Is the Oil Market Efficient?
-A Cointegration Study of Spot and Futures Prices-

Author: Mattias Nilsson
Thesis advisor: Darush Yazdanfar
Abstract
The oil market is arguably the most influential commodity market in the world, in that it has an effect on all economic variables in one way or another. Due to oil’s central role in the world economy, it is of the utmost importance that all parts of society strive to increase the understanding of how the market works. This study has analysed the efficiency of the oil market in the period 1986 to 2008, with the efficient market hypothesis as the theoretical framework. Data on the prices of spot and futures contracts on crude and heating oil has been collected from the New York Mercantile Exchange, and tested for cointegration, with the underlying assumption being that cointegration is a sign of weak form efficiency. The results implies that the spot and futures prices have not been cointegrated during the studied period, and thus we conclude that the oil market has not behaved in accordance with the weak form of the efficient market hypothesis.
Acknowledgements

I would like to thank everyone who has offered me support during the process of writing this thesis. I would like to acknowledge the contribution of opponents on seminars along the way, who have all contributed valuable and constructive criticism, and who have helped me by offering their unique perspectives on my work. I would especially like to thank my supervisor Darush Yazdanfar, without whom this thesis would never have been finished.
Figures
Figure 1 - Historical development of the crude oil price, 1861 – 2006 13

Tables
Table 1 – Results of the descriptive sample analysis 17
Table 2 - Results of the ADF-test performed on the crude oil data 18
Table 3 - Results of the ADF-test performed on the heating oil data 18
Table 4 - Results of the cointegration test performed on the crude oil data 19
Table 5 - Results of the cointegration test performed on the heating oil data 19
1. Introduction
1.1 Background

The development of the oil price has for a long time had major effects on the world economy. The importance of the oil market has been proven on several occasions, not least so by the oil crisis of the early 1970s, and the political unrest in the Middle East during the last few years has reminded us that the western economies are just as easily affected by changes in the oil price today as they were a few decades ago. Studies have shown that the price of oil affects the GDP growth, as well as other macroeconomic variables, financial markets and stock price volatility (Lardic and Mignon, 2006; Gisser and Goodwin, 1986; Sadorsky, 2003). What has changed since in recent decades is the extent to which futures contracts are being used to stabilise the market. Introduced in the late 1970s and early 1980s, oil market futures function as a tool for price discovery as well as for risk sharing (Garbade and Silber, 1983). By employing strategies using these oil market futures, spanning from short one month contracts to longer contracts spanning several months, agents on the market have increased the predictability of the oil price movements and thus decreased the insecurity otherwise present on the market. An underlying assumption if this system is to work reasonably well is that the market in question is efficient, i.e. there is no way of earning extraordinary profits without taking on extraordinary risk. If the market is not efficient, the predicting power of the trade in futures contracts is flawed, and thus insecurity is going to be an ever present problem on the market (Kumar, 2004). Furthermore, an inefficient market will open the door for speculative behaviour, which will lead to prices reflecting not only the fundamental information regarding oil prices, but also factors important to short term speculative interests, which will increase the insecurity further. This, in turn, will drive agents involved in activities such as hedging out of the market, which will further decrease the efficiency. Since the oil price is of such importance to all sectors of the economy, from employment to the purchasing power of consumers, and from both a macro and a micro perspective, it is of the utmost importance that we understand the workings of the oil market.

This thesis will test the efficiency of the oil futures market by analysing spot and futures prices, and examine the two for cointegration. In an efficient market, the spot and futures prices are based on the same information, rationally processed by agents
on the market. Thus the prices will be cointegrated if the market is efficient (Bekiros and Diks, 2008).

1.2 Previous Studies

Previous studies performed on the subject have reached various conclusions. For example, Gülen (1998) found that spot and futures prices tend to be more cointegrated if the futures contract considered is a short one. Green and Mork (1991) performed a study analysing the efficiency of the crude oil market between 1978 and 1985. They found that efficiency had to be rejected for the period as a whole, but that improvements could be detected over time. Lorne N. Switzer and Mario El-Khoury (2007) analysed the efficiency of the New York Mercantile Exchange (NYMEX) during volatile periods and reached the conclusion that spot and futures prices had been cointegrated, and thus that the market had been functioning efficiently. The same result was found by Ewing et al. (2006), who performed a cointegration study where they analysed the relationship between spot and three month futures contracts between 1986 and 2004. Dempster et al. (2006) tested the valuation and hedging of spread options where the two assets used were cointegrated, and found that it is important to investigate cointegration, since it affects the way strategies can best be employed. A similar conclusion was drawn by Alexander (1999), in a study performed on financial assets.

1.3 Purpose of Study

The purpose of this thesis is to analyse the efficiency of the oil market during a period spanning from 1986 to the present. This will be done by studying daily spot and futures prices on crude and heating oil traded on the NYMEX. The spot and futures prices will be checked for cointegration, with the underlying assumption that cointegration is a sign of weak form efficiency. Hopefully the findings of this study will increase our understanding of the workings of the oil market during this period of time.

