



UNIVERSITY OF GOTHENBURG
SCHOOL OF BUSINESS, ECONOMICS AND LAW

Structural Breaks in Mean Reverting Processes

Empirical study of WTI-Brent futures spreads

Alexander Djurberg and Zakarias Svenmyr

Graduate School

Master of Science in Finance
Master Degree Project No. 2012:90
Supervisor: Tayland Mavruk

Abstract

The purpose of this study is to examine the implication of structural breaks in mean reverting processes on the expected return of spread trading. Previous research focuses on the effectiveness of threshold filters in mean-reverting models when deciding trading strategies to exploit arbitrage opportunities within the spread of two highly correlated commodity futures. It is often assumed that high levels of co-integration are persistent in these futures prices, therefore ignoring the risks associated with structural breaks. Conducting an event study, this thesis uses an intensity-based model to measure the risk and return associated with structural breaks and changes in the properties of the spread process. The relationship between the two oil futures WTI and Brent have recently experienced considerable structural changes after a long period of stable relation, making the point of change an interesting event to study. The results aim to show how the changes in the mean levels affect risk and return.

Key Words: Ornstein-Uhlenbeck, Mean Reversion, WTI, Brent, Spread, First-time hitting density, Expected return, Futures

Contents

1	Introduction	3
1.1	Background	3
1.2	Research Objectives	6
1.3	Research Question	6
2	Literature Review	7
3	Theoretical Review	9
3.1	Efficient Markets	10
3.2	The mean reverting process	10
3.3	First-time hitting density	11
3.4	Spread position from hedging	11
3.5	Spread Trading	12
3.6	Relation between assets underlying a spread position	14
3.7	Structural breaks	15
3.8	The event	15
4	Data	16
5	Method	20
5.1	Test of unit root	20
5.2	Spread calculations	20
5.3	Test of co-integration	20
5.4	Analysis of spread characteristics	20
5.5	Calibrating the Ornstein-Uhlenbeck (Vasicek) process	21
5.6	Calculation of expected return using first-time hitting density	22
6	Results	28
7	Discussion	37
8	Conclusion	40
9	References	41

1 Introduction

1.1 Background

Crude oil futures have experienced a surge in the volume of contracts traded on the NYMEX (see Appendix A, Figure 1) partly due to their low correlation with other asset types. Commodity futures provide investors, physical consumers and producers the ability to speculate and hedge positions respectively, with futures contracts becoming increasingly available for a number of commodities, of which oil contracts inherit the highest level of market liquidity¹. A high level of liquidity accompanied by the importance of oil products increases the relevance of studying this particular commodity market. The collapse of OPECs administrated pricing system in 1986 changed the pricing system for oil benchmarks, so-called spot prices of oil, from an area specific pricing process to a market pricing process where traded instruments are the basis for benchmark pricing. In the market pricing process, futures and other derivative contracts converge to the products specific benchmark price when the contract approaches maturity. It is argued that a market based pricing model is more efficient for the purpose of price discovery due to the speed of information incorporated into the price². The pricing of oil is faced by various complexities, whereby the physical commodity and futures prices to a certain extent determine each other. The power of price discovery within the oil market and its different types of oil benchmarks is not reflected by the level of production. Although there are hundreds of different grades of oil globally, but there are only a handful of benchmarks that are used in real purpose of pricing, the main of which are Brent, WTI (West Texas Intermediate), Maya and the Dubai/Omani benchmark; of which the two key oil benchmarks are WTI and Brent.

WTI and Brent futures are traded globally on both the NYMEX (New York Mercantile Exchange) and the ICE (Intercontinental Exchange), based respectively in New York and London. Originally, Brent was solely traded on the ICE due to the proximity of the North Sea, where Brent is extracted, while WTI was traded on the NYMEX. The difference in trading location implied that there were only a few hours during the day from which the trading of both contracts overlapped, which decreased the efficiency of comparative analysis. This problem is now surpassed with both contracts currently traded on the NYMEX and ICE.

The price differential between two contracts are often referred to as a spread³. Prices on similar commodity futures contracts would, according to the efficient market hypothesis be co-integrated and the spread would stay at a level that reflects the differences related to properties and location of the two underlying assets. A position in spreads consists of a long and short position in two

¹The level of liquidity for the spectrum of future contracts on different underlying assets <http://www.traders.com/index.php/sac-magazine/departments/futures-liquidity>

²This information is transferred relatively fast since expectations of prices of the underlying are transferred directly to market participants offering price in futures contracts

³The spread created when using similar commodities from different areas is referred to as location spread. Location spreads are the also the main spread type analyzed in this thesis

contracts simultaneously, returns and losses arise only when the spread level changes and are not dependent on the actual contract price levels.

