary relationship between WTI and Brent implies that this high margins in production are gone for U.S. refiners. This leads us to set as a subject for further research the following question: What is the impact on the U.S. downstream business after finding itself competing again with equal crack spreads?

Bibliography

- Adelman, M. A. (1984). International Oil Agreements. *The Energy Journal, 5*(3), 1-9.
- Alizadeh, A. H., & Nomikos, N. (2004). Cost of carry, causality and arbitrage between oil futures and tanker freight markets. *Transportation Research Part E, 40*, 297-316.
- Aruoba, S. B., Diebold, F., & Scotti, C. (2008). *Real-Time Measurement of Business Conditions.* NATIONAL BUREAU OF ECONOMIC RESEARCH. Cambridge: NBER.
- Büyükşahin, B., Lee, T., Moser, J., & Robe, M. (2012). *Physical Market Conditions, Paper Market Activity, and the WTI-Brent Spread.* U.S. Energy Information Administration. U.S. Energy Information Administration.
- Baumeister, C., Guérin, P., & Kilian, L. (2014). *Do High-Frequency Financial Data Help Forecast Oil Prices? The MIDAS Touch at Work.* International Economic Analysis Department. Ontario: Bank of Canada.
- BBC. (n.d.). Retrieved from BBC: http://www.bbc.co.uk/timelines/zqgxtfr#zs64kqt
- BIMCO. (2016, May 2). Retrieved August 17, 2016, from BIMCO: https://www.bimco.org/Reports/Market_Analysis/2016/0502_SH_Interview.aspx
- Bloomberg L.P. (2016, August 15). 3:2:1 Crack Spreads for WTI Brent LLS from 7/1/2005 to 12/8/2016. New York, New York, United States of America.
- Bloomberg L.P. (2016, July 22). Crude Oil Production for Nigeria, Angola, Libya, Algeria, UK, Norway from January 2005 to May 2016. New York, New York, United States of America.
- Bloomberg L.P. (2016, July 22). Weekly Baltic Dry Index from 7/1/2005 to 27/5/2016. New York, New York, United States of America.
- Chen, W., Huang, Z., & Yi, Y. (2015, June 16). Is there a structural change in the persistence of WTI–Brent oil price spreads in the post-2010 period? *Economic Modelling*, 64-71.
- Chow, G. C. (1960). Tests of Equality Between Sets of Coefficients in Two Linear Regressions. *Econometrica, 28*(3), 591-605.

- Darbouche, H., & Fattouh, B. (2011). *The Implications of the Arab Uprisings for Oil and Gas Markets.* Oxford: Oxford Institute for Energy Studies.
- Deutsche Bank. (2013). *Oil & Gas for Beginners.* London.
- Dickey, D., & Fuller, W. (1979, June). Distribution of the Estimators for Autoregressive Time Series With a Unit Root. *Journal of the American Statistical Association, 74*(366), 427-431.
- EIA. (2016, June 16). Retrieved July 10, 2016, from U.S. Energy Information Administration: http://www.eia.gov/energyexplained/index.cfm?page=oil_home
- Fattouh, B. (2010). The dynamics of crude oil price differentials. *Energy Economics, 32*, 334-342.
- Fattouh, B. (2011). *An Anatomy of the Crude Oil Pricing System.* Oxford Institute for Energy Studies.
- Federal Reserve Bank of Philadelphia. (2016). Retrieved from https://www.philadelphiafed.org/research-and-data/real-time-center/business-conditions-index
- Fielden, S. (2013, February 3). Retrieved July 24, 2016, from RBN Energy: https://rbnenergy.com/brent-wti-reconciliation-dreams-shattered-post-super-bowl-analysis
- Fielden, S. (2013, April 4). Retrieved July 24, 2016, from RBN Energy: https://rbnenergy.com/crude-loves-rock-n-rail-east-coast-delivery-terminals
- Gülen, S. (1999). Regionalization in the world crude oil market: further results. *Energy Journal*, 125-139.
- Gulen, S. G. (1997). Regionalization in the world crude oil market. *Energy Journal*, 109-126.
- Hammoudeh, S. M., Ewin, B., & Thompson, M. (2008). Threshold Cointegration Analysis of Crude Oil Benchmarks. *The Energy Journal, 29*, 79-95.
- Hayashi, M., & Hughes, L. (2012, December). The Fukushimanuclearaccidentanditseffectonglobal energy security. *Energy Policy, 59*(2013), 102-111.
- Heier, M., & Skoglund, S. (2014). *Crude Oil Price Differentials.* Norwegian School of Economics, Department of Financial Economics, Bergen.