1.4 Hypotheses

The tested hypotheses in this thesis are concerned with the issues of nonstationarity and cointegration respectively. A time series is considered to be stationary if its properties, such as the mean and the variance, are constant over time. Conversely, a
time series is nonstationary if any of these properties are not constant. It is important to test whether or not the data is nonstationarity, since if that indeed turns out to be the case; the equations employed might be spurious. This would imply that nonstationary variables might be highly correlated for noncasual reasons (Studenmund, 2006). The formal hypothesises tested in this thesis against adjusted t-values are presented below.

• $H_0$: The data is nonstationary.
• $H_1$: The data is stationary.

Cointegration between variables is when there exists a long-run equilibrium relationship between them (Studenmund, 2006). As mentioned earlier, cointegration between spot and futures prices on a market is an indicator that said market is efficient in the weak form, thus the test is very useful when dealing with issues of market efficiency. The tested hypotheses are presented below.

• $H_0$: The series are not cointegrated.
• $H_1$: The series are cointegrated.

1.5 Outline of Study
The study will start with a presentation of the methodology of the dissertation (2). This section will describe how the study has been made and what the reasoning behind specific choices made in the study has been. This will be followed by an introduction to the theoretical framework employed (3). After this a section giving a brief presentation of the oil futures market will follow (4), before a section describing the empirical findings is presented (5). The last section of the dissertation will describe the conclusions drawn from the findings (6).

1.6 Limitations
It is important to always be aware of the limitations of the methods employed when performing a study. In this case, it is important to see which version of the efficient market hypothesis is affected by the data we are analysing. It seems evident that we are looking at historical figures, why it must be concluded that only the weak form of the efficient market hypothesis can be tested using the cointegration method employed in this dissertation.
2. Methodology

2.1 Methodical Considerations

In this section, a brief explanation of the choices made regarding the approach of study, the objects of study and the handling of information will be presented.

2.1.1 Approach of Study

When attempting to analyse the efficiency of a market, there are several methods that can be employed. This thesis has primarily used a positivistic approach, in that the focus has been on analysing objective relationships amongst the variables in question. It should be mentioned that this approach have received some criticism for being too one-sided (Sayer, 1992). To test the market efficiency, the concept of cointegration, which means that there is a pattern or a relationship between two variables to which they will tend to return over time (Balke and Fomby, 1997), has been employed. In a perfectly efficient market, all available information will be reflected in the present price of a commodity, thus the futures price will be an unbiased predictor of future spot prices and the two will be cointegrated. The use of cointegration studies is widely spread when analysing the efficiency of a given market, and have for example been employed by Crowder and Hamed (1993) to study the oil market, and by Chen and Lin (2004) to study the market for lead on the London Metal Exchange.

2.1.2 Handling of Information

The information used in this study has primarily been found using online resources. The data on historical oil prices has been collected from the online database provided by the Energy Information Administration1, presenting official statistics from the U.S. government. This data has subsequently been analysed in Stata and Excel. As always when calculations are involved, there is an obvious risk that mistakes are made which lowers the reliability of the study. This risk have been minimised by going through the material and redoing the calculations several times. The articles used to describe the theoretical framework used in the study have primarily been found through the databases jstor.com and Google Scholar, and the literature used has been found through library search programs or through recommendations.

---

2.1.3 Selection of the Objects of Study
The objects studied in this dissertation have been chosen for various reasons. Crude oil has been chosen since it is the most important form in which oil is traded. It trades in the largest volumes and for the highest total nominal amount. This makes crude oil the most important indicator of oil market efficiency, and is as such an obvious choice when attempting to analyse the market in question. Heating oil has been chosen to increase the validity of the dissertation by including one more object of study, but also because heating oil has a direct impact on the lives of many people. A price change on heating oil directly changes the price of heating a large portion of homes, why it is an important market that needs to be studied.

2.1.4 Validity and Reliability of Study
The validity of a thesis can be defined as the extent to which the study measures what it purports to measure (Carmines and Zeller, 1979). When analysing a complex and to a certain extent subjective matter such as market efficiency, it is important to be aware of the always existing validity problem. It is impossible to guarantee perfect validity in any study, and efforts must be made continuously to evaluate the severity of potential problems (Druckman, 2005). To minimise validity problems in this thesis, a method of analysis has been chosen that is widely accepted by the academic community. This in itself does not guarantee validity, but it does function as an indicator that the method has been accepted and deemed suitable in the past. The reliability, i.e. the extent to which a study yields identical results on repeated trials (Carmines and Zeller, 1979), shares the characteristics of validity in that a study can never be guaranteed to be free from reliability problems. As has been mentioned above, the obvious risk of errors in calculations is ever present in quantitative research, and great efforts have been made to avoid such errors in the analysis. Apart from this potential problem the reliability of the study is arguably relatively high, due to the nature of the analysis. Focusing on objectively given data using an accepted and well defined methodology makes repeated studies with high precision possible.