The WTI-Brent futures spread (location spread) has drawn increasing attention over the past two years, with a consistently exacerbating change in the spread level for the two futures prices, while the role of WTI as the world leading oil price determinant is losing power to Brent. The WTI-Brent futures spread is explicitly modeled by Dempster, Medova et al. (2008) using a mean-reverting model to show that arbitrage opportunities prevail. Earlier studies of spread processes have often shown significant high level of co-integration between WTI and Brent, not experienced post 2010. The spread level between WTI and Brent has shown to be consistently changing during the last two years, thereby making the effect of changing mean levels a non-trivial matter in today's context.

Understanding the dynamics of price differentials that can be assumed to follow a mean reverting process and the risk associated with a spread position, such as the WTI-Brent spread, may present informative results from both a hedging and speculative aspect. This thesis intends to provide a way of measuring the risk associated with spread positions and structural breaks within spread processes. Surprisingly so, spread trading or Spreading has maintained little spotlight within academic literature, though this method of hedging and profit extraction has been to a great extent applied by traders (Fennell 2010). Furthermore, it is constituted by Girma and Paulson (1999) that the study of commodity spread trading is highly complex and is therefore under-researched, which further implicates the interest of the subject. Pricing of commodity futures has been studied by Schwartz (1997) and others, but the price relationship among different commodities has not been studied to the same extent. To better understand the implications of structural breaks for spreads in terms of return and the inherent risk, this thesis delves into the under-researched area of a spread trading model applied to the WTI-Brent spread. Previous research focuses on the existence of arbitrage among spreads from a range of commodities, though risks associated with spread positions are not explicitly shown.

Spread positions are often risky due to the commonly occurring high level of volatility in the spread process. The high level of volatility can generate high levels of return, but can also incur high losses. The expected return from a spread position is zero in the long-run without a trading strategy, but the risk is highly significant for positions with shorter maturities. Hedgers are subjected to the risk associated with volatility since hedging positions are often only taken for a shorter period of time to hedge temporary exposure to changes in asset prices. A measurement for risk associated with spread positions is therefore highly relevant for firms with spread positions. Using models closely associated with the analysis of spreads and credit risk, this thesis implements a framework that will show the return and risk that these spread positions incur.

1.2 Research Objectives

This thesis focuses on the mean-reverting movements in spreads, applying a first-time hitting density approach to calculate the expected return and losses where return from the mean reverting process will be calculated using density and probability approximations from Linetsky (2004). The spread between futures prices on Brent and WTI have experienced structural changes in the price relation and mean level of spread during the last two years, which make it possible to study the influence on expected return in the case of structural breaks. In terms of modeling spreads, threshold filters have been consistently researched, though the risk aspect associated with spread positions is not discussed to the same extent. By using a simple trading strategy, this thesis focuses on the risk and return associated with changes in properties of mean reverting processes.

1.3 Research Question

How does a structural break influence expected return from a spread position, when assuming a Ornstein-Uhlenbeck mean reverting process for spread levels?

Is it possible to predict the outcome of a structural break?

2 Literature Review

To understand the functioning and movements of spreads, it is important to understand the factors that affect prices on the underlying futures. In the case of commodities, futures contracts play an important part in the pricing discovery process of actual commodity prices. Models created to predict movements in futures prices on commodities have evolved considerably over time in order to produce the best description of the pricing process these contracts follow, given the nature of commodities. The fundamental pricing of these claims shown by Gibson (1990), implied that the inclusion of a convenience yield⁴ was vital in pricing the claims precisely, while Cortazar and Schwartz (1994) provide empirical results on the pricing of commodities. It was not until Schwartz (1997) that a multi factor model was used to describe the movements of futures contracts. The futures pricing model presented by Schwartz describes the contract price as a stochastic process comprising of other stochastic dependent variables. Some of the fundamental factors playing a role in the pricing process of futures contracts are the instantaneous interest rates, convenience yields and the spot prices, all described as stochastic variables by Schwartz (1997).

Factors affecting the WTI-Brent spread level are explored by Milonas and Henker (2001), whose findings show that the largest disparities in the spread level arise from the convenience yield and transportation costs, while exhibiting that arbitrage opportunities exist from the spread positions. This arbitrage could, according to Milonas and Henker (2001), only be achieved during a few hours of the trading day due to the fact that WTI was traded on the NYMEX and Brent was traded separately on the ICE. This issue is now resolved with both contracts traded at the same time simultaneously.

Mean reversion in the context of financial products was first introduced by Merton (1971), whereby the dynamics of a portfolio can be demonstrated in a stochastic differential equation. Further on, Vasicek (1977) would interpret the term structure of interest rates introducing the diffusion⁵ (Markov) process to describe the movement of interest rates, which follows the same theoretical underlyings as the Black-Scholes model used in Merton (1971). The use of mean reverting processes would then be implemented by Schwartz (1997) to describe the behavior of future prices and factors affecting commodity future prices.

