universität wien

MAGISTERARBEIT

Titel der Magisterarbeit

"Decomposition of Oil Price Changes"

Verfasser

Arda Saygın Köstem

angestrebter akademischer Grad

Magister der Sozial- und Wirtschaftswissenschaften
(Mag. rer. soc. oec.)

Wien, im November 2009

Studienkennzahl lt Studienblatt:	A 066 913
Studienrichtung lt. Studienblatt:	Magisterstudium Volkswirtschaftslehre
Betreuer / Betreuerin:	Neil Foster, Ph.D.

Acknowledgements

I would like to thank my supervisor Neil Foster and Nadir Gürer for useful comments, Chris Kennedy for providing the NYMEX futures price series, Şenay Eriş, Erol Güçlü, and Berrin Barlan for their support throughout my studies and finally my dear family.

Teşekkür

Bu çalışma süresince gösterdikleri ilgi ve yardımdan dolayı danışmanım Neil Foster, Nadir Gürer, NYMEX future fiyat serisini sağladığı için Chris Kennedy, eğitimim süresince sağladıkları destek için Şenay Eriş, Erol Güçlü, Berrin Barlan ve son olarak sevgili aileme teşekkür ederim.

To my grandmother Öznur Korürek

Anneannem Öznur Korürek’e

CONTENTS

List of Abbreviations

AIC	Akaike Information Criteria
CRDW	Co-Integration Regression Durbin Watson
DOLS	Dynamic Ordinary Least Squares
ECM	Error Correction Model
FOB	Free on Board
OECD	Organization of Economic Cooperation and Development
OPEC	Organization of Petroleum Exporting Countries
NYMEX	New York Mercantile Exchange
VAR	Vector Autoregressive
WTI	Western Texas Intermediate

List of Figures

List of Tables

1 Introduction

The world has been increasingly relying on oil as a source of energy. Despite the development of alternative energy sources, the dominance of oil will clearly not be challenged in the near future (Figure 1). Considering its wide use in different areas, there is a strong link between crude oil prices and global GDP growth. Thus, the fluctuations in the crude oil price feed directly into the global economy. This strong influence on the global economy is what makes crude oil unique and worth investing time to investigate the factors driving its price. The correct understanding of the oil market is not only a necessity for assessing the implications of oil market developments but also to develop appropriate policies to ensure a stable growth rate of the world economy.

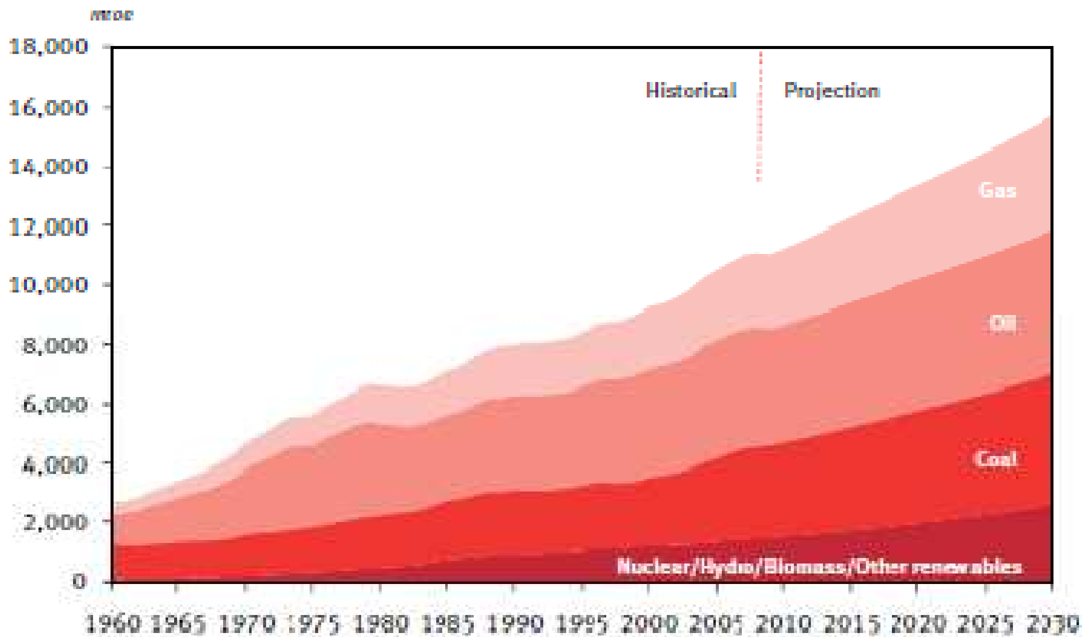

Figure 1. World Supply of Primary Energy by Fuel Type

There are a number of theories which characterize oil price formation (cf. Hotelling, 1931; Hubbert, 1956; Pindyck, 1999; Wirl, 2008 among others). These theories could be broadly categorized as long term price formation and short term price formation theories. In this thesis, short term dynamics of oil prices are investigated, therefore it's worth mentioning some mainstream ideas about them. In order to understand the background on which these theories are based, a brief history of the oil market is presented in the following paragraphs.

Oil was in use as early as the beginning of the 19th century. However it was not until World War II that oil was used as the main source for fuel. Rather in the 1860s it was used as “a disinfectant, a vermin killer, hair oil and boot grease”. Before World War II there was a plentiful supply of oil and domestically produced oil was satisfying the demand in producing countries. In fact, until the end of the 1940s, the United States and Latin America provided the oil used by Europe. The change in the American position from exporter to the largest importer in the world is a good indicator of how the oil market and the world economy evolved over time. As economies became more and more dependent on oil a stable supply became a necessity. However, the importance of having a safe supply of oil did not become clear until World War II.

In the pre-war period, the major oil companies were coordinating to keep the prices up which allowed them to accrue large profits. This is the period when the so-called Seven Sisters[1] emerged. These companies were “integrated”, which means that they were conducting every step of oil supply circuit from extraction of the crude oil, to its shipment, refinery processing and final delivery to customers. Before World War 2 broke out, Americans paid a significant amount to Saudi Arabia for an oil concession. The Saudis were convinced that they sold the Americans sand, as their closest ally, the British, did not believe that there was oil in the area. For five years they did not notice that they handed the Americans the largest oil deposits at the time. At the end of five years of disappointment Americans discovered oil. This discovery was perhaps the most rewarding of all times. At about the same time in March 1938, on the other side of the world, the Mexican government overtook the assets of the western oil companies in the country. This was the first example of the risks associated with investing in under-developed countries.

[1] Seven private oil companies which possessed monopoly power in the oil market: Exxon, Shell, BP, Mobil, Chevron, Texaco, Gulf Oil

With the outbreak of World War II it became clear that the oil supplied by Saudi Arabia would not be enough to satisfy global demand. The U.S. government gave every incentive for more companies to enter the Saudi market. Soon more American companies dug for oil in Saudi Arabia. To take advantage of the supply shortage and match the already ahead American oil industry, the British established a vast oil industry in Iran. This was the largest British asset in a foreign country at the time.

In 1951, the Iranians nationalized their oil industry, which was a big blow to Britain both politically and financially. Naturally, this step had its consequences. Prime-minister Mosaddeq was overthrown by a military coup on 19 August 1953. Although Mosaddeq paid a heavy price for his bold policy, he paved the way for national oil policies in oil producing countries. Nationalism kept rising in the Middle East, and the balance of oil power started to shift. OPEC was formed in 1960 through the leadership of Venezuela and Saudi Arabia. OPEC acted unilaterally to raise prices for the first time in 1963.

In 1970, Libya demanded higher petroleum prices, a greater share of revenues, and more weight and rights in the managerial decision process. Foreign petroleum companies agreed to a more than three-fold price increase in the early 1971. This deal made Libya the first producing country to force the West to accept their conditions. In December, the Libyan government suddenly nationalized the holdings of British Petroleum in Libya due to a foreign policy dispute. In 1973 the Libyan government announced the nationalization of a minimum of 50% of shares in all other petroleum companies operating in the country. However, total nationalization was not an option, given the need for foreign expertise and funds in oil exploration, production, and distribution. The absolute power of the American oil giants were over. The Middle East assumed the control. However, the major oil companies continued to drive oil prices from the end of World War II until the Arab Israeli War (the Yom Kippur War) in 1973. The Arab oil embargo caused a panic in the West, as oil prices sky rocketed. This was the first instance that OPEC fully exercised its power. It proved to be effective as prices quadrupled in a matter of months. Western companies turned elsewhere to find oil. Soon after the search, more oil was being produced from the North Sea than from many OPEC countries.

The Shah fell in the Iranian Revolution in 1979, and Western influence over Mid-East oil shrank even further. Two weeks after the fall of the Shah, the anti-western Ayatollah Khomeni came to power, who made it clear that Iran would not be a reliable source of oil anymore. Indeed, the turbulence and the revolution interrupted the Iranian supplies and the second oil shock hit the West in 1979.

This time there was panic but in reality there was not a permanent shortage of oil. Oil touched $40 a barrel, and OPEC felt that it could raise oil prices at will. It was the height of OPEC's market control. In the meantime the efforts of Western countries started to pay off as oil was discovered at the North Sea and in Alaska. By the early 1980s, a quarter of U.S. production was coming from the fields in Alaska. Higher prices resulted in increased exploration and production outside of OPEC. From 1980 to 1986 non-OPEC production increased by 10 million barrels per day. OPEC was faced with lower demand and higher supply from outside the organization. By 1983, four years after the Second Oil Shock, it was clear there were plenty of supply and no increase in demand, so OPEC had to cut prices. OPEC decided to introduce production quotas to limit production and keep prices up. In 1986, oil prices fell from $29 to $10 a barrel. Market forces now governed the market, as oil prices were set in the futures markets in New York (that started in 1983). OPEC was no longer in the driver's seat. From 1982 to 1985, OPEC attempted to set production quotas low enough to stabilize prices. These attempts met with repeated failure as various members of OPEC produced beyond their quotas. During most of this period Saudi Arabia assumed the lead role as the swing producer cutting its production in an attempt to slow the free fall in prices. By August of 1985, the Saudis were tired of this role because other OPEC members were not acting according to their quota commitments. They linked their oil price to the spot market for crude and by early 1986 increased production from 2 mbd to 5 mbd. Crude oil prices plummeted below $10 per barrel by mid-1986. Despite the fall in prices Saudi revenues remained about the same with higher volumes compensating for lower prices. In July 1986 OPEC reinstated quotas and since then OPEC continued to have mixed success in controlling prices. Despite the mistakes in the timing of quota changes as well as the usual problems in maintaining production discipline among its member countries, OPEC continues to play an important role in oil price formation.

Based on this brief history, the history of the oil market can be separated in to the Seven Sisters period and the OPEC period. Note that this division is based on the market structure and not prices; indeed there was not even a world price before World War II as Adelman notes "There was no world crude oil price before WWII…, which in 1948 mandated a competitive market price for crude oil sold out of the Persian Gulf to Europe" (Adelman, 2002).

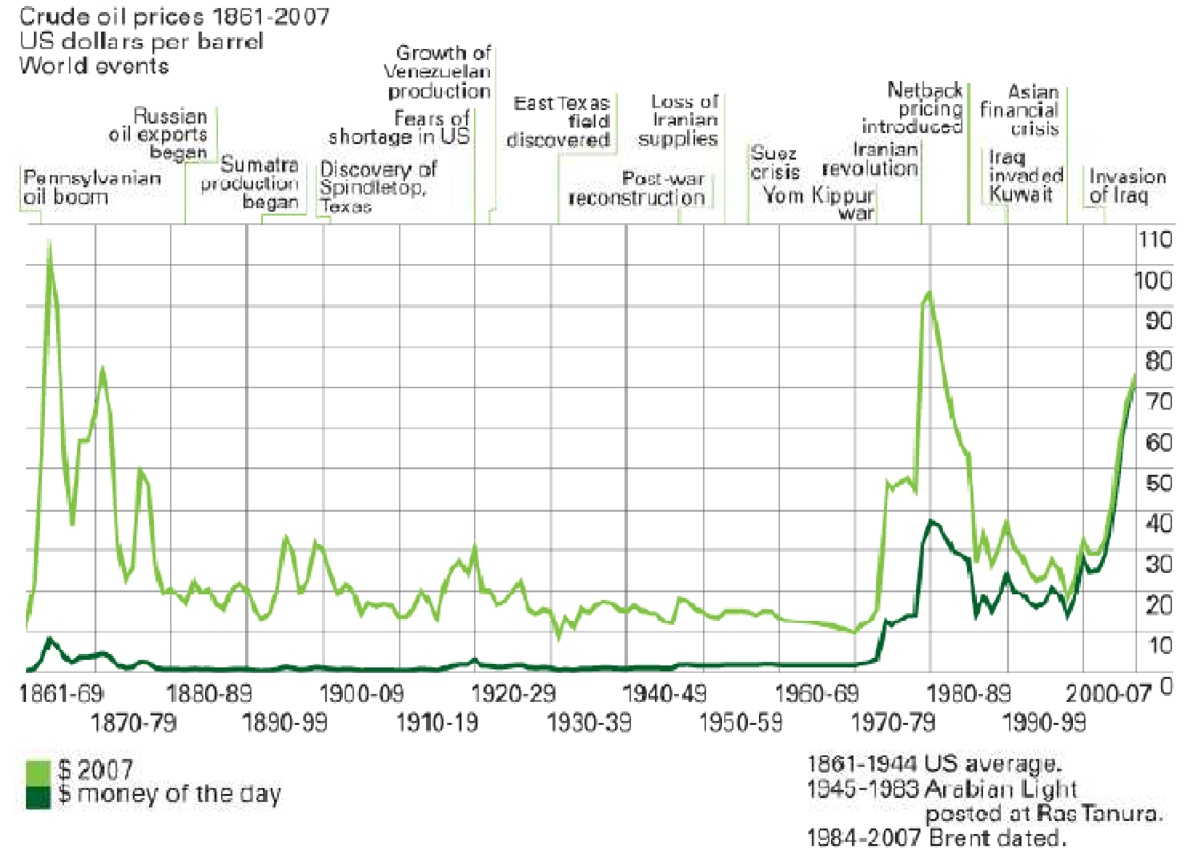

Figure 2. Crude Oil Prices (2008 US$)

Following World War II, the inflation-adjusted price of crude oil was stable and declining until 1970. As nationalization reached a level at which oil producing countries controlled most of the industry, OPEC took over the role of the Seven Sisters as the price maker. The rising political tensions and OPEC's embargoes sharply increased the prices in the 1970s and the oil market stayed quite unstable ever after.