2.2 How the Study Has Been Made
To study the efficiency of the oil market, daily price figures have been collected on crude and heating oil spot and futures contracts from June 2 1986 to April 15 2008. The futures contracts considered are one month-ahead contracts. The period in
question has been chosen since it arguably is sufficiently long for the results to be valid, while it is not long enough for structural changes to have a major impact and distort the results. Through the employment of the augmented Dickey-Fuller (ADF) unit root test the data was tested for stationarity (Studenmund, 2006). This was done through the regression of lagged time series, whereby hypothesis tests were performed. After this, the spot and futures contracts were tested for cointegration by the regression of lagged residuals, found through regular regression analysis, a procedure also known as the Engle-Granger test (Studenmund, 2006). This of course also was followed by testing the null hypothesis. A step-by-step summary of these calculations are presented in Appendix A, while a full view of the data output can be found in Appendix B. To increase the familiarity of the sample, the data was also subjected to a descriptive analysis, where measures such as variance, standard deviation, mean value, covariance and correlation were calculated. Details on these calculations are also presented in Appendix A.

2.3 Performed Tests

When analysing time series it is important to test whether the series in question are stationary or nonstationary. The reason for this is that an analysis of two nonstationary time series can produce spurious regression, i.e. an apparent relationship that in actuality does not exist (Wonnacott and Wonnacott, 1979). This problem can, when known, be avoided by using first differences of the studied series (Stock and Watson, 2007).

As has been previously mentioned, the ADF-test was used to test the data for stationarity. The equation used when employing this test is dependent on the characteristics of the time series in question. The alternatives, each corresponding to a different equation are (Studenmund, 2006):

- The time series has no constant and no trend.
- The time series has a constant but no trend.
- The time series has a constant and a trend.

The data analysed in this thesis has a nonzero mean, and no obvious trend, thus the constant no trend equation is chosen (specified in detail in Appendix A).
As mentioned above, the test chosen to examine the data for cointegration was the Engle-Granger test. This test is performed by employing the ADF-test on residuals found by running a regression of one series on the other. If we conclude that the residuals are stationary, the two time series tested are cointegrated (Stock and Watson, 2007).

3. The Efficient Market Hypothesis

The below sections will provide an introduction to the efficient market hypothesis (EMH), some critique against the theory, and the response of efficient market theorists to that critique. Throughout the sections some results of previous studies will also become evident.

3.1 The Birth of the EMH

The efficient market hypothesis was introduced to the academic community as a uniform theory by Eugene F. Fama in his 1970 survey *Efficient Capital Markets: A Review of Theory and Empirical Work*. Fama here defined an efficient market as one in which prices at all times fully reflect all available information. Theoretically, a market is assumed to be sufficiently efficient if there is no transaction costs; all available information is free and available to all market participants at the same time; and all market participants agree on the implications of information on the current and future price of securities (Fama, 1970). The rationale behind the efficient market hypothesis is that information is spread quickly throughout the investment community and that it is incorporated into the price of securities without delay. Thus, only new information affects the price of financial assets, and since new information by definition is unpredictable, equity prices move in an unpredictable manner. This implies that it is impossible to consistently achieve risk adjusted returns above that of the market as a whole (Malkiel, 2003). Put in another way; since all investors have the same information, one must accept a higher level of risk if one is to achieve above average returns.
3.2 Versions of the EMH

The efficient market hypothesis exists in three versions considering different subsets of information, the weak form; the semi-strong form; and the strong form (Fama, 1970). These different versions of the hypothesis are presented below.

3.2.1 Weak Form of the EMH

The basic difference between the three versions of the EMH is that they include different types of information in the expression “all available information” (Bodie et al. 2005). The weak form of the EMH states that equity prices fully reflects all available information contained in historical prices (Fama, 1970). The rationale behind this statement is that past equity price movements are readily and cheaply available to all investors, thus if a certain pattern could be detected it would immediately be exploited and become obsolete. For example, if the past performance of a stock told the market that a one week price increase always was followed by a 10% drop in the price, that information would be acted upon without delay, and the information would lose its value. One implication of the weak form of the EMH is that it makes technical analysis, or charting as it is also called, obsolete as a tool for predicting future equity prices (Malkiel, 1989). This is necessarily so since technical analysis is based on the idea that reliable signals about future prices can be found by analysing historical price movement patterns, which is exactly what the EMH states is not the case.

3.2.2 Semi-Strong Form of the EMH

Taking matters one step further, the semi-strong form of the EMH includes not only past performance in the expression “all available information”, but also all publicly known information regarding the prospects of a financial instrument, a firm or a commodity (Fama, 1970). In stocks, this can include information about such matters as a firm’s management team, balance sheet data or the competence of the R&D teams. In commodities, it can include information such as that regarding future increases in demand, a supply crisis in a market of substitute commodities or increases in cost of production due to for example strong worker’s unions. The semi-strong form of the EMH implies that there is no value in fundamental analysis when it comes to predicting future prices (Bodie et al. 2005). This is so since fundamental analysis concerns itself with analysing firm, industry and economy data in order to
find attractive investment opportunities likely to yield above average risk adjusted returns. If all that information already is in the hands of all investors, this is fruitless since the equity prices will already reflect the information. Ironically though, although the EMH states that fundamental analysis is of no use when trying to predict future prices, it is fundamental analysis that makes the market efficient (Hirt and Block 2006). If fundamental analysis did not exist, the available information would not be analysed and thus, not reflected in the current price of equities.