- Horsnell, P., & Mabro, R. (1993). *Oil markets and prices: the Brent market and the formation of world oil prices.* UK: Oxford University Press.
- IHS. (2014). *US Crude Oil Export Decision: Assessing the Impact of the Export Ban and Free Trade on the US Economy.* Houston: IHS Global Inc.
- Kaufmann, R. K., & Banerjee, S. (2014, November). A unified world oil market: Regions in physical, economic, geographic, and political space. *Energy Policy, 74*, 235-242.
- Keystone XL. (2016). Retrieved July 24, 2016, from TransCanada Keystone XL: http://www.keystone-xl.com/about/the-keystone-xl-oil-pipeline-project/
- Kilian, L. (2009). Not All Oil Price Shocks Are Alike: Disentangling Demand and Supply Shocks in the Crude Oil Market. *American Economic Review, 99*(3), 1053-1069.
- Kilian, L. (2014). *The impact of the shale oil revolution on U.S. oil and gasoline prices.* Frankfurt: CFS Working Paper Series.
- Kleit, A. N. (2001). Are Regional Oil Markets Growing Closer Together?: An Arbitrage Cost Approach. *The Energy Journal, 22*(2), 1-15.
- Klovland, J. T. (2002). *Business cycles, commodity prices and shipping freight rates.* Bergen: SNF NH.
- Lee, H. B. (2008). Using the Chow Test to Analyze Regression Discontinuities. *Tutorials in Quantitative Methods for Psychology, 4*(2), 46-50.
- McRaey, S. (2015). University of Michigan, Department of Economics,.
- Milonas, N., & Henker, T. (2001). Price spread and convenience yield behaviour in the international oil market. *Applied Financial Economics, 11*, 23-36.
- Narayan, P., Narayan, S., & Sharma, S. (2013). *An analysis of commodity markets: What gain for investors?* Deakin University, School of Accounting, Economics and Finance, Melbourne,.
- Newbery, D. (1984). *The Manipulation of Futures Markets by a Dominant Producer.* (R. Anderson, Ed.) Lexington, MA: Lexington Books.
- Pesaran, H., Shin, Y., & Smith, R. (1999). Pooled mean Group Estimation of Dynamic Heterogeneous Panels. *Journal of the American Statistical Association*, 621-634.
- Phillips, P. C., & Perron, P. (1988, June). Testing for a Unit Root in Time Series Regression. *Biometrika, 75*(2), 335-346.

- Purvin & Gertz. (2010). *The Role of WTI as a Crude Oil Benchmark.* CME.
- Quandt, R. E. (1960, June). Tests of the Hypothesis that a Linear Regression System Obeys Two Separate Regimes. *Journal of the American Statistical Association, 55*(290), 324-330.
- Reboredo, J. C. (2011). How do crude oil prices co-move? A copula approach. *Energy Economics, 33*, 948-955.
- Seaway Pipeline. (2016). Retrieved July 24, 2016, from Seaway Crude Pipeline Company LLC (Seaway): http://seawaypipeline.com
- SØRENSEN, L. Q. (2009). *Oil Price Shocks and Stock Return Predictability.* Norwegian School of Economics and Business Administration, Department of Finance and Management Science, Bergen.
- The Wall Street Journal. (2015, 12 16). Retrieved 8 15, 2016, from http://www.wsj.com/articles/congressional-leaders-agree-to-lift-40-year-ban-on-oil-exports-1450242995
- Thomson Reuters. (2016, July 11). Spot prices for WTI Cushing and Dated Brent from 16/12/2010 until 31/5/2016. United States of America.
- U.S. Department of State. (2013, March). Retrieved July 19, 2016, from U.S. Department of State: https://keystonepipeline-xl.state.gov/documents/organization/221218.pdf
- U.S. EIA. (2012, July 16). Retrieved July 15, 2016, from U.S. Energy Information Administration: http://www.eia.gov/todayinenergy/detail.cfm?id=7110
- U.S. EIA. (2012b). Retrieved July 11, 2016, from U.S. EIA: http://www.eia.gov/petroleum/data.cfm
- U.S. EIA. (2014). *World Oil Transit Chokepoints.*
- U.S. EIA. (2015, August 10). Retrieved July 15, 2016, from http://www.eia.gov/todayinenergy/detail.cfm?id=22432
- U.S. EIA. (2016, May). Retrieved July 14, 2016, from U.S. Energy Information Administration: http://www.eia.gov/petroleum/data.cfm
- U.S. EIA. (2016). Retrieved July 15, 2016, from U.S. Energy Information Administration: http://www.eia.gov/state/maps.cfm?src=home-f3
- U.S. EIA. (2016). Retrieved July 24, 2016, from U.S. EIA: http://www.eia.gov/petroleum/data.cfm#imports