Considering the change in the market structure from the private oligopoly of seven integrated oil giants (aka "Seven Sisters") to the formation of OPEC, oil price formation evolved over time. Perhaps it would be more accurate to say that the weights of different variables in the price discovery function altered due to the changing roles of the market players. Other factors, such as the introduction and ever increasing role of financial markets on oil prices are also to blame for the unstable oil prices. How much role each variable plays, has always been a subject of debate, however the fact that not all times of crises were followed by oil price rises indicate that market forces dominate the price formation process after all. A simple investigation of Figure 2 provides the evidence that price rises cannot always be linked to political events. As long as these events do not disturb the supply and demand balance in anyway, prices remain at the level required by the market equilibrium. Just as with political events, any factor has to have some direct or indirect effect on the demand and supply balance. With this in mind, the following econometric analysis will be designed around variables which are believed to reflect the demand and supply balance of the market. However, it must be stressed that human behavior does influence investor decisions which are shaped by the expectations of the market participants. Note that these expectations are also expectations about the demand and supply balance.

2 Literature Review

Existing studies of oil prices are commonly based on econometric models which use variables commonly cited as proxies for market conditions. One can group these variables as reflecting the demand and supply conditions as well as expectations of the investors. Variables reflecting all three market conditions are used in this thesis. Some of the recent studies which have used the same variables as the ones in this thesis and their results are discussed in the following paragraphs. Then follows a discussion of the mainstream econometric methods used in oil price studies.

There is a consensus on the counter-cyclical relation between inventories and price levels (cf. Kaufmann 1995, Balabanoff 1995, Dees et. al. 2004) however the effect and magnitude of the other variables are disputed in academic circles. Perhaps the most disputed of all is the role of OPEC. While there is an agreement that OPEC does have some influence on the oil prices, the magnitudes suggested are hardly similar. Despite disagreements about the power of OPEC, there seems to be an agreement about a decline in OPEC's power in recent years. Merino and Ortiz (2005) find that OPEC marginally influences the price due to the lack of spare capacity to cope with a significant disruption of supplies. This situation has not changed since then as Roubini (2008) and Kesicki (2009) note that the absence of significant spare capacity inside OPEC constrains the ability of the cartel to control prices. These discussions keep going on while one thing is clear: OPEC is found to be a factor contributing to oil prices to some degree.

One of the recent studies which dispute the validity of the previous findings is the work by Kaufmann et al. (2008). They show that the relationship between refinery utilization, which is commonly cited as one of the contributors to the price rise of crude oil, and oil prices is negative such that higher refinery utilization rates actually reduce the price of oil. A similar result is obtained in this thesis and an accompanying discussion is presented in the relevant section.

There are two main approaches to modeling oil prices in the literature; multivariate and univariate. The former are mostly based on Vector Autoregressive (VAR) models and the latter are mostly based on Dynamic Ordinary Least Squares (DOLS). Although different results are obtained from the two approaches from time to time, one can often attribute these varying results to differences in sample periods and frequencies.

Kaufmann et al. (2004), Merino and Ortiz (2005), Yanagisawa (2008), Stevans and Sessions (2008), Asali (2008) and Geman and Ohana (2009) are among the recent studies which have used VARs to analyze oil prices. All five studies confirm a one directional causality from inventories to prices while the first two find a one directional causality from OPEC capacity utilization and refinery utilization to prices as well. Merino and Ortiz find a two way causality between spread (as described in this thesis) and the prices which is contrary to the result obtained in the later sections of this thesis. This conflict could possibly be due to the different sample periods, January 1992 to June 2004 in Merino and Ortiz and March 1995 to February 2009 in this thesis. They also show that oil prices increased hand in hand with the rise of the non-commercial long positions on crude oil and note that backwardation has informative content for the premium[2] over the entire sample period.

Geman and Ohana (2009) have especially complementary information for this thesis. As the sample periods overlap for the most part, their findings are valid during the sample period of this thesis. They find that three quantities are of crucial importance for all commodities, namely the spread of the forward curve (spread), the available inventory and the spot price volatility. Among these, they find the highest correlation between spread and the inventory levels. Figures 9 and 10 of this thesis show this tendency.

The most notable univariate models in recent years are Möbert (2007), Kaufmann et al. (2008), and Chevillon and Rifflart (2009). The latter two uses DOLS while the former uses OLS on the first difference of the variables. Using OLS on the first difference of the series could perhaps give accurate results for short term dynamics; however it is not a reliable method for long sample periods such as the one considered here. Kaufmann et al. (2008) find that most of the price increase in the 2004-2006 period can be attributed to concerns about future oil market conditions, which experienced a switch from backwardation to contango market. A similar result is found in Möbert

[2] Defined as the deviation of the oil price over the price required by the inventory levels.

(2007), which uses three different variables to account for the futures market conditions (i.e. expectations). He finds all three to be statistically significant, spread having the largest magnitude. On the other hand, Kaufmann et al. (2004) and Möbert (2007) disagree about the role of OPEC. The former states that OPEC has a major influence on the prices whereas the latter argues that the effect is only a modest one. The two studies report conflicting results regarding the effect of refinery utilization as well. Kaufmann et al. (2008) report that the refinery utilization is negatively related to prices while Möbert (2007) finds the opposite. The latter uses different levels of refinery utilization above 95% as a dummy variable, all of which carry coefficients with positive signs. Möbert defines refinery utilization as the global demand relative to refining capacity in mbd. This definition could be misleading because it ignores the fact that some of the global demand is satisfied by inventories. The resulting refinery utilization rates are frequently near and sometimes at full capacity which can imply a false positive relation between the prices and the refinery utilization rates. Using the US refinery utilization rate I find similar results concerning the relationship between the refinery utilization rates and prices.

Considering the theoretical expectations and the results obtained from different econometric methods it can be said that there is a general agreement on the fundamental factors effecting oil prices, however the relative weights of each factor changes due to the changes in the structure of the market. Askari and Kirchene (2008) explain this phenomenon in their study by concluding that oil price dynamics experienced breaks, or jumps in recent years. The results from different samples in this thesis reflect such changes in the market structure.

3 Econometric Analysis

In attempting to decompose the effects of different factors on the fluctuation of the price of crude oil, time series tools are used. This section will explain these tools, the reason for their choice and finally present the results of the econometric model. Before engaging in econometric modeling I shall first introduce the sample and the variables chosen for the model in sections 2.1 and 2.2 respectively. In section 2.3, the model and the results will be presented. 2.4 will conclude this section with robustness analysis and a discussion of causality.

3.1 Sample Period

Monthly data from March 1995 until February 2009 are used to estimate the price equation. The sample period is motivated by a number of factors. In its broadest sense, the oil market can be separated into two phases. The first phase is the non-competitive non-volatile crude oil market which ended with the cartelization among OPEC states. The second phase has been more market oriented and characterized by more volatility and continues today. The sample period considered in this thesis falls into the second phase. The reason for this as Bacon explains is that there is no point in modeling the first phase; "Before 1973 prices were largely administered and were roughly constant in nominal terms so that there has been little interest in modeling the behavior of the world oil prices before this date." (Bacon, 1991)

The current market structure is shaped following the price collapse of 1986, when OPEC reinstituted quotas after changing its announced strategy from defending an official price into defending market share. In 1995 crude oil options contracts were introduced in the New York Mercantile Exchange (NYMEX) which have been used in ever increasing amounts.

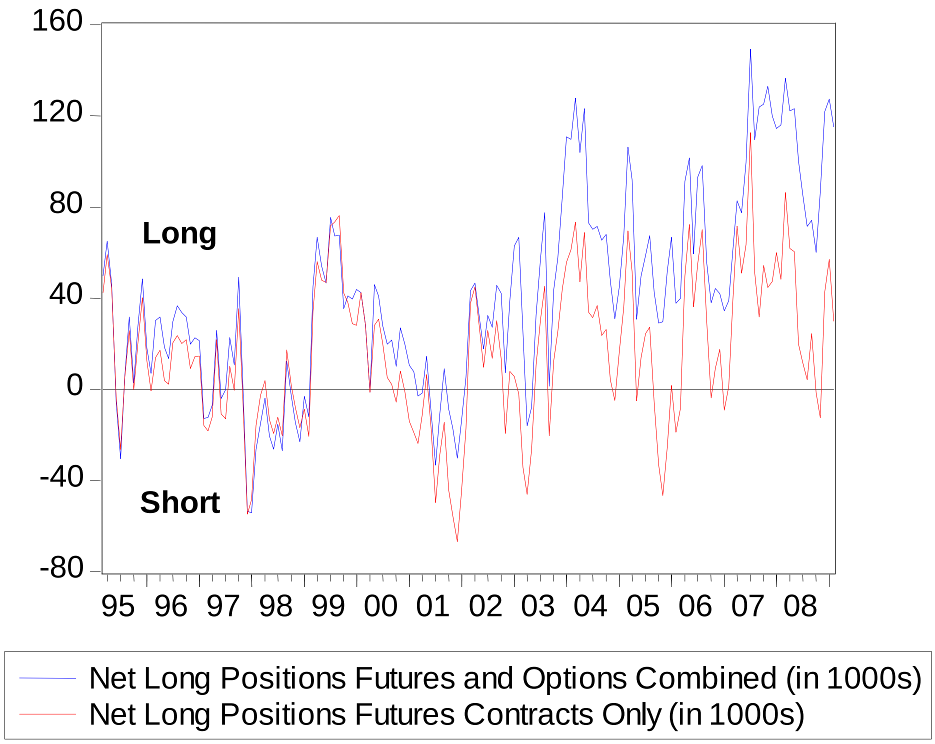

Figure 3. Net Long Positions at the NYMEX

Figure 1 shows that the introduction of options contracts has increased the volume of net long positions in the NYMEX. Furthermore, this observed boom in the net long positions is proved to influence the prices as presented by the report submitted to the U.S. Senate Committee on Energy and Natural Resources (Verleger, December 11, 2007). The report concludes that so-called **delta hedging** magnifies the price fluctuations. Econometric evidence can be found in the same report. In light of this report, the start of the sample period is coincides with the introduction of options contacts in the NYMEX.

3.2 Variables

Relatively few studies focus on prices as the variable of interest compared to the more common use of prices as explanatory variable (cf. Kilian, 2007; Clarida et al., 2000; Hamilton and Herrera, 2004). Bacon (1991) gives a survey of factors which are central to price formation. He notes the particular importance of incorporating OPEC behavior in any supply/demand framework. Kaufmann et al. (2004) introduce the variables "OPEC *capacity utilization*[3]", "OPEC *quota*[4]" and "OPEC *cheat*[5]" to accommodate the influence of OPEC on the prices. Kaufmann et al. modify OPEC capacity utilization in their 2008 paper as discussed in the next section. This latter variable will be incorporated in the econometric model used in this thesis unless otherwise stated.

OPEC and other variables were used in different econometric specifications and different samples. Sub-samples were constructed due to structural breaks confirmed by Chow break point tests and in some instances sub-samples were intuitively constructed. Discussion of the time series properties of each variable and their expected effects follow.

3.2.1 Price Series

Two different sets of price series (Figure 4)[6] are used as the dependent variable; the real average free-on-board[7] (FOB) price of US crude oil imports and the price of West Texas Intermediate (WTI), the benchmark oil type for North America.
Price is the real average (FOB) price of US crude oil imports in 2009 US$ and WTI is the real price of Western Texas Intermediate in 2009 US$. The US consumer price index is used to adjust the nominal price series.

The first price series is motivated by the fact that average FOB prices of imported crude oil represent the price paid for physical barrels obtained from a variety of sources. Therefore this basket is unaffected by the conditions unique to a single market. For instance, the other price series West Texas Intermediate (WTI) experienced disturbances due to local circumstances such as the fire at McKee Refinery near the West Texas fields, hurricanes Katrina and Ivan. In addition to this, the average FOB price of US crude oil imports reflects the composition of the crude oil types desired. The composition of imports is shifting towards light sweet crude oil from heavy sour, which especially picked

---

[3] OPEC capacity utilization is the percentage use of OPEC capacity, which is calculated by dividing OPEC production (mbd) by OPEC capacity (mbd).
[4] OPEC quota is the production limit (per day) set by OPEC.
[5] OPEC cheat is the difference between OPEC production and OPEC quotas (million barrels per day).
[6] Note that during hurricane Katrina and Ivan in 2004 and 2005 the price series moved in opposite directions. Also the impact of the fire at McKee refinery which shut this facility down from February until July 2007 could be seen by the upward push in the WTI series at this period.
[7] The price charged at the producing country's port of loading.