3.2.3 Strong Form of the EMH
Finally, the strong form of the EMH states that all information, not only that which is public, already is reflected in the price of equities (Fama, 1970). This implies that not even actors with monopolistic access to information, such as insiders, can profit from that information. The strong form of the EMH is quite extreme and would, if it holds true, make regulatory bodies such as the SEC in the U.S. superfluous to some extent.

3.3 The EMH and Its Critics
Going back 20 years, the EMH was widely accepted in the academic community, and markets were viewed by most as being highly efficient (Malkiel, 2003). This has to some extent changed in recent years. Program trading and behavioural finance are just two examples of relatively new phenomena that have given critics of the EMH ammunition. But even earlier, the EMH’s dismissal of technical and fundamental analysis did, not unexpectedly, receive severe critique. This came mostly from practitioners, but also from some academicians. A series of articles has been published during the last couple of decades in defense of technical analysis, a recent one being published in 2000 by Lo; Mamaysky and Wang. They studied a large number of U.S. stocks in the period 1962-1996 and found that technical analysis does indeed have some ability to predict future price movements. In support of fundamental analysis a series of articles has been published, for example concluding that investors are not rational, and thus they do not make rational use of available information. In their article *Prospect Theory: An Analysis of Decision under Risk* (1979), Kahneman and Tversky claim that investors underweight positive outcomes that are just probable compared to outcomes that are certain, due to their risk aversion. This implies that there are investments that offer above average risk adjusted returns to those investors able to accurately and rationally price risk. As an example of the
new criticism, mentioned above, that has been aimed at the EMH, behavioral finance followers has concluded that psychological factors affect investors and make them less than rational. In his article From Efficient Markets Theory to Behavioural Finance (2003), Shiller concludes that excess volatility and bubbles such as that in the IT sector in the late 1990s can be ascribed to psychological factors. He claims that as investors see prices rise, they get dragged into the market by a “bandwagon effect”. This implies that investors are not rational, and that factors other than rationally analysed information determine equity prices. It also implies that the psychologically induced behaviour of investors can be predicted and used to earn extraordinary profits.

3.4 The EMH's Response to the Critique
According to efficient market theorists, most of the critiques aimed at them are clearly off target. A number of articles responding to the claims of chartists that technical analysis does work have been published in recent years. As an example, Lesmond; Schill and Zhou’s paper The Illusory Nature of Momentum Profits (2004) concludes that even if technical analysis does have the ability to predict future price movements, the predicting power is still so weak that no profit opportunity remains after transaction costs have been taken into account. This conclusion, that transaction costs makes the intensive trading often supported by technical analysis unprofitable, is also reached by Odean in his article Do Investors Trade Too Much? (1999). On the matter of behavioural finance and the irrationality of investors, Fama published the article Market Efficiency, Long-Term Results and Behavioral Finance (1998) where he claims that the alleged irrationality of investors can not be used to earn extraordinary profits since overreaction to new information seems to be as common as under reaction. Furthermore he concludes that even if specific patterns of behaviour can be identified, these patterns disappear over time and thus can not be exploited profitably. To conclude, it is evident that the debate on whether or not markets are efficient has been a fierce one, and there is nothing to indicate that it will be over anytime soon. The evidence published so far does not give a clear answer either way, and studies will surely continue to be made on the subject.
4. The Oil Futures Market

It is a well established fact that the oil market is one of the most important markets in the world. Changes in the oil price have an effect on everything from interest rates to the price of petrol and the cost of heating a house. According to the World Energy Model developed by the organisation of oil producing countries (OPEC) the daily world demand for oil was 76 million barrels in the year 2000\(^2\). The model predicts that the demand will have increased to 90.6 million barrels/day by the year 2010. Crude oil, the most actively traded type of oil, is the original, unprocessed oil that goes on to become petrol, heating oil and the other oil products used in everyday life. As is evident in the figure below, depicting the historical crude oil price, major political events such as the Iranian revolution in 1979 and the recent invasion of Iraq are closely related to oil price shocks. This further emphasises the connection that exists between the oil market and the world economy, as well as world politics.


4.1 The Importance of the Oil Market
As mentioned above, changes in the price of oil affects the world economy in a series of ways. Several studies have found statistically significant indications that an increase in the oil price is a contributing factor behind recessions (Hamilton, 1983; Barsky and Lutz, 2004). A study performed by Gisser and Goodwin (1986) found that the oil price has an effect on macroeconomic variables that exceeds that of fiscal policy, and sometimes that of monetary policy.

4.2 Introduction to the Futures Market
Futures contracts have, in different forms, been around since the Middle Ages (Hull, 2001). The early futures contracts were predominantly unstandardised agreements between farmers and merchants, but with the opening of the Chicago Board of Trade (CBOT) in 1848, the standardised contracts in existence today soon was introduced (Kolb and Rodriguez, 1996). The basic function of these contracts is as an agreement reached between two parts calling for the delivery of an asset at a specified future date and at a specified price (Hirt and Block, 2006). On the organised futures market, three types of agents can all meet their needs. Hedgers can reduce the risk of a certain position by taking an opposite one in the futures market. Speculators can use leverage to bet on the future development of the market, and lastly arbitrageurs can seek riskless profits by dealing on two or more markets simultaneously (Hull, 2001).