- U.S. EIA. (2016). Retrieved 2016, from U.S. Energy Information Administration: http://www.eia.gov/dnav/pet/hist/LeafHandler.ashx?n=pet&s=wrpexus2&f=4
- U.S. EIA. (2016, July 22). Movements between PAD Districts by Pipeline. United States of America.
- U.S. EIA. (2016, July 22). PAD District Imports by Country of Origin from 7/1/2005 to 27/5/2016. United States of America.
- U.S. EIA. (2016, July 22). Rotary Rigs in Operation from 7/1/2005 to 27/5/2016. United States of America.
- U.S. Energy Information Administration. (2012, February 7). Retrieved July 11, 2016, from U.S. Energy Information Administration: http://www.eia.gov/todayinenergy/detail.cfm?id=4890
- U.S. Energy Information Administration. (2013, 2 4). Retrieved 8 15, 2016, from U.S. Energy Information Administration: https://www.eia.gov/todayinenergy/includes/CrackSpread_Explain.cfm
- Verbeek, M. (2004). *A Guide to Modern Econometrics* (2nd Edition ed.). John Wiley & Sons Ltd.
- Weiner, R. J. (1991). Is the world oil market `one great pool'? *Energy Journal, 12*(3).
- Wooldridge, J. M. (2012). *Introductory Econometrics: A Modern Approach* (Fifth Edition ed.). Michigan: South-Western, Cengage Learning.

Appendices

I. Akaike Information Criterion

The Akaike's Information Criterion provides you with the number of lags which you can include in a model based "on a trade-off between goodness-of-fit and the number of parameters used to obtain that fit." (Verbeek, 2004, p. 285)

$$AIC(p) = \ln\left[\frac{SSR(p)}{T}\right] + (p+1)\frac{2}{T}$$

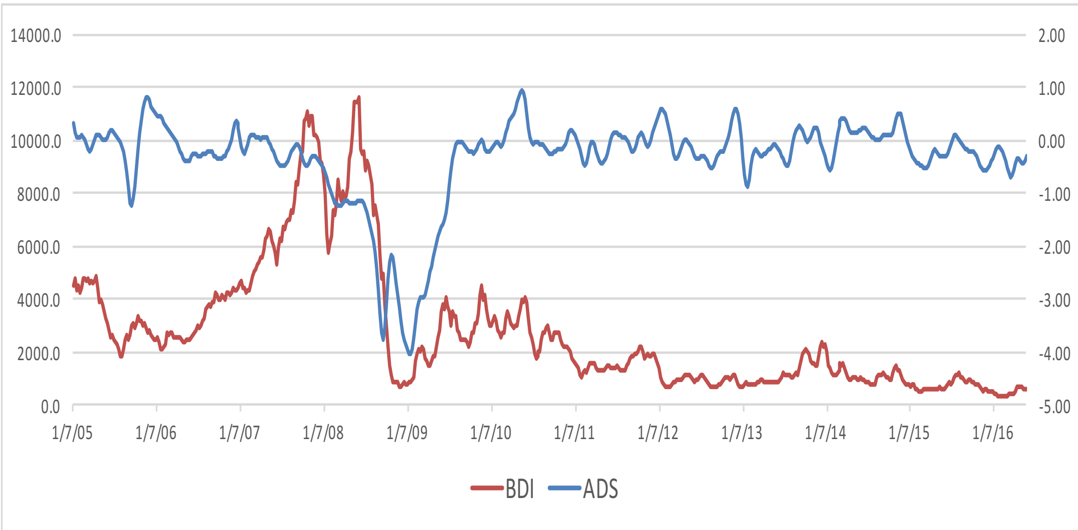
Where:
SSR(p): the sum of squared residuals for the estimated AR(p)