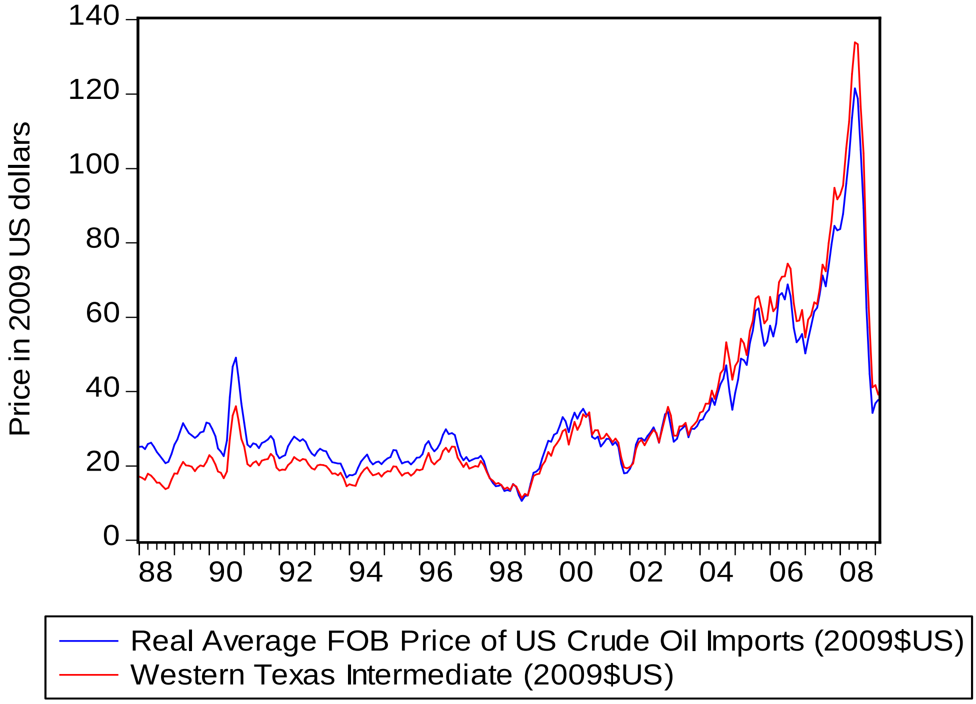

Figure 4. Import Basket Price vs. WTI

the implementation of the recent environmental regulations. The ability of the price series to reflect this shift will be crucial to the analysis of the results of the econometric model.

I begin to review the time series properties of the series by looking at the correlogram of the series. The autocorrelation function (ACF) decays very slowly which is typical of non-stationary series. Although visual review of the correlogram gives useful hints, the ACF fails to distinguish between unit root and near unit root processes (Enders, 2004). Therefore I move on to make formal test of unit root, namely the Augmented Dickey Fuller Test.

The ADF test statistic is significant at the 10% level, while the null of the price series having a unit root cannot be rejected at the 5% and 1% level. I conclude that the price series is a unit root process. Following usual convention to find the order of integration, I take the first difference of the series. The ADF test statistic of the first difference of the series exceeds the test critical value at the one percent level implying that the first difference of the series is stationary. Therefore I conclude that this price series is integrated of order one.

Table 1. ADF Test Statistic of Price Series

Price[8]	-2.600171
ΔPrice[9]	-6.411403***

Lag length is determined by the Schwarz Information Criterion
1%, 5%, 10% significance levels are denoted by ***, **, *

The other dependent variable, the WTI price series is used to evaluate the robustness of the results obtained from the import price basket. The conclusions derived from the former series should also apply to this series and there should not be any conflicting results. Furthermore the remark by Bacon justifies the validity of the use of a benchmark crude price as a proxy for the entire crude oil market; "Over the longer term the prices of all crudes tend to move sufficiently closely together to make the use of a marker or reference crude an attractive simplification for attempts to understand the workings of the world oil market" (Bacon, 1991). Undoubtedly the results will be applicable to the entire oil market.

The WTI series naturally follows the real average FOB price of US imports with a premium changing from time to time. Following the steps above I obtained the results presented in table 2. As is the case with the real average FOB price of US imports, I

Table 2. ADF Test Statistic of WTI Series

WTI	- 2.140102
ΔWTI	-7.874281***

Lag length is determined by the Schwarz Information Criterion
1%, 5%, 10% significance levels are denoted by ***, **, *

cannot reject the null hypothesis of unit root at conventional levels of significance. Since the first difference of the WTI series has a significant ADF test statistic, I conclude that the WTI series is also integrated of order one.

[8] The ADF test statistic of the levels of the variables include a constant and/or a trend when appropriate throughout the thesis.
[9] The ADF test statistics of the first difference of the series are calculated without a constant and a trend throughout the thesis.

3.2.2 Stocks

Stocks is the days of forward consumption of OECD oil stocks, which is calculated by dividing OECD stocks of commercial crude oil stocks by OECD demand for crude oil. Note that OECD commercial crude oil stocks include crude products but excludes government controlled inventories, such as the US strategic petroleum reserve (SPR). SPR management, or any other government controlled inventory for that matter, does not normally take market conditions into account and therefore does not reflect market conditions. It might be argued that fill policies of governments could influence market prices, however the amount of crude that can be withdrawn from the market or released from SPR is small relative to the market size, and it is reasonable to assume that such policy effects are negligible. There still is a debate about the true impact of these policies, however deeper discussion of this issue is beyond the scope of this thesis. The reader can refer to Verleger's report dated December 11, 2007 submitted to the US committee on energy and natural resources for a thorough discussion of the effects of SPR management on oil prices.

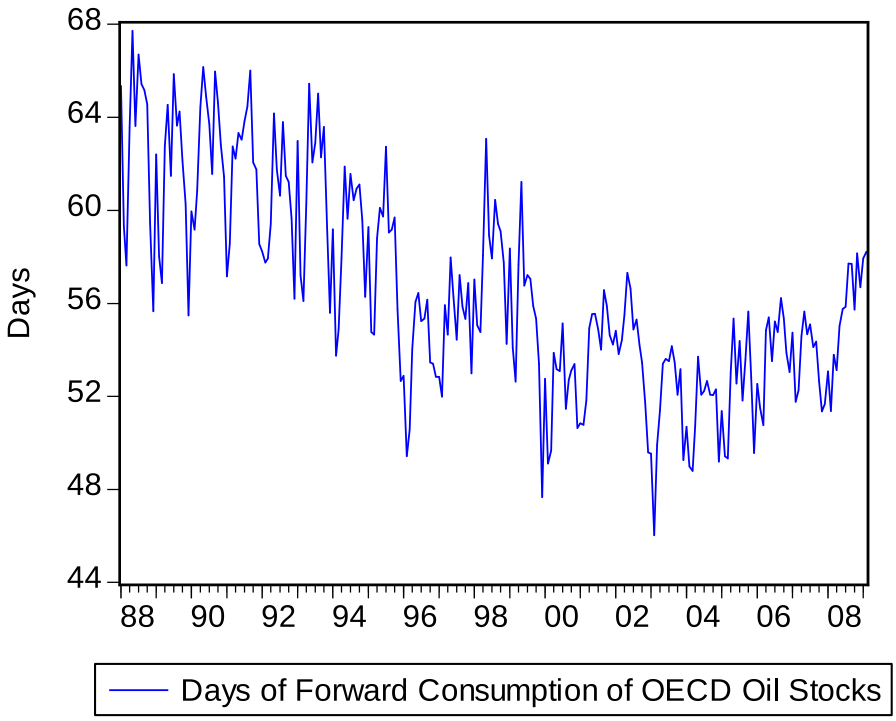

Figure 5. OECD Inventory Levels

The time series behavior of the series begins with a review of the series' correlogram. The ACF decays slowly and follows a wavy pattern, which is typical of non-stationary series. Carrying out the formal ADF test for unit root, the following results are obtained (Table 3). Since the ADF test statistic of the level of the variable does not

Table 3. ADF Test Statistic of Stocks Series

Stocks	-2.571302
ΔStocks	-3.381086***

Lag length is determined by the Schwarz Information Criterion
1%, 5%, 10% significance levels are denoted by ***, **, *

exceed the test critical values at any of the conventional levels of significance I conclude that the stocks series is a unit root process. First differencing of the series turns out to be enough to reach a stationary series, therefore the stocks series is integrated of order 1. Historically higher inventory levels are associated with lower prices. Figure 6 shows the counter-cyclical relation of prices and inventories. The only exceptions to the counter-cyclical behavior are observed during the gulf war and the 2002-2006 periods. The first one is the result of the rapid precautionary inventory built up preceding the war which can be interpreted as a demand shock. This temporary disruption did not last long and the natural counter-cyclical relation between inventory and price levels returned immediately after the war.

The more recent cyclical relation is harder to explain. One has to take into account the role of financial markets in order to correctly assess this phenomenon. One can clearly observe the increasing activity in the NYMEX from figure 7 below. Although there has been an increase in the number of contracts ever since the introduction of futures (and later options) contracts, this increase picked up especially after 2002. This corresponds to the hand in hand increase in prices and inventories.

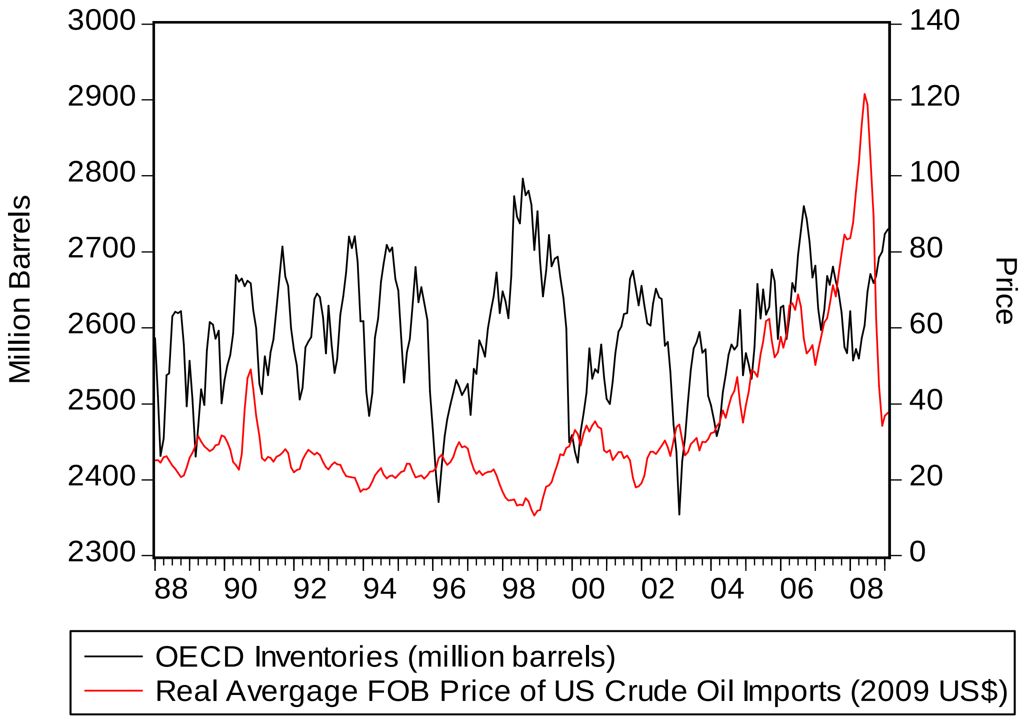

Figure 6. OECD Inventory Levels vs. Real Average FOB Price of US Crude Oil Imports

Closer inspection of the 2002-2006 period reveals the shift of investors towards long positions (i.e. betting on higher prices) which creates incentives for investors to build up stocks. When this inventory build-up takes place simultaneously by a large number of investors, demand is obviously boosted resulting in a price hike. In order to check for the effects of financial markets on the price of crude oil, a sub-sample with the variable spread will be modeled. Next we look at the variable spread, the motivation behind its choice, its relation with the investor behavior and its time series properties.