As mentioned above, futures markets were originally developed primarily for farming products, with grain being the first product traded on the CBOT. These contracts were soon followed by others, but not until 1978 was the first futures contract on oil introduced, when heating oil contracts became available for trading on the NYMEX. These were to be followed five years later by crude oil contracts, which today are the most liquid forum for oil trading, as well as being the largest volume futures contract on a physical commodity in the world.

4.2.1 Why the Oil Futures Market Matters
When, as we have seen above, a resource is as important for the world economy as oil is, small changes in the price can have large effects on the welfare of people. This

---

knowledge can lead to insecurity amongst market participants, which in turn can lead to less than optimal decisions being made, due to overly defensive strategies being employed. This insecurity is to an extent eliminated by the use of futures contracts. Futures markets have been proven superior to cash markets in determining prices since the price determination process takes place in a central market with open competitive bargaining (Kofi, 1973). This higher efficiency in price setting stabilises the market and reduces insecurity. As has been stated in studies performed by, for example, Bopp and Sitzer (1987) and Garbade and Silber (1983), the primary function of futures markets is to facilitate risk sharing. This allows agents on the market to adjust the risk they face to a preferred level. As a consequence of the sharing of risk, the market also works as a predictor of prices. It is thus evident that the futures market in of great importance to the economy, and that we need to strive for a better understanding of its functions.

4.3 Futures Market Efficiency and Important Concepts

Relevant to mention while discussing market efficiency and the futures market are concepts strongly related to the core issue. One such concept is that of basis risk, which arises in hedging situations when there is a discrepancy in how the asset to be hedged and the futures contract used move together (Haushalter, 2000). The formal statement of basis risk, as stated by Hull (2001), is:

• **Basis = Spot price of asset to be hedged – Futures price of contract used**

The potential basis risk faced by hedgers can for example be due to a difference in the expiration date of the futures contract and the date the asset to be hedged is to be bought or sold. But it can also be due to inefficiency in the futures market, i.e. if the spot and futures prices are not cointegrated (Hull, 2001).

Another important topic related to futures market efficiency is that of arbitrage, i.e. a riskless profit obtained by entering transactions in two or more markets simultaneously (Valdez, 1993). For arbitrage possibilities to be present there need to be inconsistencies in the pricing on the relevant markets. Arbitrage can also be present in the futures market if said market in not functioning efficiently. If, for example, the futures price is above the expected future spot price, a riskless profit can be made by
shorting a futures contract while at the same time buying the underlying asset (Hull, 2001). A particular arbitrage technique occasionally used by market participants is the futures spread. This technique is applied by for example buying a futures contact, while at the same time selling another contract on the same underlying commodity (Kolb and Rodrigues, 1996). The aim is to take advantage of inconsistencies in the price spread between the two positions, while being partially protected by a hedge. In an efficient market with strong cointegration between spot and futures prices, profiting from arbitrage is not possible, since no discrepancies in prices are large enough to offset transaction costs.

As the above section makes evident, there is a major difference in how market participants relate to the concept of efficiency. While the arbitrageur can exist only if there is some form of inefficiency on the market (Valdez, 1993), the hedgers and speculators thrive at an efficient market. Indeed, the entire concept of efficient hedging is based upon the assumption that the market itself is efficient. If it is not, the measure of taking an opposite position to hedge against unbeneificial price movements becomes fundamentally flawed (Hull, 2001). As a consequence of this, an inefficient market will lead to less hedgers participating in the market, which will set of a negative spiral leading to ever lower levels of efficiency.

5. Results of the Study
This section will present the results of the calculations done on the collected data. First, the results of the descriptive analysis will be presented. Following this, the results from the augmented Dickey-Fuller (ADF) test are presented for both crude and heating oil. After this the results of the cointegration tests are presented.

5.1 Descriptive Sample Analysis
As has been mentioned above, a descriptive analysis, i.e. analysing the properties of the sample, familiarises one to the data. The table below presents the results of the analysis performed on the crude and heating oil price time-series.
From the above results it is evident that the variance and standard deviation, measuring variability in the sample, of spot and futures prices is almost identical. This is to be expected since the spot and futures prices react to the same information in similar ways. It is also evident that the variability is significantly higher for the heating oil data than for the crude oil data. The reason for this could be that the price of heating oil, as can be seen in the mean value calculations, is significantly higher than the price of crude oil, due to heating oil being a refined product. Thus identical percentage point changes results in larger nominal changes in the price of heating oil.

The correlation, measuring the direction and the strength of a linear relationship, is close to 1.0 for both the crude and the heating oil data. This indicates that they move closely together, and in the same direction. The covariance, measuring how much spot and futures prices change together, follows the same pattern as the variance, with a higher covariance associated with the heating oil data. The overall picture painted by the descriptive analysis is that the spot and futures prices do share the same characteristics. It is important to note though, that this introductory analysis in no way contains sufficient information to conclude whether or not spot and futures prices are cointegrated, or for that matter whether or not the market is efficient.