II. Gauss-Markov Assumptions

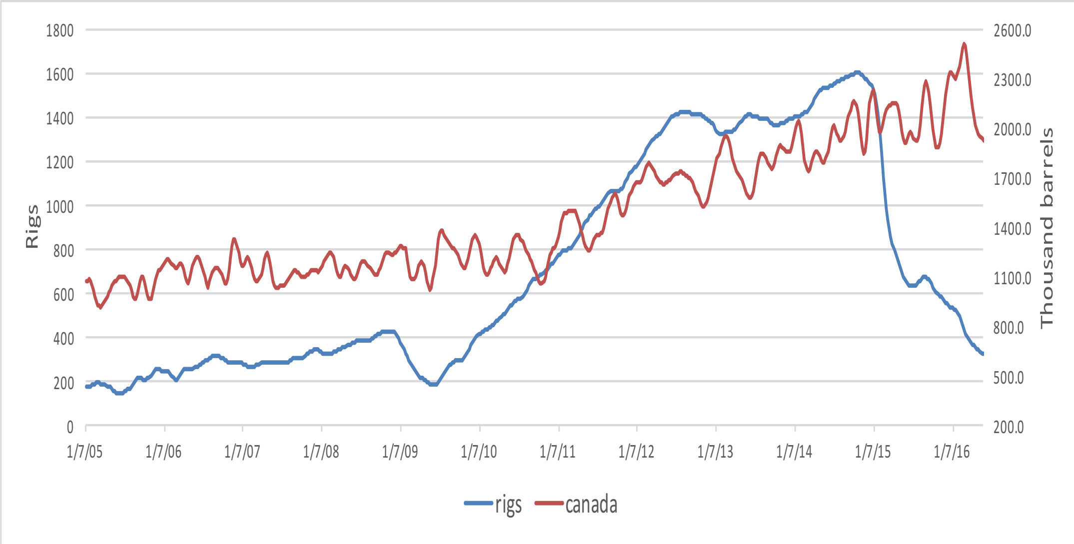
These are the requirements that that time series data should meet when they are regressed with Ordinary Least Squares:

- Linear parameters for the population
- No perfect collinearity
- No heteroscedasticity in the error term
- No serial correlation
- Zero means
- Normality

III. ADS-BDI indexes



IV. U.S. crude oil production and Canadian imports



V. Pipe oil flows from PADD2 to PADD3

VI. Sweet light oil production

VII. Chow test values and statistical significance

Break Date	Chow Test Sequence	p-values
15/01/14	15.68	(0.0000)***
16/01/14	15.59	(0.0000)***
17/01/14	15.63	(0.0000)***
20/01/14	16.21	(0.0000)***
21/01/14	16.86	(0.0000)***
22/01/14	16.95	(0.0000)***
23/01/14	16.69	(0.0000)***
24/01/14	16.44	(0.0000)***
27/01/14	16.68	(0.0000)***
28/01/14	17.17	(0.0000)***
29/01/14	16.67	(0.0000)***
30/01/14	16.8	(0.0000)***
31/01/14	16.8	(0.0000)***
03/02/14	16.72	(0.0000)***
04/02/14	16.63	(0.0000)***
05/02/14	16.62	(0.0000)***
06/02/14	16.56	(0.0000)***
07/02/14	17.08	(0.0000)***
10/02/14	17.15	(0.0000)***
11/02/14	17.26	(0.0000)***
12/02/14	17.01	(0.0000)***
13/02/14	16.66	(0.0000)***
14/02/14	16.92	(0.0000)***
17/02/14	16.85	(0.0000)***
18/02/14	16.86	(0.0000)***
19/02/14	16.7	(0.0000)***
20/02/14	16.5	(0.0000)***
21/02/14	16.3	(0.0000)***
24/02/14	16.4	(0.0000)***
25/02/14	16.45	(0.0000)***
26/02/14	16.6	(0.0000)***
27/02/14	16.56	(0.0000)***
28/02/14	16.39	(0.0000)***
03/03/14	16.48	(0.0000)***
04/03/14	16.44	(0.0000)***
05/03/14	16.34	(0.0000)***
06/03/14	16.6	(0.0000)***
07/03/14	16.55	(0.0000)***
10/03/14	16.62	(0.0000)***
11/03/14	16.83	(0.0000)***
12/03/14	17.31	(0.0000)***
13/03/14	**18.06**	**(0.0000)*** **
14/03/14	17.87	(0.0000)***
17/03/14	17.98	(0.0000)***
18/03/14	17.97	(0.0000)***
19/03/14	17.23	(0.0000)***
20/03/14	16.75	(0.0000)***
21/03/14	16.98	(0.0000)***
24/03/14	17.29	(0.0000)***
25/03/14	17.09	(0.0000)***
26/03/14	17.39	(0.0000)***
27/03/14	16.67	(0.0000)***
28/03/14	16.68	(0.0000)***
31/03/14	16.58	(0.0000)***
01/04/14	16.45	(0.0000)***
02/04/14	16.78	(0.0000)***
03/04/14	16.13	(0.0000)***
04/04/14	16.26	(0.0000)***
07/04/14	16.39	(0.0000)***
08/04/14	16.2	(0.0000)***
09/04/14	15.94	(0.0000)***
10/04/14	16	(0.0000)***
11/04/14	15.97	(0.0000)***
14/04/14	15.94	(0.0000)***
15/04/14	15.92	(0.0000)***