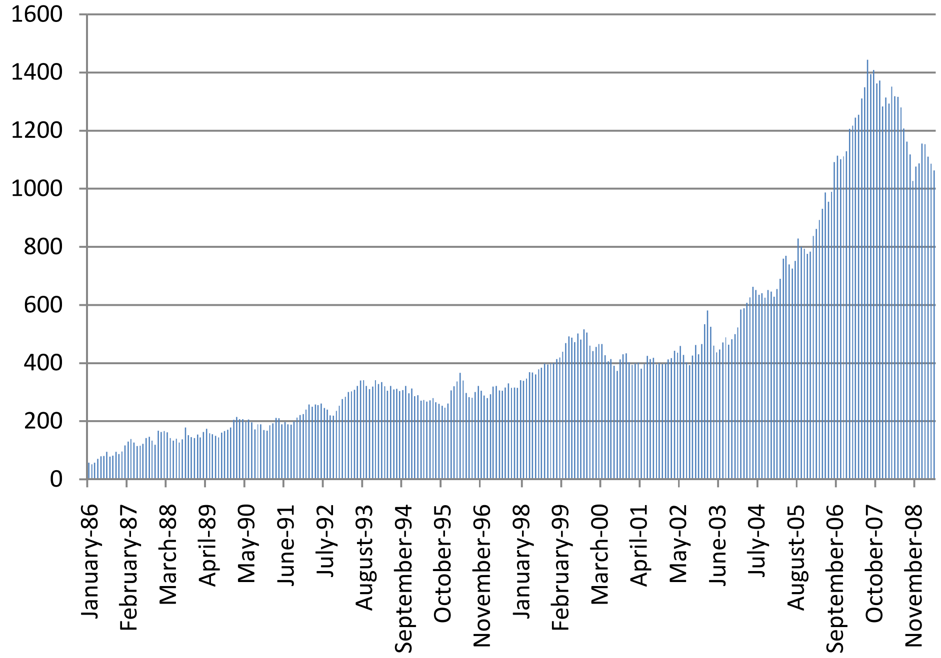

Figure 7. Number of Total Reportable Long Positions at the NYMEX (in thousands)

3.2.3 Spread

Speculators are frequently blamed as the reason behind large price shifts in the oil market. It is still a highly debated topic with respected economists defending opposing views. Considering the significant rise in the open interest (i.e. number of contracts traded at the NYMEX) of the crude oil contracts after 2002, I shall focus on this period. In his May 12, 2008 New York Times article titled “The Oil Non-Bubble” Paul Krugman argued that there is no evidence of a speculative oil bubble because if speculators are driving up the price of oil above the price “justified by fundamentals”, then a market adjustment would occur where “drivers would cut back on their driving; homeowners would turn down their thermostats; owners of marginal oil wells would put them back into production…[and the resulting] excess supply would….drive prices back down”. On the other hand Prof. Eckaus of the MIT Center for Energy and Environmental Policy Research points out that hedge funds along with other speculators are very active in the oil market and their activity has raised the volume of oil transactions far above the volume warranted by ordinary commercial transactions. He argues that there is no reason based on current (referring to summer 2008) and expected supply and demand that justifies the current price of oil. Finally, he concludes that the oil price is a speculative bubble.

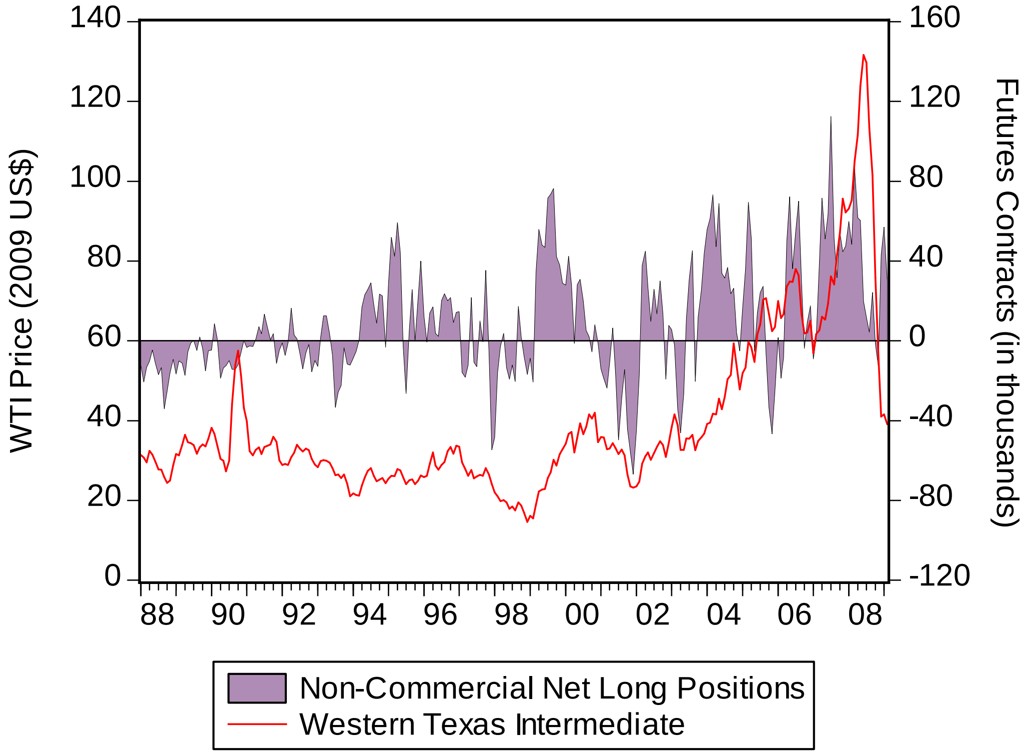

Figure 8. Futures Contracts vs. WTI Price

In order to look for evidence in support of either view I plot (Figure 8) the "best" measure available to know whether speculators are betting on rising or falling prices, namely the net long positions of non-commercial traders in light sweet crude oil at the NYMEX, and the price level of the corresponding crude oil, namely the Western Texas Intermediate (WTI).

The dominance of the long positions over the short positions is obvious throughout the 2002-2008 period. The behavior of the inventories follows roughly that of the positions at the NYMEX until 2008 (Figure 9). As the financial crises hits the global economy the inventories keep increasing due to the negative demand shock while prices collapse. Figures 8 and 9 suggest that the expectations of the investors shape their decisions about inventories. Namely an expectation of higher prices causes hoarding in anticipation of higher future profits. In order to account for expectations which are the source of inventory management practices I use the variable spread.

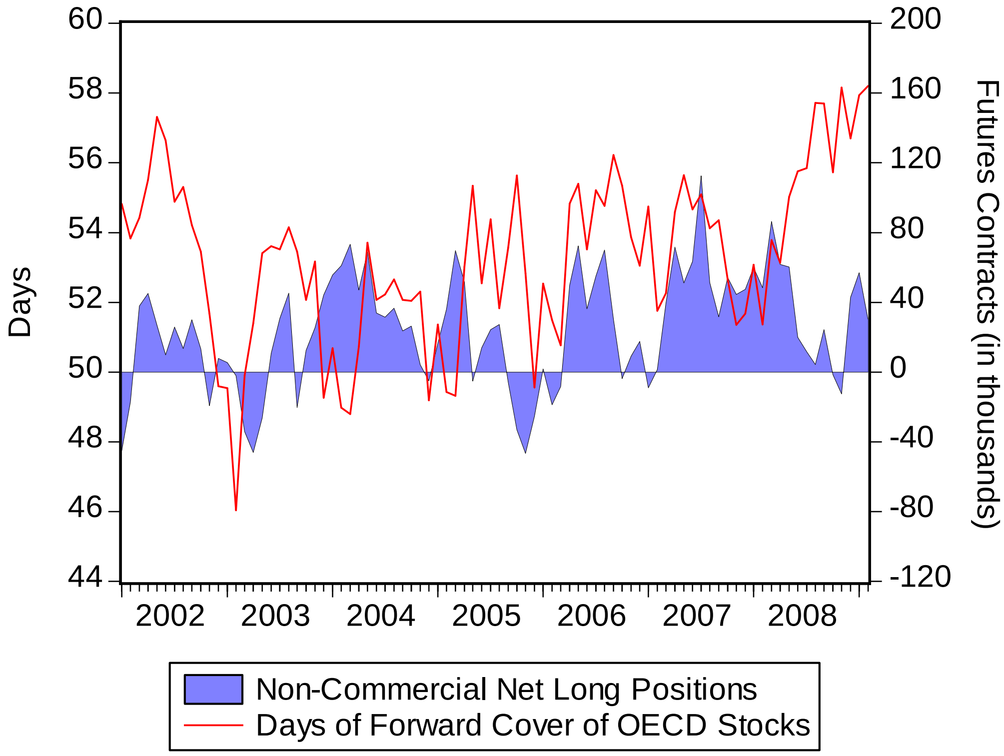

Figure 9. Inventories vs. Futures Contracts

Spread is the difference between WTI futures contracts with 4 months and 1 month to maturity. This variable reflects the expectations of investors. Inventory management policies are based on these expectations.

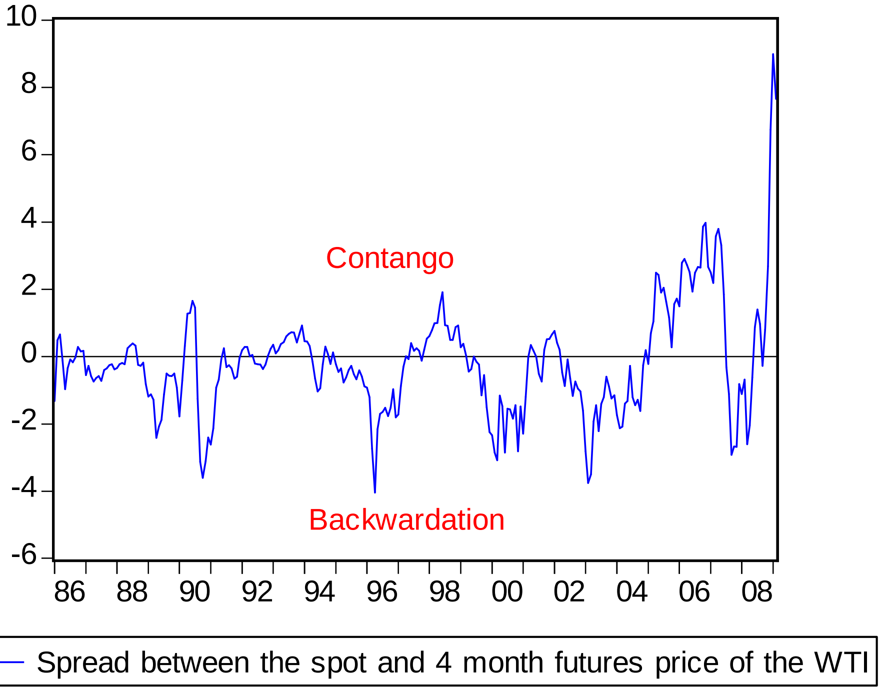

Figure 10. Spot-Futures Spread

Following the usual order, the inspection of the correlogram hints that the series is unlikely to be non-stationary with a fairly rapid decay in the ACF. Formal tests for unit roots confirm this observation (Table 4). Since the ADF test statistic exceeds the critical

Table 4. ADF Test Statistic of Spread Series

Spread	-2.491752**

Lag length is determined by the Schwarz Information Criterion

1%, 5%, 10% significance levels are denoted by ***, **, *

value at the 5% level, the null hypothesis of a unit root is clearly rejected and I conclude that spread is a stationary process for the sample period between March 1995 and February 2009. This characteristic of the series makes it unsuitable for the DOLS method used in the econometric analysis later in this paper. Nevertheless, sub-samples which focus on more recent developments in the oil market may include spread as an explanatory variable in the DOLS estimation since the spread series behave more like a non-stationary process in the recent years. The ADF test result for the sub-sample January 2002 – February 2009 is presented below;

Table 5. ADF Test Statistic of Spread Series (Jan. 03 – Feb. 2009)

Spread	-1.931434
ΔSpread	-6.193981***

Lag length is determined by the Schwarz Information Criterion

1%, 5%, 10% significance levels are denoted by ***, **, *

I cannot reject the null hypothesis of a unit root at conventional levels of significance for the restricted sample. This is not surprising since prices are stable until 2000 and keep rising thereafter. In particular the 2002-2009 period sees a 500% rise in prices. Expectations are altered and the futures market experiences the longest contango[10] period in its history (Figure 10). It can be argued that due to increased activity in the futures market and the increasing trend in prices in the 2000s the spread variable is more relevant than in any other time in the history of crude oil after the introduction of the futures market. The first difference of the series is stationary since the ADF test statistic is significant at the 1% level implying that the spread variable is integrated of order one during the 2002-2009 period. I expect a positive influence of the spread on the prices in this period.

3.2.4 OPEC Capacity Utilization

Caputil is capacity utilization by OPEC. This variable is built using the same convention as in Kaufmann et al. (2008). It is calculated by dividing OPEC production by OPEC capacity, multiplying this quotient by OPEC's share of global oil production, and dividing this product by the rate at which OPEC cheats on its quota[11]. As explained in Kaufmann et al. (2008) OPEC capacity utilization is designed in this way in order to account for the two different effects of OPEC cheat on the market; cheating could increase capacity utilization by OPEC but reduce oil prices by increasing supply relative

[10] The situation in which the futures prices exceed the spot prices. The opposite case is called "backwardation".
[11] OPEC cheat is calculated by dividing the difference between OPEC production and OPEC quota by world oil demand

to the quota that balances world oil demand and non-opec production. Dividing the OPEC capacity utilization by OPEC cheat allows increased rates of cheating to reduce oil prices even as cheating causes capacity utilization to rise.