### 5.2 Testing for Nonstationarity

The critical values used to test the null hypothesis have been provided by the program STATA. As has been stated earlier, it has been assumed that the data is constant with no trend, and the level of significance has been set at 5 %. The results of the calculations have been used to perform hypothesis tests, in this case concerning the possible nonstationarity of the data. The null hypothesis in the ADF-test states that the data indeed is nonstationary, while the alternative hypothesis states that it is stationary.
5.2.1 Crude Oil

The table below presents the results of the ADF-test performed on the data collected on crude oil.

<table>
<thead>
<tr>
<th></th>
<th>Critical value</th>
<th>Calculated t-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Sign. level .05</td>
<td>lag. value</td>
</tr>
<tr>
<td>ADF-test; spot</td>
<td>-3.41</td>
<td>-2.17</td>
</tr>
<tr>
<td>ADF-test; futures</td>
<td>-3.41</td>
<td>-2.10</td>
</tr>
</tbody>
</table>

Table 2. Results of the ADF-test performed on the crude oil data.

As is evident in the table above, the calculated test statistics, on spot as well as on futures prices, are both above the critical values associated with the significance level analysed. Thus the null hypothesis can be rejected, and the conclusion that the data is stationary can be drawn.

5.2.2 Heating Oil

As with the crude oil, the table below presents the results of the calculations done on the heating oil data.

<table>
<thead>
<tr>
<th></th>
<th>Critical value</th>
<th>Calculated t-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Sign. level .05</td>
<td>lag. value</td>
</tr>
<tr>
<td>ADF-test; spot</td>
<td>-3.41</td>
<td>-4.47</td>
</tr>
<tr>
<td>ADF-test; futures</td>
<td>-3.41</td>
<td>-3.79</td>
</tr>
</tbody>
</table>

Table 3. Results of the ADF-test performed on the heating oil data.

Compared to the data on crude oil, the heating oil data is more ambiguous. The t-values associated with the lagged values are below the critical value, indicating that the null hypothesis can be accepted and that the data is nonstationary. However, the data associated with the differenced lagged values are above the critical value, indicating that the data is stationary. Thus, there is some ambiguity about whether or not the data is nonstationary, and no certain conclusion can be drawn.
5.3 Testing for Cointegration

When testing the data on spot and futures prices for cointegration, the same criteria, and thus the same critical value, as was applied to the ADF-test have been used. The table below presents the critical value along with the calculated t-values for crude oil.

**Results of the cointegration test on crude oil data.**

<table>
<thead>
<tr>
<th>Critical value</th>
<th>Calculated t-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sign. level .05</td>
<td>lag. value</td>
</tr>
<tr>
<td>Crude oil</td>
<td>-3.41</td>
</tr>
</tbody>
</table>

Table 4. Results of the cointegration test performed on the crude oil data.

The t-value associated with the lagged values is well below the critical value, thus indicating that the null hypothesis stating that there is no cointegration should be accepted. The differenced lagged values t-value is also below the critical value, although not with the same margin. The data thus points in the same direction, and the conclusion drawn is that the spot and futures prices on crude oil has not been cointegrated during the studied period, given the significance level considered.

**Results of the cointegration test on heating oil data.**

<table>
<thead>
<tr>
<th>Critical value</th>
<th>Calculated t-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sign. level .05</td>
<td>lag. value</td>
</tr>
<tr>
<td>Heating oil</td>
<td>-3.41</td>
</tr>
</tbody>
</table>

Table 5. Results of the cointegration test performed on the heating oil data.

The table above depicts the calculated t-values of the heating oil time series, compared to the critical value. As the crude oil data, the heating oil data is uniform in that both the lagged values and the differenced lagged values points in the same direction, with both t-values being numerically smaller than the critical value. Thus, the null hypothesis can not be rejected, and it can be concluded that the heating oil time series, like the crude oil time series, are not cointegrated.
6. Conclusions

The finding of the empirical analysis, that the data on crude oil prices are stationary, and that the data on heating oil prices can not be proved nonstationary, can at first glance seem unexpected. It is easy to assume that the oil price should move in a nonstationary fashion. Despite this, the results are less surprising when a deeper analysis is undertaken. The fact that the data during the studied period has primarily been stationary implies that there exists arbitrage opportunities on the market, and are thus in accordance with the arbitrage theory, if we assume that there are no transaction costs and no taxes. Thus the question arises how big these arbitrage opportunities are, and how high transaction costs and taxes agents on the market can bear and still profit from these opportunities. These issues can not be answered by the results of this thesis, but are interesting questions to follow up on in future studies.