Notes: The break date where the test gives the highest number should be treated as our structural break date. P-values are shown for each test inside the parenthesis and stars(*, **, ***) indicate different statistical significance levels (10%, 5%, 1% respectively)

Chow F test for structural change on 13th March 2014

Samples	Periods	Sum Square Residuals
Full Sample	(16/12/2010-31/05/2016)	2534.4
Sub-Sample1	(16/12/2010-13/03/2014)	1479.6
Sub-Sample2	(14/03/2014-31/05/2016)	992.4

$$F = \frac{(SSRr - SSRu)/k}{SSRu/(T-2k)} = \frac{(SSRr - SSR1 - SSR2)/k}{(SSR1 + SSR2)/(T-2k)} =$$

$$F = \frac{(253.4 - 1479.6 - 992.4)/2}{(1479.6 + 992.4)/[1424 - (2 * 2)]} =$$

$$F = 18.06$$

$$Fcrit = Fk, T - 2k, 0.01 = 4.61$$

$$F > Fcrit$$

We reject our null hypothesis of no structural break on 13/03/2014.

VIII. SSR and coefficients for restricted and unrestricted models

Full sample

Dependent Variable: WTI_BRENT
Method: Least Squares
Date: 07/27/16 Time: 20:22
Sample (adjusted): 12/17/2010 5/31/2016
Included observations: 1423 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.155709	0.059156	-2.632197	0.0086
WTI_BRENT(-1)	0.984153	0.004754	207.0195	0.0000

R-squared	0.967907	Mean dependent var	-9.966810
Adjusted R-squared	0.967885	S.D. dependent var	7.452269
S.E. of regression	1.335500	Akaike info criterion	3.417894
Sum squared resid	2534.441	Schwarz criterion	3.425287
Log likelihood	-2429.831	Hannan-Quinn criter.	3.420655
F-statistic	42857.06	Durbin-Watson stat	2.611890
Prob(F-statistic)	0.000000		

Sub-Sample 1

Dependent Variable: WTI_BRENT
Method: Least Squares
Date: 07/27/16 Time: 20:27
Sample (adjusted): 12/17/2010 3/13/2014
Included observations: 845 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.357841	0.115039	-3.110609	0.0019
WTI_BRENT(-1)	0.975689	0.007317	133.3418	0.0000

R-squared	0.954733	Mean dependent var	-14.44213
Adjusted R-squared	0.954680	S.D. dependent var	6.223331
S.E. of regression	1.324857	Akaike info criterion	3.402850
Sum squared resid	1479.672	Schwarz criterion	3.414067
Log likelihood	-1435.704	Hannan-Quinn criter.	3.407148
F-statistic	17780.03	Durbin-Watson stat	2.460110
Prob(F-statistic)	0.000000		

Sub-Sample 2

Dependent Variable: WTI_BRENT
Method: Least Squares
Date: 07/27/16 Time: 20:29
Sample: 3/14/2014 5/31/2016
Included observations: 578