The cubic specification of caputil is also borrowed from Kaufmann et al. (2008), which allows for two turning points (inflection points). Since the sample period starts with the introduction of options contracts and extends until recent months, and the data frequency is monthly as opposed to quarterly data used by Kaufmann et al. (2008) it is of interest whether this cubic specification will yield similar results. In that study the results

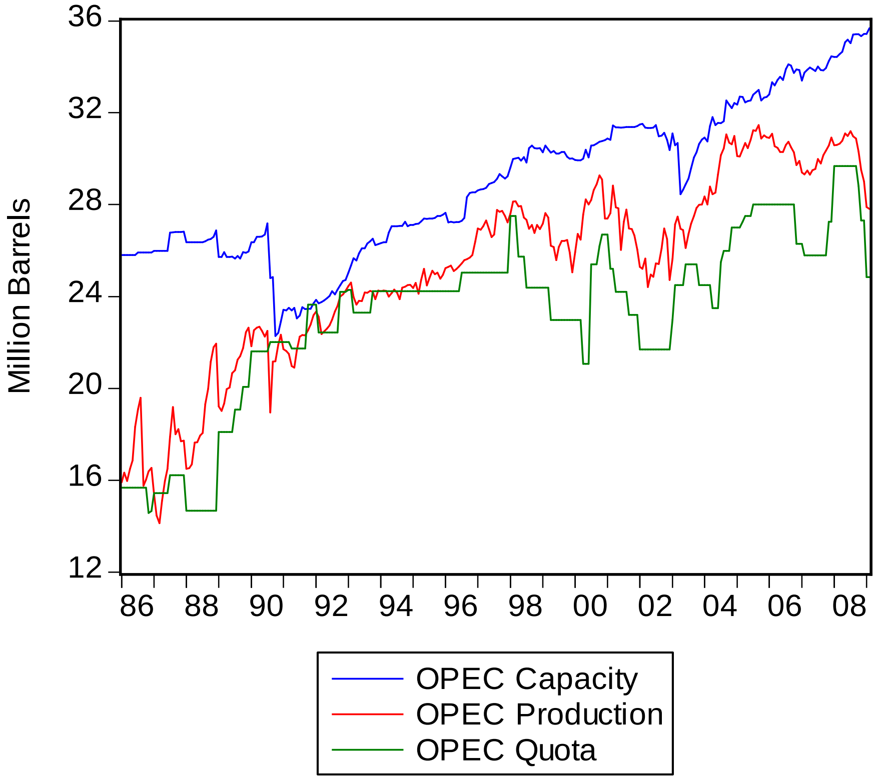

Figure 11. Individual OPEC Variables

confirmed the expectations that the coefficient of the squared variable would be negative and the other two positive.

These results suggested that prices increase exponentially up to the first turning point and increase exponentially after the second turning point. Between these turning points there is a normal operating range in which changes in capacity utilization have a small impact on prices (Kaufmann et al., 2008). The market forces behind this relation between capacity utilization and price is explained thoroughly in Kaufmann et al. (2008) and the following paragraph is quoted from the mentioned work:

"This non-linear relationship is based on the assumption that producers prefer to lift oil within a normal operating range. At levels well above this range, high utilization can interfere with field maintenance that is needed to ensure a well's long-run productivity. Similarly, operators are reluctant to pump at very low utilization rates because the fixed costs of production are very much greater than operating costs – so long as prices remain above operating costs owners prefer to operate their wells at capacity to pay their fixed costs. Preferences to operate within this range coupled with inelastic price elasticity of demand imply that oil prices must change significantly to move OPEC capacity utilization rates back towards the normal operating range. As capacity utilization rises beyond normal operating conditions and supplies become tight, inelastic demand implies that large price increases are needed to reduce demand and thereby bring utilization rates back to the normal range. On the downside, inelastic price elasticity of demand imply very large price reductions are needed to increase consumption (or make it economical to decommission capacity) to move capacity utilization back towards the normal operating range."

Turning to the time series properties of the caputil series, the correlogram does not give any initial clue about the stationarity of the series. I turn to the ADF test to identify the series. The test results are presented below;

Table 6. ADF Test Statistic of Capacity Utilization

Caputil	-1.644416
ΔCaputil	-6.836433***

Lag length is determined by the Schwarz Information Criterion
1%, 5%, 10% significance levels are denoted by ***, **, *

The ADF test statistic is not significant at any of the conventional significance levels, therefore I conclude that caputil is a unit root process. The first difference of the series turns out to be stationary, therefore I identify this variable as integrated of order one.

Besides measuring the influence of OPEC on the oil prices, caputil can also be interpreted as a proxy for demand. After the announcement of a quota mechanism in the mid 80's OPEC's official strategy has been defending the market share instead of directly setting the price. In order to monitor the prices OPEC instead announced a target price corridor. This strategy meant that OPEC set the production to match the expected difference between demand and non-OPEC supply, which is determined largely by non-OPEC capacity. As price takers, non-OPEC producers generally operate at or near capacity. This means that any increase in OPEC utilization has to be the result of a demand increase. Contrarily OPEC can increase production to suppress prices as well; however this ability has been significantly diminished in recent years. OPEC introduced a target price corridor between 22 and 28 dollars in March 2000. Any price increase beyond $28 was to be prevented by increased OPEC production. However, after prices continued to rise above $28 for more than a year OPEC suspended it corridor in June 2005. The inability of OPEC to keep prices in the announced corridor casts serious doubts about the cartels ability to control prices. Since this ability depends solely on OPEC spare capacity which enables the cartel to produce beyond the level that balances supply and demand, the available spare capacity reveals how strongly OPEC can alter the supply-demand. In any case, the increased utilization of OPEC signals "tightness" in the market and a positive effect of caputil is expected on the prices. However, I do not rule out the possibility of a negative effects before the year 2000 for the reason mentioned above, namely the increased utilization also corresponds to an increase in supply.

3.2.5 Refinery Utilization

Refutil is the US refinery utilization rate. The refining process plays a crucial role in the oil market. Consumer are not interested in the crude oil, instead the refined products such a gasoline, diesel, heating oil and kerosene. Refinery bottlenecks put tremendous upward pressure first on the refined product prices and then on crude oil prices. Often these pressures cause jumps in the crude oil price. With the emergence of

Asian giants such as China and India the world's refineries are troubled to meet demand which has increased faster than the refiners can expand their capacities. The refining sector is capital intensive in nature and it takes years before the actual refining takes place. As a result the situation which developed in recent years can be expected to last at least in to the near future. Intuitively, I expect that refutil would increase prices. However, as mentioned in Kaufmann et al.(2008), there is an alternative approach which

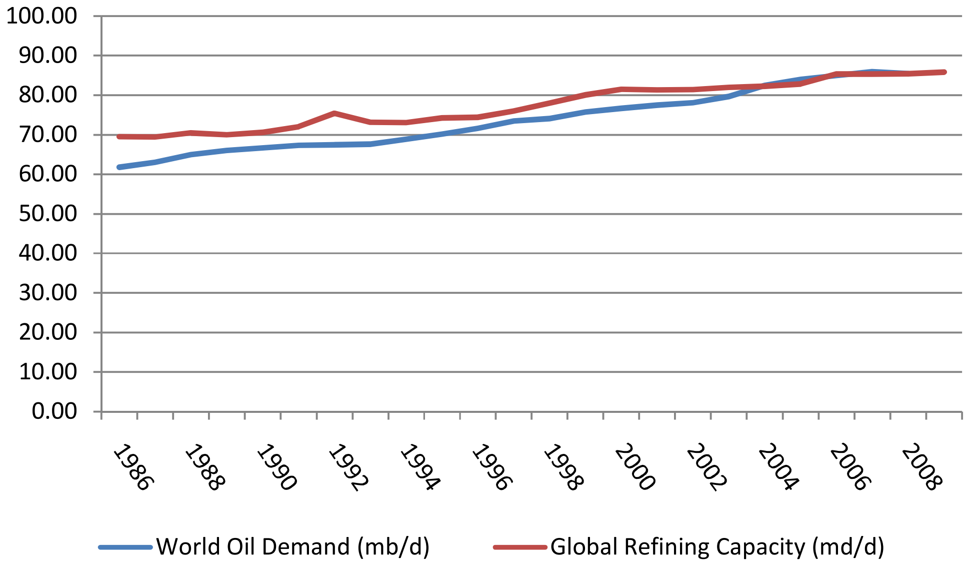

Figure 12. Demand vs. Refining Capacity

one should consider in evaluating the results of the econometric analysis. Their argument is based on the notion that the effects of refinery utilization rates on the prices of crude oil are associated with changes in the quality of crude oils produced and the ability of refineries to convert these crudes into refined petroleum products. The quality of a crude oil based largely on its density and sulfur content. The density of crude oil is measured by the API gravity index, which measures the density of crude oil relative to water. An API value greater than 10 indicates that crude floats on water, with larger values indicating a reduction in density (i.e. lighter crude). In general, light grades of crude oil are of a higher quality because they generate more valuable products than heavier grades. Crudes with a high sulfur content (so-called sour crudes) are of a lower quality because they increase refinery maintenance costs due to enhanced corrosion associated with the sulfur. Due to these differences the price of light sweet crudes are higher than sour heavy ones.

Producers around the world provide different quality crudes to the market and refiners use different kinds of crudes based on the product that they produce. Light sweet crudes are desired by both the producers and the refiners since they increase the revenue for the producers and decrease the maintenance costs for the refiners. Of course, the more higher quality products (i.e. gasoline, diesel) the refiners produce, the greater the revenues they generate as well. This common interest of both sides makes light sweet the primary choice for both sides. Because much of the world's refining capacity is set up to use light sweet crudes, increasing refinery utilization rates generally stimulate the production of heavy and sour crudes (Kaufmann et al., 2008). For example, Saudi Arabia increased its production of crude oil from 7.52 million barrels per day (mbd) in 1999 to

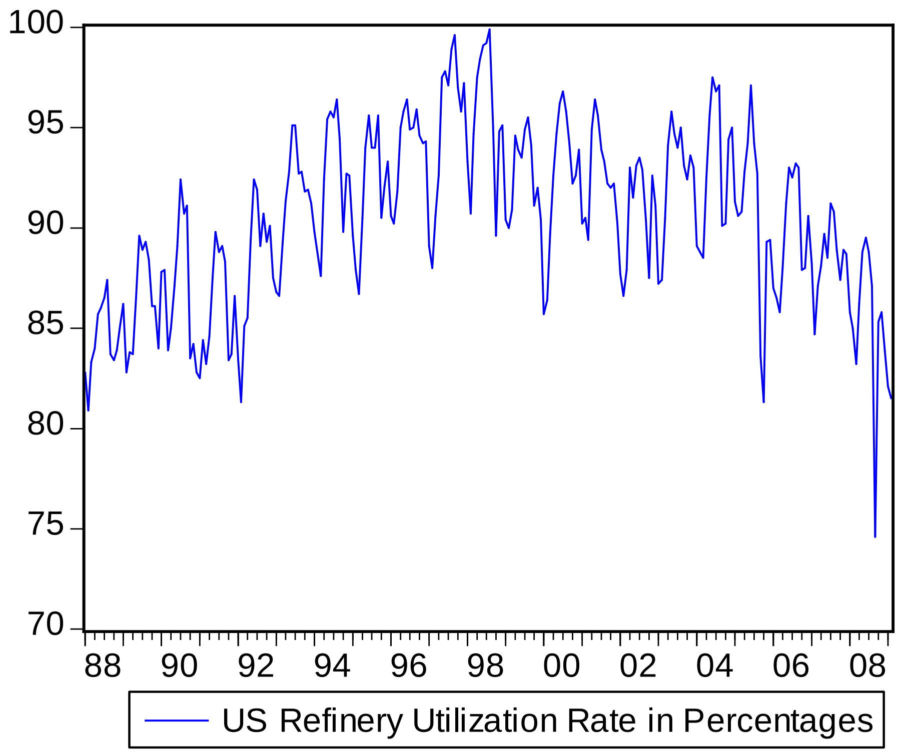

Figure 13. US Refinery Utilization Rate

9.15 mbd in 2005. Of this 1.63 mbd increase, 1.58 mbd came from medium grades of crude oil - the production of light grades of crude oil increased by only 0.053 mbd (Eni

Spa, 2006). Because of the price spread among the crude oils, this change in the composition of crude oil reduced the average price of crude oil produced by Saudi Arabia (Kaufmann et al., 2008).

Taking into account the effect of refinery utilization on the composition of crude oil imported to the United States, I expect to see the real average FOB price of US imported crude oil to decline as utilization picks up. On the other hand, the effect of refinery utilization on WTI should not be influenced by the composition effect mentioned above. It is of great interest to examine whether refutil has an opposite effect on the WTI series.

Turning to the time series properties of the refutil series, I run the ADF test as usual. The ADF test statistic of the level of the series does not exceed the test critical values at any of the conventional significance levels so I cannot reject the null hypothesis of a unit root. I take the first difference of the series and check for stationarity. The ADF test statistic of the first difference of the series is significant at the 1% level implying that the differenced series is stationary.

Table 7. ADF Test Statistic of Refinery Utilization

Refutil	-1.145586
ΔRefutil	-5.593664***

Lag length is determined by the Schwarz Information Criterion

1%, 5%, 10% significance levels are denoted by ***, **, *

3.3 Variable Summary

In the previous sub-sections we reviewed the time series properties of the series which are to be used in the following econometric analysis. Table 8 and Table 9 summarize the results of the stationarity tests and expectations about these variables.