The result of the cointegration test on the heating oil data, that the null hypothesis assuming no cointegration can be accepted with a large margin on both lagged and differenced lagged values, implies that the market during the studied period has not functioned in accordance with the weak form of the efficient market hypothesis. This is also in accordance with the finding that the data is stationary and that arbitrage opportunities have been present. The result of the cointegration test on the crude oil data, depicted in the previous section, also concluded that there has been no cointegration between spot and futures prices during the studied period. The t-value associated with the lagged data indicated that the null hypothesis should be accepted with a large margin, while the lagged differenced value data lead to the same conclusion, but with a margin substantially smaller than that on the heating oil data and on the lagged crude oil data. It could be argued that the fact that the heating oil market shows clearer signs of inefficiency is due to the larger volumes traded on the crude oil market. This tends to increase the efficiency of the market, even though these results show that neither market has been efficient enough during the period studied. But it must also be mentioned that this is purely speculation, and that the undertaken study does not contain information sufficient to draw definite conclusions regarding the underlying reasons for the results.

The results found in this thesis that the oil market has not been working efficiently during the studied period, are similar to the findings of Lo, Mamaysky and Wang.
They stated, as has been mentioned above, that existing inefficiencies on markets could make technical analysis profitable when correctly applied. Although their study considered stocks, it clearly pointed out flaws in the efficient market hypothesis, relevant to not only the stock market but to the commodity markets as well. The findings are also in line with Green and Mork’s (1991) conclusion that oil market efficiency had to be rejected for the period 1978 to 1985. The results also seem to agree with the conclusions made by Shiller (2003), that financial and commodity markets tend to be inefficient due to less than rational behaviour of actors on the markets. However, the conclusion that the spot and futures prices have not been cointegrated is in contrast to the results of studies such as the above mentioned study performed by Gülen (1998), where he stated that cointegration could be found on shorter futures contracts during the period 1983 to 1995. Regarding the conclusion that arbitrage opportunities have existed on the market during the studied period, it could be argued that similar to the finding of Odean (1999) mentioned above, transaction costs are in fact sufficiently high to make profitable trading based on trends discovered in historical prices movements impossible, despite the apparent inefficiencies.

A consequence of the results of this study is, as was mentioned above, that arbitrage opportunities have been present during the studied period, even though we can not conclude whether or not it has been possible to profit from these opportunities. The market inefficiency also affects the other agents present on the oil futures market. As has been discussed earlier in the thesis, hedging strategies can not be employed with the same degree of accuracy on an inefficient market as they can on an efficient one. If this condition of inefficiency remains for a substantial period of time, and if it is severe enough, there is a tangible risk that sections of the hedging portion of the market will disappear, possibly along with speculators who also favours efficient markets. The effect of this would be a thinning of the market, with the possible consequence that the market becomes even less efficient, unless an influx of arbitrageurs trades price discrepancies away. It is also the case that no matter what happens to the oil futures market if hedgers and speculators become unwilling to participate, it would be a severe blow to large segments of practitioners and users of the market. For example institutions, individuals and government functions that rely on hedging strategies as a way to reduce their risk, would in a market with severe
inefficiency find great difficulty to continue their operations. This, of course, would have negative consequences not only to those directly affected, but also to the rest of society if the lack of effective risk sharing functions reduces the economic activity.

The fact that the oil futures market have not been functioning efficiently during the studied period, raises the question whether the results are generalisable to different markets and different periods. One needs to be aware that the oil market does, to a certain degree, have a position in modern western society that not many other markets have. This makes it plausible that the results of an oil futures market study does not necessarily transfer to other commodity futures markets such as those of gold, copper and silver. On the other hand, the essential characteristics of the markets are similar, and since the most liquid market is often the most efficient one, and the crude oil futures market is the most liquid futures market in the world, the results might be a cause for concern also for other markets. Regarding time periods, it is worth mentioning that the oil futures market has been shown to have fairly strong seasonal patterns. As a simple example of this one can just think of the market for heating oil in northern countries during the winter, when demand for obvious reasons is stronger than during the summer. These seasonal patterns might result in market efficiency being strong during certain periods of the year, and weak during others. Since this study considers a relatively long time period, such seasonal trends would not be discovered. However, the long time period does allow us to generalise the results when it comes to different time periods, such as a five year period or a decade, in a way that a study focused on seasonal patterns would not.

To summarise, the purpose of this thesis was to study the efficiency of the oil market by checking spot and futures prices for cointegration. The results show that, based on this study, we can not conclude that the time series data have been cointegrated during the studied period. This implies that the market has not behaved in accordance with the weak form of the efficient market hypothesis. This, in turn, implies that agents on the market have been able to earn extraordinary profits without taking on extraordinary risks. It also implies that information contained in historical price movements might have been able to produce extraordinary profits, since weak form efficiency has not been proven. It is hard to tell if this lack of efficiency is due to some intrinsic flaw of the oil market as a whole, or if it is due to some specific
characteristic of the studied period, but that is certainly another interesting issue to analyse in future studies.

6.1 Ideas for Further Studies
As has been mentioned above, the apparent lack of efficiency on the oil futures market raises the question of whether other commodity futures markets share the same efficiency problem. It would be interesting to conduct studies on important markets such as those of gold and copper. As has also been discussed above, the question of potential seasonal patterns affecting the efficiency has not been dealt with by this analysis, and studies conducted on this matter could give valuable insight into how the market functions. Another interesting question for further studies is that of arbitrage opportunities, and whether they are big enough to allow profitable trading when transaction costs and taxes are accounted for.
7. Bibliography


Online Sources


Appendix A

Step-by-step presentation of the calculations

Descriptive analysis:

• Mean value = \( \frac{x_n}{n} \), where \( x_n \) is the sum of all \( x \) and \( n \) is the number of observations.