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.383674	0.085871	-4.468013	0.0000
WTI_BRENT(-1)	0.883905	0.019268	45.87462	0.0000

R-squared	0.785113	Mean dependent var	-3.424170
Adjusted R-squared	0.784740	S.D. dependent var	2.829231
S.E. of regression	1.312653	Akaike info criterion	3.385432
Sum squared resid	992.4816	Schwarz criterion	3.400517
Log likelihood	-976.3899	Hannan-Quinn criter.	3.391314
F-statistic	2104.481	Durbin-Watson stat	2.681957
Prob(F-statistic)	0.000000		

IX. ADF and PP tests before and after structural break date

Table: Sample before the structural break-ADF test
We cannot reject the null hypothesis of a unit root at the 5%
Null Hypothesis: WTI_BRENT has a unit root
Exogenous: Constant
Lag Length: 3 (Automatic - based on AIC, maxlag=20)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-2.684540	0.0771
Test critical values:	1% level	-3.437892	
	5% level	-2.864759	
	10% level	-2.568538	

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(WTI_BRENT)
Method: Least Squares
Date: 07/27/16 Time: 20:32
Sample (adjusted): 12/22/2010 3/13/2014
Included observations: 842 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
WTI_BRENT(-1)	-0.019251	0.007171	-2.684540	0.0074
D(WTI_BRENT(-1))	-0.249109	0.034499	-7.220698	0.0000
D(WTI_BRENT(-2))	-0.074858	0.035462	-2.110976	0.0351
D(WTI_BRENT(-3))	0.102568	0.034374	2.983862	0.0029
C	-0.287240	0.112699	-2.548723	0.0110

R-squared	0.087407	Mean dependent var	-0.006413
Adjusted R-squared	0.083045	S.D. dependent var	1.334338
S.E. of regression	1.277732	Akaike info criterion	3.333971
Sum squared resid	1366.486	Schwarz criterion	3.362093
Log likelihood	-1398.602	Hannan-Quinn criter.	3.344748
F-statistic	20.04160	Durbin-Watson stat	2.005815
Prob(F-statistic)	0.000000		

Sample before the structural break-PP
Our statistic gives a t-stat lower than 5% significance level
Null Hypothesis: WTI_BRENT has a unit root
Exogenous: Constant
Bandwidth: 5 (Newey-West automatic) using Bartlett kernel

		Adj. t-Stat	Prob.*
Phillips-Perron test statistic		-2.852661	0.0515
Test critical values:	1% level	-3.437865	
	5% level	-2.864746	
	10% level	-2.568531	

*MacKinnon (1996) one-sided p-values.

Residual variance (no correction)	1.751091
HAC corrected variance (Bartlett kernel)	1.213767

Phillips-Perron Test Equation
Dependent Variable: D(WTI_BRENT)
Method: Least Squares
Date: 07/27/16 Time: 20:33
Sample (adjusted): 12/17/2010 3/13/2014
Included observations: 845 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
WTI_BRENT(-1)	-0.024311	0.007317	-3.322510	0.0009
C	-0.357841	0.115039	-3.110609	0.0019

R-squared	0.012926	Mean dependent var	-0.006899
Adjusted R-squared	0.011755	S.D. dependent var	1.332713
S.E. of regression	1.324857	Akaike info criterion	3.402850
Sum squared resid	1479.672	Schwarz criterion	3.414067
Log likelihood	-1435.704	Hannan-Quinn criter.	3.407148
F-statistic	11.03908	Durbin-Watson stat	2.460110
Prob(F-statistic)	0.000931		

Table: Sample after the structural break-ADF
We reject the null hypothesis at 5%
Null Hypothesis: WTI_BRENT has a unit root
Exogenous: Constant
Lag Length: 4 (Automatic - based on AIC, maxlag=18)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-2.973346	0.0381
Test critical values:	1% level	-3.441434	
	5% level	-2.866322	
	10% level	-2.569376	

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(WTI_BRENT)
Method: Least Squares
Date: 07/27/16 Time: 20:38
Sample: 3/14/2014 5/31/2016
Included observations: 578