Table 8. ADF Test Results Summary

	Price	WTI	Stocks	Spread	Caputil	Refutil
Level	-2.600171*	-2.140102*	-2.571302	-1.931434	-1.644416	-1.145586
Δ	-6.411403	-7.874281	-3.381086	-6.193981	-6.836433	-5.593664

Lag length is determined by the Schwarz Information Criterion
1%, 5%, 10% significance levels are denoted by ***, **, *

Table 9. Summary of Variables and Hypothesis

	Description	Order of Integration			Expected Sign
Dependant Variable		Sample 1	Sample 2	Sample 3	
Price	Real Average FOB Price of US Crude Oil Imports	I(1)	I(1)	I(1)	
WTI	Price of Western Texas Intermediate	I(1)	I(1)	I(1)	
Independent Variables					
Stocks	Days of Forward Consumption of OECD Oil Stocks	I(1)	I(1)	I(1)	(-)
Opecutil	OPEC Capacity Utilization	I(1)	I(1)	I(1)	(+)
Caputil	OPEC Capacity Utilization Modifies by OPEC Cheat	I(0)	I(1)	I(1)	(-)
Refutil	US Refinery Utilization Rate (as percentage)	I(1)	I(1)	I(1)	(-)
Spread	Spread between price of 4 months and 1 month futures contracts at the NYMEX	I(0)	I(0)	I(1)	(+)

Sample 1: January 1988 - February 2009, Sample 2: March 1995 - February 2009, Sample 3: January 2002 - February 2009

3.4 Model

Considering the frequency and length of the sample period, it is difficult to fit a model which could capture different factors causing the price to fluctuate under different circumstances. The possibility of structural breaks make the results of such a model unreliable nor is such a model desirable since my primary interest lies in the inspection of the recent price movements. However in order to check whether my expectation about the variables hold, a general model spanning the entire sample is investigated;

$$\text{Price}_t = \alpha + \beta_1 \text{stocks}_t + \beta_2 \text{opecutil}_t + \beta_3 \text{refutil}_t \quad (1)$$ [12]

As indicated in section 2.1, the dependent and independent variables contain a stochastic trend. The presence of I(1) trends invalidates the blind application of ordinary least squares (OLS) because the diagnostic statistics generated by OLS will indicate a meaningful relationship among unrelated I(1) variables more often than implied by random chance (Granger and Newbold, 1974). Such relations are termed spurious regressions[13]. To avoid falling into the spurious regression trap, the relationships in Eq. (1) are evaluated by determining whether they co-integrate. Following the method proposed by Engle and Granger (1987), Eq. (1) is evaluated by OLS to determine whether two (or more) variables co-integrate. The regression error is analyzed for a stochastic trend using the ADF test. If the test statistic fails to reject the null hypothesis, the non-stationary residual indicates that the regression is spurious. Otherwise, the regression error is stationary, and therefore the variables co-integrate. In this case, Eq.(1) can be interpreted as a co-integrating relationship. Table 10 reports the result of the ADF test of the regression residuals of the Eq.(1)

[12] Note that opecutil (OPEC production relative to OPEC capacity) is used instead of caputil since for caputil I reject the null of a unit root for the sample January 1988 - February 2009

[13] For more on “Spurious regressions” refer to Appendix A.

Table 10. ADF Test Result of Eq. (1) Residuals

OLS regression residuals of Eq.(1)		-4.3487***
Test critical values:	1% level	-2.574513
	5% level	-1.942136
	10% level	-1.615828

However, even if the variables co-integrate, the OLS estimate of the co-integrating vector will contain a small sample bias and the limiting distribution will be non-normal with a non-zero mean (Stock 1987). To avoid confusion associated with this bias, the co-integrating relationship among non-stationary variables in Eq.(1) is estimated using dynamic ordinary least squares[14] (DOLS) (Stock and Watson, 1993). DOLS generates asymptotically efficient estimates of the regression coefficients for variables that co-integrate, it is computationally simple, and it performs well relative to other asymptotically efficient estimation of the co-integrating vector. Lags and leads used to estimate the DOLS version of Eq.(1) are chosen using the Akaike Information criterion (Schwarz, 1978). Table 11 reports the DOLS estimates of the Eq.(1).

Table 11. DOLS estimate of Eq. (1)

Dependent Variable	US FOB Price
constant	682.8232***
	(88.64407)
stocks	-3.724029***
	(0.463023)
opecutil	0.574342*
	(0.311062)
refutil	-5.399462***
	(0.777881)
R^2	0.727013

Newey-West standard errors in parenthesis. 1%, 5%, 10% significance levels are labeled by ***, **, *. CRDW[15] is the Co-integration Regression Durbin Watson statistic. The number of lags and leads are chosen using the Akaike Information Criterion.

[14] For more on DOLS refer to the Appendix B.
[15] For the table of CRDW critical values see Appendix C.

All variables have the expected signs. Although Eq.(1) does not have the cubic caputil specification and the spread variable, it is consistent with the theory and various previous studies (cf. Kaufmann et al., 2004; Möbert, 2007; Kaufmann et al. 2008). Since I would like to include the effects of the introduction of the options contracts on the crude oil market and especially want to focus on the 2002-2009 period I move to make sub-samples which correspond to my period of interest, namely March 1995 – February 2009. All variables introduced in section 2.2 are I(1) in the sample except for spread. The core model of this paper is provided below;

$$\text{Price}_t = \alpha + \beta_1 \text{Stocks}_t + \beta_2 \text{Caputil}_t + \beta_3 \text{Caputil}_t^2 + \beta_4 \text{Caputil}_t^3 + \beta_5 \text{Refutil}_t + \beta_6 \text{Spread}_t \qquad (2)$$

Excluding spread, I can use DOLS to estimate this regression if the regression errors of the OLS estimate of (2) – excluding spread – are stationary. The ADF test results of the regression error series of the OLS estimate of (2) –excluding spread– is below.

Table 12. ADF Test Result of Eq. (2) **–** excluding spread – Residuals

OLS regression residuals of Eq.(1)		-3.9364***
Test critical values:	1% level	-2.746354
	5% level	-1.786861
	10% level	-1.427485

Since the null hypothesis of a unit root can be rejected at the 1% level, the regression error series of the OLS estimate of (2) – excluding spread – is stationary which makes DOLS a suitable method to apply.

Eq.(2) – excluding spread – is estimated for the sample mentioned above. The results are summarized in Table 13.

Table 13. DOLS estimates of the Eq. (2) –excluding spread– March 95–Feb. 09

Dependent Variable	US FOB Price
Constant	692.3555***
	(70.86937)
Stocks	-2.963154***
	(1.013563)
Caputil	3.538831***
	(1.189863)
$Caputil^2$	-0.073443***
	0.027704
$Caputil^3$	0.000250***
	9.57E-05
Refutil	-5.701487***
	(1.001566)
R^2	0.830935
CRDW	0.420889

Newey-West standard errors in parenthesis. 1%, 5%, 10% significance levels are labeled by ***, **, *. CRDW is the Co-integration Regression Durbin Watson statistic. The number of lags and leads are chosen using the Akaike Information Criterion.

All coefficients are consistent with the previous literature and carry the expected signs mentioned in the sections 2.2. Considering the increase in the activity in the NYMEX I suspect that the spread variable could possibly have an important effect in the sub-sample which corresponds to the period of a 500% rise and fall of the prices, namely the 2002-2009 period. Since spread is non-stationary during this period I can now run (2) with the spread, however I check first the stationarity of the regression error series of (2) –including spread – to see if DOLS is applicable. Table 14 shows that the ADF test

Table 14. ADF Test Result of Eq. (2) –including spread– Residuals

OLS regression residuals of Eq.(1)		-4.2154***
Test critical values:	1% level	-2.786359
	5% level	-1.890341
	10% level	-1.435648

statistic of the residual series of the OLS regression of (2) exceed the test critical values, so I can reject the null hypothesis of a unit root and conclude that the error series is stationary.

The results (Table 16) for this sub-sample do indicate that the spread variable is significant and indeed is the single largest contributor to the price hike in this period. The large coefficient of the spread variable could be attributed to the fact that the price spread usually is not very large between the spot and futures prices.

Table 15. DOLS estimates of the Eq. (2) for sub-sample Jan 2002 – Feb. 2009

Dependent Variable	US FOB Price
Constant	701.5715***
	(134.1937)
Stocks	-8.679329***
	(1.090550)
Caputil	-34.44433***
	(6.311274)
Caputil^2	2.211522***
	(0.452999)
Caputil^3	-0.036136***
	(0.009133)
Refutil	-0.579780***
	(1.358936)
Spread	13.76556***
	(2.523493)
R^2	0.940508
CRDW	0.4765163

Newey-West standard errors in parenthesis. 1%, 5%, 10% significance levels are labeled by ***, **, *. CRDW is the Co-integration Regression Durbin Watson statistic. The number of lags and leads are chosen using the Akaike Information Criterion.

Note that in the 2002-2009 period the cubic specification of caputil is somehow altered. This change can be interpreted as a change in the price mechanism during this period. To check if there is a structural break corresponding to the beginning of the sub-sample (January 2002) I carry out a Chow Break point test on the full sample of Eq.(2), namely March 1995 – February 2009. Test results confirm that a structural break takes place in

January 2002 (Table 11). The intuitive sub-sample is now confirmed with econometric evidence and further analysis of recent price changes could be based on this sub-sample.

Table 16. Chow Break Point Test

F-statistic	2.175177	Prob. F(41,83)	0.001387
Log likelihood ratio	1.204026	Prob. Chi-Square(41)	0.000000

The coefficients estimated by DOLS represent the long run relationship among variables. DOLS does not estimate the short-run dynamics – it is not necessary for asymptotically efficient estimation of the co-integrating relation. To examine the short run relationship among the variables in Eq. (2), OLS is used to estimate the following error correction model (ECM);

$$\Delta \text{Price}_t = \kappa + \rho\eta_{t-1} + \sum_{i=1}^{s} \lambda_{1i}\Delta \text{Stocks}_{t-i} + \sum_{i=1}^{s} \lambda_{2i}\Delta \text{Caputil}_{t-i} + \sum_{i=1}^{s} \lambda_{3i}\Delta \text{Caputil}_{t-i}^2 + \sum_{i=1}^{s} \lambda_{4i}\Delta \text{Caputil}_{t-i}^3 + \sum_{i=1}^{s} \lambda_{5i}\Delta \text{Refutil}_{t-i} + \sum_{i-1}^{s} \lambda_{6i}\Delta \text{Price}_{t-i} \quad (3)$$

in which Δ is the difference operator, and η is the difference between the observed value for real oil price and the value generated by the co-integrating relationship (Eq.(2)). The number of lags(s) for the right-hand side variables in Eq.(3) are chosen using the Akaike Information Criterion (Akaike, 1973).

The statistical significance of ρ in Eq. (3) is used to test the null hypothesis that prices are not affected by disequilibrium η between observed real oil prices and the equilibrium that is implied by the right hand side variables in Eq. (2). A negative value for ρ indicates that differences between observed values for crude oil prices and the equilibrium price moves observed prices towards the equilibrium value implied by the co-integrating relationship. Under these conditions, the right-hand side variables in Eq. (2) are said to “Granger cause” real oil prices. Table 17 reports the ECM results.

Table 17. OLS estimates of the Eq. (3)

Dependent Variable	US FOB Price	
Sample	1995M3-2009M2	2002M1-2009M2
ρ	-0.052038*	-0.043825*
DW	-1.742736	-0.704821
R^2	0.488013	0.595968

t-statistics in parenthesis. 1%, 5%, 10% significance levels are labeled by ***, **, *. DW is the Durbin-Watson statistic. The number of lags and leads are chosen using the Akaike Information Criterion.

The ECM estimation results of Eq. (3) for both the entire and the sub-sample periods indicate that the co-integrating relationship given by Eq. (2) can be interpreted as an equation for the long-term determinants of price. The error-correction term η has a negative sign and is significant in both samples which mean that the right-hand side variables adjust (albeit slowly, with only 5% correction each period, i.e. every month) to the difference between the observed value of oil prices and the value indicated by the equilibrium. These results prove that OECD stocks, capacity utilization, refinery utilization and expectations at the NYMEX "granger cause" the real oil prices. The models adequacy, the results and causality will be discussed in detail in the next section.

3.5 Model Robustness and Causality

In order to check that the model in the previous section is satisfactory in terms of standard time series assumptions I checked for residual normality, misspecification, omitted variables and parameter stability. The model passed all the diagnostic tests (Table 13) which makes me confident that the model is well defined and one can rely on the results of the model for accurate conclusions.

Table 18. Testing Non-Linearities and Stability

	Ramsey Reset 1	Ramsey Reset 2	Omitted Variables	Redundant Variables
F-statistic	1.797956	1.487341	2.418436	12.98768
Log likelihood	4.262202	7.377748	44.81327	31.38126
Prob. F	0.0000	0.0412	0.003378	0.00008
Prob. Chi-Sq.	0.0000	0.0250	0.000148	0.00000

To check for the direction of causality between price and the right hand side variables I run granger causality tests in all directions. I construct a Vector Autoregressive (VAR) model with four endogenous variables; Prices, Stocks, Caputil and Refutil. The direction of causal order is determined by restricting the lagged values of an endogenous variable to zero. Rejecting the null hypothesis that the coefficients are zero (i.e. the excluded variable does not granger cause the dependent variable) indicate that the lagged values of the excluded variable have information about the dependent variable beyond that contained in lagged values of the dependent variable and the other right-hand side variables. Restrictions on the lagged value(s) of a variable are evaluated with a Wald Test. Table 19 reports the results.

Table 19. Test of Causal Order on the VAR

	Independent Variables			
Dependent Variable	Price	Stocks	Caputil	Refutil
Price		11.35796**	10.315376*	13.16752**
Stocks	5.624565		8.562771	9.80084
Caputil	3.198106	4.066737		7.272634
Refutil	4.14295	6.44444	5.739726	

Numbers represent Chi-squared values with 6 degrees of freedom
1%, 5%, 10% significance levels are labeled by ***, **, *
The lag length of VAR is chosen using the AIC

The results indicate that OECD stocks, OPEC capacity utilization, and Refinery utilization rate “granger cause” real prices while there is no evidence of any causality from prices to any right hand side variable.

Combined, the above results confirm that the DOLS assumptions are satisfied, Eq. (2) constitute a long-term relationship between price and the right-hand side variables and the model is reliable as far as the econometric specification is concerned. The interpretation of the results from section 2.3 involves reference to various other market indicators. The next section will discuss the results in depth with reference to the relating market information.

4 Evaluation of the Results

The econometric model specified in chapter 2 does agree with the previous literature and expectations. To see if the results really are justified by the explanations put forth in section 2, I re-estimate Eq.(2) with WTI prices as the dependent variable. The results are in table 15.