• Variance: \( \text{Var}(x) = \text{E}((x-\mu)^2) \) where \( \text{E}(x) = \mu \) is the expected value (mean) of variable \( x \).

• Standard Deviation: \( \text{SD}(x) = \sqrt{\text{Var}(x)} \).

• Covariance: \( \text{Cov}(x,y) = \text{E}((x-\mu)(y-v)) \) where expected values are expressed by \( \text{E}(X) = \mu \) and \( \text{E}(Y) = v \), and \( x \) and \( y \) are the variables.

• Correlation: \( p(x,y) = \frac{\text{Cov}(x,y)}{\sigma(x)\sigma(y)} \) where \( p \) is the correlation coefficient, \( x \) and \( y \) the variables, \( \text{Cov}(x,y) \) the covariance of \( x \) and \( y \), and \( \sigma \) the standard deviation.

Stationarity:

• The method used to test the data for stationarity is the augmented Dickey-Fuller test. We assume that the time series are constant with no trend, thus we use a test equation that has a constant term: \( \Delta y_t = \alpha + \gamma y_{t-1} + \sum \alpha_s \Delta y_{t-s} + v_t \). Here, \( \Delta y_t \) represents the change in variable \( y \) (the oil price) from period \( t-1 \) to \( t \). The constant \( \alpha \) represents the intercept while \( v_t \) is the error term. The expression \( \sum \alpha_s \Delta y_{t-s} \) indicates that a lagged difference is added to the equation. It is also assumed that a certain proportion (\( p \)) of the variable \( y_{t-1} \) is contained in the present period’s price \( y_t \). This phenomenon is in the equation accounted for by the inclusion of \( \gamma \), since \( p-1 = \gamma \).

• The data collected on spot and futures prices has been lagged in Excel using the method described below:
<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spot price</td>
<td>Differenced Spot</td>
<td>Lagged Spot</td>
<td>Differenced Lagged Spot</td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>42,16</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>43,96</td>
<td>=A3-A2</td>
<td>=A2</td>
</tr>
<tr>
<td>4</td>
<td>43,41</td>
<td>-0,55</td>
<td>43,96</td>
</tr>
<tr>
<td>5</td>
<td>45,51</td>
<td>2,1</td>
<td>43,41</td>
</tr>
<tr>
<td>6</td>
<td>45,32</td>
<td>-0,19</td>
<td>45,51</td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
<td>45,32</td>
</tr>
</tbody>
</table>

• The lagged data have been used to estimate a regression, the Differenced Spot in this example has been used as the Y-range and the Lagged Spot and the Differenced Lagged Spot are used as the X-range. This results in a Summary Output where t-statistics are provided.

• To test the null hypothesis that the data is nonstationary, the calculated t-value is compared to a critical value that can be found in most econometrics books. If the t-value is numerically larger than the critical value the null hypothesis is rejected.

**Cointegration:**

• The method used to test the data for cointegration is the Engle-Granger test.

• A regression is estimated using the spot prices as the Y-range and the futures prices as the X-range.

• The residuals obtained by the regression are lagged using the method presented in the table above.

• These lagged figures are used to estimate another regression using the same method that we used when testing for stationarity. This results in a Summary Output that provides us with the t-values necessary to test the null hypothesis, which states that the series are not cointegrated.
Appendix B

Data output:

Results of the ADF-test on crude oil data.

<table>
<thead>
<tr>
<th></th>
<th>Critical value</th>
<th>Calculated t-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Sign. level .05</td>
<td>lag. value</td>
</tr>
<tr>
<td>ADF-test; spot</td>
<td>-3.410</td>
<td>-2.17</td>
</tr>
<tr>
<td>ADF-test; futures</td>
<td>-3.410</td>
<td>-2.10</td>
</tr>
</tbody>
</table>

Results of the ADF-test on heating oil data.

<table>
<thead>
<tr>
<th></th>
<th>Critical value</th>
<th>Calculated t-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Sign. level .05</td>
<td>lag. value</td>
</tr>
<tr>
<td>ADF-test; spot</td>
<td>-3.410</td>
<td>-4.47</td>
</tr>
<tr>
<td>ADF-test; futures</td>
<td>-3.410</td>
<td>-3.79</td>
</tr>
</tbody>
</table>

Results of the cointegration test on crude oil data.

<table>
<thead>
<tr>
<th></th>
<th>Critical value</th>
<th>Calculated t-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Sign. level .05</td>
<td>lag. value</td>
</tr>
<tr>
<td>Crude oil</td>
<td>-3.410</td>
<td>-28.94</td>
</tr>
</tbody>
</table>

Results of the cointegration test on heating oil data.

<table>
<thead>
<tr>
<th></th>
<th>Critical value</th>
<th>Calculated t-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Sign. level .05</td>
<td>lag. value</td>
</tr>
<tr>
<td>Heating oil</td>
<td>-3.410</td>
<td>-18.43</td>
</tr>
</tbody>
</table>