Variable	Coefficient	Std. Error	t-Statistic	Prob.
WTI_BRENT(-1)	-0.055262	0.018586	-2.973346	0.0031
D(WTI_BRENT(-1))	-0.469244	0.043346	-10.82556	0.0000
D(WTI_BRENT(-2))	-0.189417	0.046868	-4.041494	0.0001
D(WTI_BRENT(-3))	-0.134849	0.046396	-2.906509	0.0038
D(WTI_BRENT(-4))	-0.076477	0.041480	-1.843700	0.0657
C	-0.160867	0.081477	-1.974398	0.0488

R-squared	0.222980	Mean dependent var	0.015675
Adjusted R-squared	0.216188	S.D. dependent var	1.352215
S.E. of regression	1.197159	Akaike info criterion	3.208106
Sum squared resid	819.7848	Schwarz criterion	3.253361
Log likelihood	-921.1427	Hannan-Quinn criter.	3.225753
F-statistic	32.82912	Durbin-Watson stat	2.000248
Prob(F-statistic)	0.000000		

Table: Sample after the structural break-PP
We reject null hypothesis even at the 1%

Null Hypothesis: WTI_BRENT has a unit root
Exogenous: Constant
Lag length: 2 (Spectral OLS AR based on SIC, maxlag=18)

		Adj. t-Stat	Prob.*
Phillips-Perron test statistic		-4.051227	0.0013
Test critical values:	1% level	-3.441434	
	5% level	-2.866322	
	10% level	-2.569376	

*MacKinnon (1996) one-sided p-values.

Residual variance (no correction)	1.717096
HAC corrected variance (Spectral OLS autoregression)	0.588495

Phillips-Perron Test Equation
Dependent Variable: D(WTI_BRENT)
Method: Least Squares
Date: 07/27/16 Time: 20:40
Sample: 3/14/2014 5/31/2016
Included observations: 578

Variable	Coefficient	Std. Error	t-Statistic	Prob.
WTI_BRENT(-1)	-0.116095	0.019268	-6.025331	0.0000
C	-0.383674	0.085871	-4.468013	0.0000

R-squared	0.059292	Mean dependent var	0.015675
Adjusted R-squared	0.057659	S.D. dependent var	1.352215
S.E. of regression	1.312653	Akaike info criterion	3.385432
Sum squared resid	992.4816	Schwarz criterion	3.400517
Log likelihood	-976.3899	Hannan-Quinn criter.	3.391314
F-statistic	36.30461	Durbin-Watson stat	2.681957
Prob(F-statistic)	0.000000		

X. Summary Statistics

Column1	spread	ADS	BDI	BDTI	rigs	canada	pipe	lightoil
Mean	-4.50793277	-0.36478388	2718.27903	918.001982	714.865573	1469.27147	8001.02559	9344.41919
Median	-1.64	-0.19105496	1902	839	521.271056	1308.42172	4008	9532.94328
Maximum	6.65	0.94759468	11612	2317	1598.40681	2510	35019.1222	11970.2457
Minimum	-29.55	-4.04938653	291	457	145.04831	918.602803	-48.528433	7034
Std.Dev.	7.65879277	0.83340857	2445.32309	309.137119	485.570199	386.232508	9081.0524	1390.45347
Skewness	-1.18335615	-2.42657988	1.67449546	1.55018434	0.54451212	0.61447707	1.38138545	-0.02702449
Kurtosis	0.50913587	6.44170172	2.37116872	3.37273678	-1.32554883	-0.8621131	0.68324023	-1.42023134
Observations	595	595	595	595	595	595	595	595
ADF p-value (level)	0.2480	0.1519	0.4209	0.0899	0.5738	0.9572	0.9823	0.8516
ADF p-value (First Dif)	0.0000	0.0000	0.0000	0.0000	0.0017	0.0000	0.0000	0.0000

XI. Correlation

	spread	*ADS*	*BDI*	*BDTI*	*rigs*	*canada*	*pipe*	*lightoil*
spread	1							
ADS	-0.20864715	1						
BDI	0.4482205	-0.12734499	1					
BDTI	0.42182453	-0.18561042	0.66795838	1				
rigs	-0.7040709	0.28031314	-0.53251387	-0.45661773	1			
canada	-0.4137689	0.1963999	-0.62010856	-0.43127137	0.67475326	1		
pipe	-0.14591905	0.179278	-0.52534998	-0.33719012	0.43320463	0.87650498	1	
lightoil	0.56683953	-0.23360899	0.62942669	0.56519227	-0.76967864	-0.85272211	-0.75699997	1