Table 20. DOLS estimates of the Eq. (2) for sample March 1995 – Feb. 2009

Dependent Variable	WTI price
Constant	747.3555***
	(70.86937)
Stocks	-2.429348***
	(0.578494)
Caputil	3.365717***
	(0.628841)
$Caputil^2$	-0.074101***
	0.14350
$Caputil^3$	0.000258***
	4.88E-55
Refutil	-4.201593***
	(1.01635)
Spread	2.534128
	(1.585265)
R^2	0.801724
CRDW	0.438354

Newey-West standard errors in parenthesis. 1%, 5%, 10% significance levels are labeled by ***, **, *.CRDW is the Co-integration Relation Durbin-Watson statistic. The number of lags and leads are chosen using the Akaike Information Criterion.

All coefficients maintain their signs and all variables are significant except for spread. Since the two price series follow each other the result is not surprising. The main question of interest lies in the variable refine since our argument regarding the negative effect of refinery utilization rate stems from the idea that as the refineries increase their utilization rate, the composition of the imported crude oil shifts in favor of less valuable crude oil types i.e. heavy sour. As I replace the dependent variable real average FOB price of crude oil imported to the US to the Western Texas Intermediate (a light sweet type) the effect of the refutil variable should erode. This is precisely what we observe in the results reported in table 15. The effect of the refutil variable, while still negative and significant, declined.

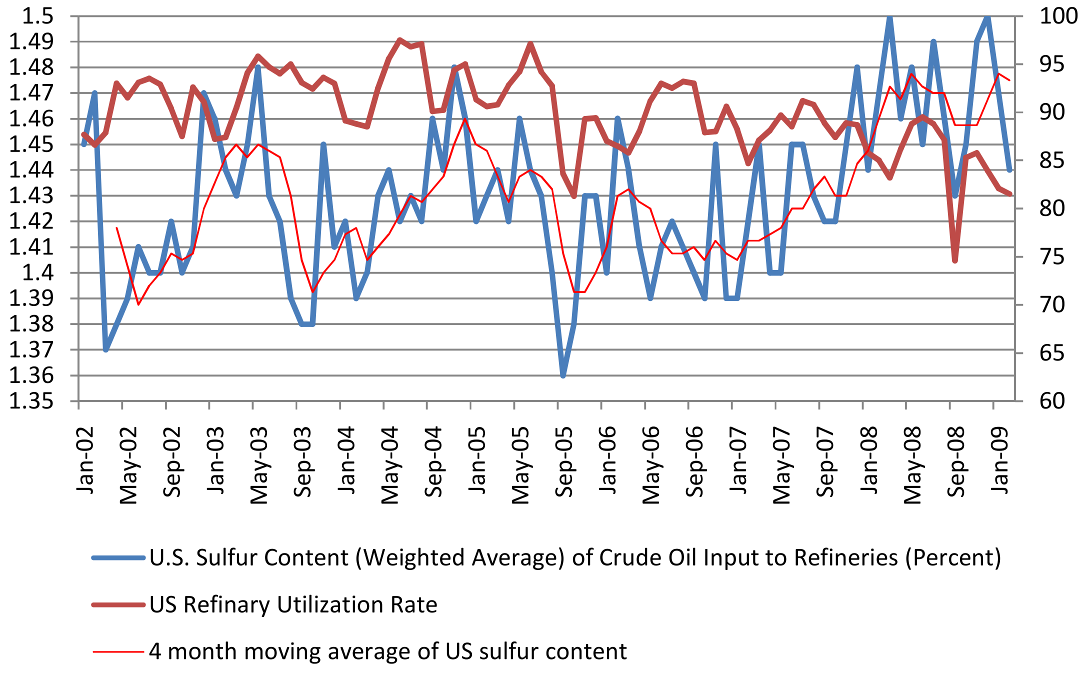

Figure 14. Sulfur Content vs. Refinery Utilization

One can observe that the sulfur content of the crude oil inputs tend to increase as the refinery utilization rates increase (Figure 14). This is consistent with the theory that refineries are forced to use lower quality crudes in order to match demand. It follows from this argument that the price difference between the lower and higher quality crudes widen in times of high refinery utilization because the refiners are not willing to pay as much for the lower quality crude oils due to lower yields and maintenance costs. To check whether this is true I plot the price difference between Saudi Arabian heavy (API gravity 27) and the Western Texas Intermediate (Figure 15). The figure does support the view that higher refinery utilization rates are associated with higher price differentials between higher and lower quality crude oils.

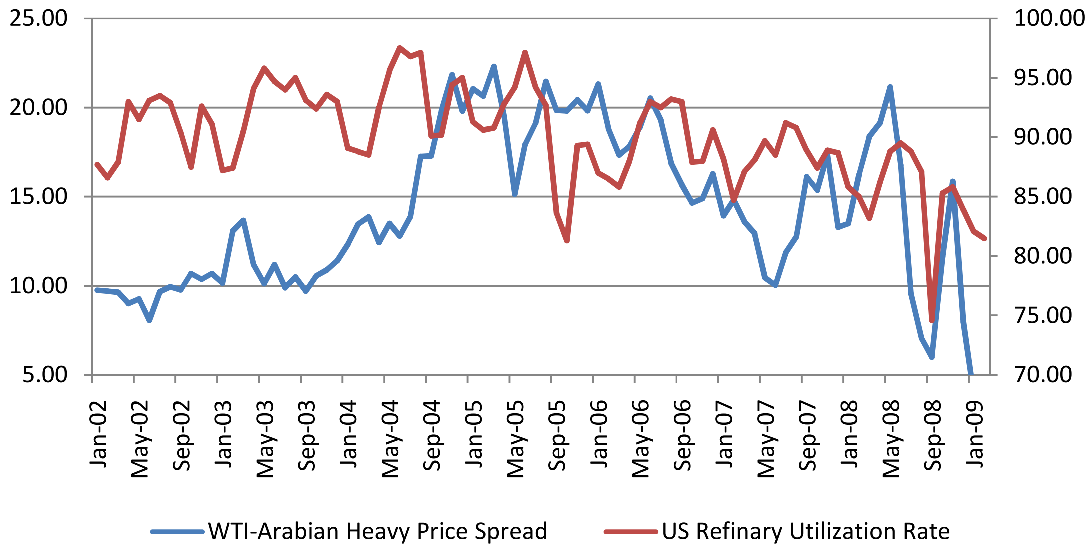

Figure 15. Heavy-Light Spread and US Refinery Utilization Rate

Since the models adequacy has been found satisfactory and the underlying assertions regarding the price discovery function has been found consistent with the data, I compare the models with the observed prices in the sample period March 1995 – February 2009

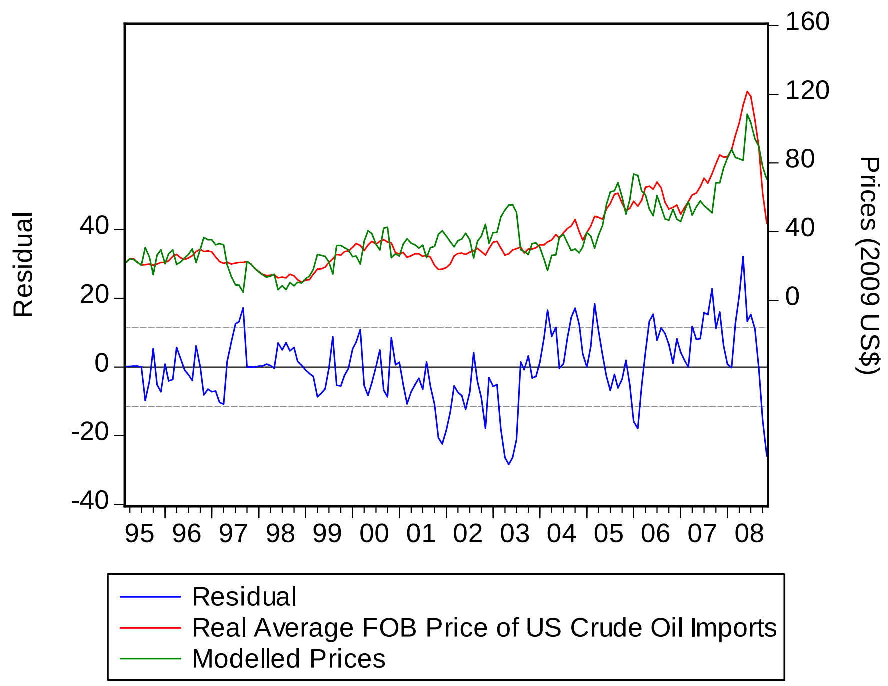

Figure 16. Modeled vs. Actual Prices

The model does a fairly good job except for the period 2001-2004 where the model consistently overestimates the prices and for the 2004-2005 period where the model underestimates the prices.

5 Conclusion

In this thesis I decomposed the factors affecting the price of crude oil. Variables common in the literature were used to run DOLS regression based on monthly data from March 1995 to February 2009. After discussing the theoretical expectations and time series behavior of each variable, I began my analysis with a very simple and general model covering the entire post 1986-quota period. Later I introduced my main sample staring with the introduction of the options contracts in the NYMEX in March 1995 and then further restricted the sample to the 2002-2009 period where a spectacular price rise took place. In all sample periods the variables carry the expected signs. Spread variables contribute the most to the oil price increase in the 2002-2009 period. This is consistent with the information presented in Figures 1, 5, 6 and 7. The rapid and large increase in the volume of oil contracts traded is reflected in the coefficient on the spread variable. A similar increase pushed the magnitude of the stocks variable about four folds. As the market becomes tighter it is reasonable to say that a marginal increase (decrease) in inventory levels have larger impact on prices. Each additional barrel withdrawn from the market will push prices ever higher as the spread between supply and demand increases. The refinery utilization rates support this claim as the refinery utilization rates reached historically high levels in the 2002-2009 period. The results also support the view that as refinery utilization increases the demand for different types of oil shift which actually applies downward pressure on the prices because lower quality crudes increase their share in the refinery inputs. Although several tests confirm that the results could be safely regarded as reliable one has to be cautious because of the possibility of conflicting results in alternative sample periods. Aware of this possibility, I conclude that financial markets did somewhat influence oil prices in the given period however market forces are the main drives in all samples as will the case be in any other sample.

Appendix A Data Sources

Nominal average FOB price of US crude oil imports and US refinery utilization rates were taken from the EIA website.

http://tonto.eia.doe.gov/dnav/pet/pet_pri_imc1_k_m.htm

Spot and futures prices of the WTI are obtained from Bloomberg Financial Services.

OECD demand, OECD commercial stocks of oil, Global demand (annually) were obtained from the IEA. Monthly global demand is based on own calculations.

OPEC quotas, production and capacity were obtained from OPEC secretariat.

Appendix B Spurious Regressions and Co-integration

The assumptions of the classical regression model necessitate that all variables be stationary and that the errors have a zero mean and a finite variance. In the presence of non-stationary variables OLS often produces a high R-squared and t-statistics that appear significant when no relationship between variables exist. While the regression seems to be good, OLS estimates are not consistent and the customary tests of statistical inference do not hold. Such regressions are named “spurious regression” after Granger and Newbold (1974).

In the special case where all variables are integrated of the same order and the residual sequence is stationary, the variables are co-integrated. Engle and Granger proposed the following method to test for co-integration which is used in this thesis.

1) Pretest the variables for their order of integration. By definition, co-integration necessitates that variables be integrated of the same order. If the variables are integrated of the same order proceed to step 2.
2) Estimate the long-run equilibrium relationship by running an OLS estimate of the variables. Save the residuals from this regression
3) Test (without a constant and a time trend) whether the saved residuals are stationary. If we can reject the null hypothesis of unit root, we can conclude that the residuals are stationary and the series are co-integrated implying that the regression results are not spurious.

Appendix C Dynamic Ordinary Least Squares (DOLS)

Instead of the inconsistent estimates of OLS in co-integrated systems, a number of methods have been proposed. Stock and Watson (1993) proposed the DOLS estimate which is used throughout this thesis, which can be written as:

$$y_{1t} = \gamma y_{2t} + \beta_0 \Delta y_{2t} + \beta_1 \Delta y_{2,t+1} + \beta_2 \Delta y_{2,t+2} + ... + \beta_p \Delta y_{2,t+p} + \beta_{-1} \Delta y_{2,t-1} \\ + \beta_{-2} \Delta y_{2,t-2} + ... + \beta_{-p} \Delta y_{2,t-p} + \varepsilon_t$$

DOLS maintains the strict exogeneity of the variables by including leads and lags of the changes of the I(1) regressors. This ensures the strict exogeneity of the levels of the I(1) regressors. This method circumvents the common problem of relaxing the classical regression assumption of no serial correlation in the error process.

The I(1) regressors used in this thesis are oil stocks, OPEC capacity utilization, refinery utilization and spread. Each of these variables takes the place of the variable y_{2t} in the above equation meaning that the level as well as the past and future changes - based on the AIC - of each of them is included in the DOLS equations estimated in this thesis.

Appendix D Co-integration Regression Durbin Watson

Co-integration Regression Durbin Watson (CRDW) is used in a single equation setting with a co-integrating relation. The CRDW critical values (Sargan and Bhargava, 1983) based on different sample sizes and number of regressors is presented in the table below.

n	T	CRDW
2	50	0.72
	100	0.38
	200	0.20
3	50	0.89
	100	0.48
	200	0.25
4	50	1.05
	100	0.58
	200	0.30
5	50	1.19
	100	0.68
	200	0.35

Source: Sargan and Bhargava, 1983.

SOURCES AND LITERATURE CITED

Abosedraa, Baghestani (2004). "On the Predictive Accuracy of crude oil futures prices." Energy Policy, 32(12): 1389-1393

Adelman, M.A. (2002). "World oil production and prices 1947–2000." The Quarterly Review of Economics and Finance, 42(2): 169-191

Agnolucci, P. (2009). "Volatility in Crude Oil Futures: A Comparison of the Predictive Ability of GARCH and Implied Volatility Models." Energy Economics, 31(2): 316-321

Al Faris, A. (1991). "The determinants of crude oil adjustments in the world petroleum market." OPEC Review, 15(3): 215-228

Alquist, Kilian (2008). "What do we learn from the price of crude oil futures?" Working Paper, University of Michigan

Asali, M. (2008). "Dynamics of petroleum markets in OECD countries in a monthly VAR&VEC model." OPEC Energy Review, 32(1): 54-87

Asche et al. (2002). "Price Relationships in the Petroleum Market: An Analysis of Crude Oil and Refined Product Prices." Energy Economics, 25: 289-301

Bacon R. (1991). "Modelling the Price of Oil." Oxford Review of Economic Policy, 7(2): 17-34

Bamberger, R., Pirog R. (2004). "The Strategic Petroleum Reserve: Possible Effects on Gasoline Prices of Selected Fill Policies." CRS Report for Congress

Bopp, Lady (1991). "A Comparison of Petroleum Futures versus Spot Prices as Predictors of Prices in the Future." Energy Economics, 13(4): 274-282

Brook et. al. (2002). *Introductory Econometrics for Finance*. Cambridge University Press

Brook et. al. (2004). "Oil Price Developments." OECD Economics Department Working Papers No. 412

Bryant, Bessler, Haigh (2002). "Causality in Futures Market." University of Maryland Working Paper No. 28574

Capuano, C. (2006). "Strategic noise traders and liquidity pressure with a physically deliverable futures contract." International Review of Economics & Finance, 15(1): 1-14

Ciner, C. (2006). "Hedging or Speculation in Derivative Markets: The Case of Energy Futures Contracts." Applied Financial Economics Letters, 2(3): 189-192

Chinn, LeBlanc, Coibion (2005). “The Predictive Content of Energy Futures: An Update on Petoleum, Natural Gas, Heating Oil and Gasoline.” NBER Working Paper. No. 11033

Cooper, J.C.B. (2002). “Price elasticity of demand for crude oil: estimates for 23 countries.” OPEC Review, 27(1): 1-8

Davidson, P. (2008). “Crude Oil Prices: "Market Fundamentals" or Speculation?” Challenge, 51(4): 110-118

Dees et al. (2007). “Modelling the World Oil Market: Assessment of a Quarterly Econometric Model.” Energy Policy, 35(1): 178-191

Eckaus, R.S. (2008). “The Oil Price Really Is A Speculative Bubble.” Center for Energy and Environmental Policy Research

Enders, W. (2004). *Applied Econometric Time Series.* John Wiles & Sons

Fattouh, B. (2006). *The Origins and Evolution of the Current International Oil Pricing System: A Critical Assessment*. Oxford: Oxford University Press

Fattouh, B. (2005). “The Causes of Crude Oil Price Volatility.” Middle East Economic Survey, 48(13)

Fattouh, B. (2008). “The Dynamics of Crude Oil Price Differentials.” Oxford Institute for Energy Studies. In Press.

French (2005). “Why and When do Spot Prices of Crude Oil Revert to Futures Price Levels?” Federal Reserve Board, Washington D.C.

Gracia, E. (2006). Bubbling Crude: Oil Price Speculation and Interest Rates. E-Journal of Petroleum Management and Economics

Hamilton (2008). “Understanding Crude Oil Prices.” UCSD

Hammoudeh et al. (2008). “Threshold Co-integration Analysis of Crude Oil Benchmarks.” The Energy Journal, 29(4): 79-95

Hayashi, F. (2000). *Econometrics*. Princeton, NJ: Princeton University Press.

Holland, S. (2008). “Modelling Peak Oil.” The Energy Journal, 29(2): 61-78

Horsnell, P., and Mabro, R. (1993). *Oil Markets and Prices: The Brent Market and the Formation of World Oil Prices*. Oxford: Oxford University Press

Hotelling, H. (1931). “The Economics of Exaustable Resources.” Journal of Political Economy, Vol.39, 137-175

Hwang, M.J., Yang, C.W. 11-14 October 2001. "Unstable Price Elasticity and High World Oil Price." The 52nd International Atlantic Economic Conference, Philadelphia, Pennsylvania.

Jalali-Naini, A. and Asali, M. (2004). "Cyclical Behaviour and Shock Persistance: Crude Oil Prices." OPEC Review, 28(2): 107-131

Johansen, S. (1988). "Statistical Analysis of Co-integration Vectors." Journal of Economic Dynamics and Control, 12(2): 231-254

Johansen, S., Juselius, K., (1990). "Maximum Likelihood Estimation and Inference on Co-integration with Application to the Demand for Money." Oxford Bulletin of Economics and Statistics, 52(2): 169-209

Jumah et al. (2009). "Modelling and Forecasting Oil Prices: The Role of Assymetric Cycles." The Energy Journal, 30(3): 137-145

Kaufmann et. al. (2004). „Assessing the Factors Behind Oil Price Changes." ECB Working Paper Series No. 855

Kaufmann et. al. (2004). "Does OPEC matter? An Econometric Analysis of Oil Prices." The Energy Journal, 25(4): 67-90

Kaufmann, Dees, Mann (2009). "Horizontal and vertical transmissions in the US oil supply chain." Energy Policy, 37(2): 644-650

Kaufmann, Laskowski (2005). "Causes for an asymmetric relation between the price of crude oil and refined petroleum products." Energy Policy, 33(12): 1587-1596

Kaufmann, Ullman (2009). "Oil Prices, Speculation, and Fundamentals: Interpreting Causal Relations Among Spot and Futures Prices." Energy Economics, 31(4): 550-558

Kilian, L. (2007). "Not All Oil Price Shocks Are Alike: Disentangling Demand and Supply Shocks in the Crude Oil Market." University of Michigan and CEPR

Krugman, P.(2008, May 12). The Oil Nonbubble. The New York Times

Kuper, G.H. (2002). "Measuring Oil Price Volatility." University of Groningen

Maslyuk, S. and Smyth, R. (2009). "Co-integration Between Oil Spot and Futures Prices of the Same and Different Grades in the Presence of Structural Change." Energy Policy, 37(X): 1687-1693

Masters, M. W. and White, A. K. (2008). “The Accidental Hunt Brothers: How Institutional Investors Are Driving Up Food and Energy Prices”, Special Report for US Congress Permanent Committee on Investigations

Moshiri, S. and Foroutan, F. (2006). “Forecasting Nonlinear Crude Oil Futures Prices.” The Energy Journal, 27(4): 81-95

Möbert, J. (2007). “Crude Oil Price Determinants.” Darmstadt Discussion Papers in Economics Nr. 186

Murat, A. and Tokat E. (2009). “Forecesting Oil Price Movements with Crack Spread Futures.” Energy Economics, 31(X): 85-90

Okogu, B. E. (1989). “A theoretical and econometric study of the determinants of crude oil spot prices.” OPEC Review, 13(4): 349-361

Pindyck (1999). “The long-run evolution of energy prices.” Energy Journal , 20(2): 1-27

Reitz, Slopek (2008). “Nonlinear Oil Price Dynamics - A Tale of Heterogeneous Speculators?” Deutsche Bundesbank

Stevans, Sessions (2008). “Speculation, Futures Prices, and the U.S. Real Price of Crude Oil.” Munich Personal RePEc Archive

Stock, J.H., Watson, M.W. (1993). “A Simple Estimator of Cointegrating Vectors in Higher Order Integrated Systems..” Econometrica, 61(4): 783-820

Venkatesh, M. R. The Real Reason Oil Prices Are Rising. 1 May 2008. www.rediff.com

Verleger, P.K. Explaining the 2008 Crude Oil Price Rise. 18 February 2008. www.pkverlegerllc.com

Verleger, P.K. (2007). “How Wall Street Controls Oil.” The International Economy, 21(1): 14-17

Vogelvang B. (2005). *Econometrics Theory and Application with Eviews.* Pearson Education

Walt, V. (2009). Oil Is Plentiful, Demand Weak. Why Are Gas Prices Going Up? Time Magazine

Weiner, R. J. (1991). “Is the World Oil Market “One Great Pool?” The Energy Journal, 12(3): 95–107

Weiner, R.J. (2002). “Sheep in wolves' clothing? Speculators and price volatility in petroleum futures.” The Quarterly Review of Economics and Finance, 42(2): 391-400

Wirl F. (2008). “Why do oil prices jump (or fall)?” Energy Policy, 36(3): 1029-1043

Wooldridge, J.M. (2006). *Introductory Econometrics: A Modern Approach*. Thomson South Western

Yanagisawa, A. (2008). “Decomposition Analysis of Souring Crude Oil Prices.” The Energy Data and Modelling Center

Yang, C.W., Hwang, M.J., and Huang, B.N. (2002). “An analysis of factors affecting price volatility of the US oil market.” Energy Economics, 24(2): 107-119

Abstract

This master thesis decomposes the effects of various factors on the fluctuation of the price of the world's most traded commodity both in physical and paper terms, namely crude oil, based on monthly data between March 1995 and February 2009. Time series analysis is employed in order to reveal the driving forces behind the oil price fluctuations and their magnitude. A broad set of variables are employed to test the impact on price changes. Considering the frequency and length of the sample period, sub-samples were considered in anticipation of structural breaks. Consequently a general model is built to check the expectations about the fundamental variables and then a sub-sample from January 2002 to February 2009 is used in order to reveal the factors behind the spectacular price rise and fall during this period.

Abstract

Rohöl ist die am meisten gehandelte Ware, unabhängig davon ob der Handel in Barrels oder in Dollars gemessen wird. Diese Arbeit untersucht verschiedene Einflussfaktoren auf den Rohölpreis anhand von monatlichen Daten zwischen März 1995 und Februar 2009. Mittels einer Zeitreihenanalyse wird versucht die treibenden Kräfte hinter der Preisbildung zu identifizieren. Hierzu wird ein breites Set an Variablen eingesetzt; in Erwartung eines Strukturbruchs werden diese in Sub-Samples unterteilt. Mit Hilfe eines allgemeinen Models werden dann die Erwartungen über die Fundamentaldaten überprüft. Im Abschluß wird mittels des gebildeten Sub-Samples von Januar 2002 bis Februar 2009 der spektakuläre Preisanstieg sowie Preisverfall in dieser Zeitperiode auf mögliche Faktoren untersucht.

Arda S. Köstem

Pelzgasse 12/12 1150 Vienna, Austria • Handy #: 0676 3602595 • E-mail: ardaskostem@hotmail.com

---

Persönliche Daten

Geburtsdatum	: 04.07.1982
Geburtsort	: Ankara/Türkei
Familienstand	: Ledig kinderlos
Staatsbürgerschaft	: Türkei

Ausbildung

2007-2009 Masterstudium Volkswirtschaftslehre, Universität Wien
Studienrichtung: Ökonomisch-mathematische Methoden

2000-2006 Bachelorstudium Volkswirtschaftslehre, West Virginia University
Studienrichtung: Allgemeine Volkswirtschaft

1989-2000 Matura, METU Gymnasium
Studienrichtung: Mathematik-Türkisch

Berufliche Erfahrung

2005-2006 Forschungsassistent, West Virginia Bureau of Business and Economic Research

- Teammitglied bei den bundesstaatlichen wissenschaftlichen Aufgaben im Bereich der mathematischen Modellierung der makroökonomischen Prozesse
- Assistent für das Projekt "Development of a New School Aid Formula for the State of West Virginia"

2005-2006 Forschungsassistent, West Virginia University, Department of Economics

- Assistent für Prof. Mehmet S. Tosuns Presentation "Explaining the Variation in Tax Structures in the MENA Region"
- Projektbezogene statistische Datenanalyse und Forschung

Praktikums

08-12/2009 United Nations Industrial Development Organization, Wien

- Teammitglied, Investment Promotion Unit, die sich mit Beförderung der Investitionen in Entwicklungsländer beschäftigt. Für diesen Zweck habe ich Daten sortiert, analysiert und regressionen Methoden verwendet.

06-08/2006 Ökonomische Forshung Abteilung, Türkisches Ministerium für Finanzen

- Beschäftigt mit Datenselektion, -analyse und –eintrag (SPSS)

06-08/2005 Statistische Abteilung, Türkische Notenbank

- Beschäftigt mit Datenselektion, -analyse und –eintrag (MS Excel)

Zusätzliche Kurse

03/2008 Transition und Privatisierung in Mittel und Osteuropa (Intensive), Wirtschaftsuniversität Praha

Außerschulischen Erfolge und Ehrungen

2000-2006 WVU Athletische Stipendium
2002-2005 Türkische Schwimm-Nationalmannschaft
2002-2005 Türkischer Meister in 200m Brustschwimmen
2002-2005 Teilnehmer in World University Olympics, Europaische Mesiterschaften und Mittelmeer Spielen

Spezielle Kenntnisse

Sprachen: Turkisch (Muttersprache), Englisch (fließend), Deutsch (fließend)
Software: Eviews, Stata, SPSS, MS Office
Führerschein Gruppe B

Referenzen

Auf Anfrage