Factors used by Schwartz to explain the movements in futures prices will subsequently affect spread levels as well. Changes in factors explaining futures prices of Brent and WTI will indeed affect spread levels, but since commodities are more than just financial assets, the price can be influenced by factors connected to the real commodity product. The spread trader therefore needs to consider more factors than those presented by Schwartz (1997). The spread trading concept within academic literature originated from that of Working (1949). The concept surrounds the theory that

⁴Convenience yield is the change in the pricing within a non-arbitrage environment adjusting for the cost of carry

⁵Stochastic process

arbitrage opportunities exist when two closely related products contain disparities in the cost of storage. However, there are other factors that influence arbitrage possibilities such as transportation costs, seasonality, temporary demand/supply convergences and volatility of the underlying cash commodity (Milonas and Henker 2001). Butterworth and Holmes (2002) examine the pricing of spread options⁶, whereby findings indicate there are difficulties in performing profitably due to transaction costs and the inability to liquidate⁷

⁶In this case we are considering an inter-market spread, which means the spread from two non-identical products

⁷Has to find suitable buyer to buy the physical product retained in the contract

3 Theoretical Review

The most common commodity derivative traded are futures due to the liquidity this market has to offer, while deals can be leveraged with margins as high as 80% (Tucker 2000). There are several types of spread positions that can be undertaken:

- Crack Spread: The spread between futures on refined oil and crude oil
 - Location spreads: The spreads between futures with an underlying in different locations
- Intermarket Spread: The simultaneous sale and purchase of the same futures contract in terms of maturity and type on two different exchanges
- Calendar spread/Intramarket/interdelivery Spread: The purchase of a futures contract with a certain maturity and the sale of the same futures contract with a different delivery date

Different types of spreads have different properties in terms of volatility, rate of mean reversion, correlation, stationarity, liquidity etc. But these properties are also dependent on the two types of assets underlying the spread. For the purpose of this thesis, the location spread between Brent and WTI will be studied. Spreads will from now on be referred to as location spread if nothing else is mentioned.

The location spread for WTI and Brent futures can be defined according to Milonas and Henker (2001) as equation (1).

$$SPR_{(t,T)} = FWTI_{(t,T)} - FBrent_{(t,T)} \quad (1)$$

$SPR_{(t,T)}$ is the price for WTI futures at time t with a maturity of T and $FBrent_{(t,T)}$ is the corresponding price for Brent futures. WTI is able to be refined into greater quantities of gasoline and has a higher gravity, thus deeming it superior and more likely to lead the spread relationship which is shown historically on data.

Upon observation of crude oil benchmark prices in Appendix A, Figure 4, it is reasonable to argue that it is WTI that moves away from a long-run relationship, while the relationship between Dubai and Brent remains constant suggesting that the increasing disparity in the WTI-Brent is largely due to WTI. This observation that it is WTI that deviate from market prices is supported by Figure 5 where the spread for WTI-Brent and WTI-Dubai seems to diverge from a long-run mean, while Brent-Dubai seems to maintain its trend.

3.1 Efficient Markets

Two similar commodities should maintain a stable relationship according to Johansen (1991). A widening gap in prices should accrue, for the most part from increased transportation costs. Future spot prices should reflect all information available if the market is efficient. Ma (1989) indicates that energy futures traded on the NYMEX are better suited to forecast oil prices due to the liquidity and ability in reflecting real oil prices. However, Brent futures are considered more relevant in terms of hedging non-US oil, since it reflects the global price of oil (Milonas and Henker 2001). The globalization thesis implies that the spread level of crude oil prices should remain at a stable level (Fattouh 2008). However, studies have shown that oil markets are not as integrated as originally stipulated. Considering the similarities in WTI and Brent with only a marginal difference in sulphur content and gravity⁸ the two should follow the same trend, however over the last 2 years this relationship has changed, presented in futures prices in Figure 2. The relationship between similar commodity products, has in previous literature shown to return to an equilibrium level, implying that an opportunity for arbitrage exists (Milonas and Henker 2001). Even if opportunities for arbitrage do exist, they are not riskless. Girma and Paulson (1999) found that positive and significant arbitrage opportunities in spread trading using historical data from 1984 through to 1994 though convey doubt in its persistence, since arbitrage is quickly removed by market participants. The research concerned in this thesis would therefore concern at least a semistrong form of efficiency assuming that publicly available information is reflected in futures prices.

3.2 The mean reverting process

A problem when studying the mean reverting process of prices on futures in different locations is that prices must have the same time to maturity. Since futures contracts often matures with a time span of one month, i.e. one month between maturity of two futures of the same type, only monthly historical observations can be used when analysing the mean reverting process of futures prices. In the case of modelling spread processes, the time to maturity is not of the same relevance as long as the two underlying assets have the same time to maturity. This increases the number of observations available; one can use daily data instead of monthly when calibrating models.

The futures spread for WTI-Brent in this thesis is assumed to follow an OU (Ornstein-Uhlenbeck) process, proposed by Vasicek (1977) where the spread level follows a mean reverting process to evaluate the term structure of interest rates. However, the first to describe a mean reverting process was Merton (1971) whereby the dynamics of a stochastic process of a portfolio P over time is described by a diffusion process, described in equation (2).

$$dP = \{[w(\alpha - r) + r]P - C\}dt + w\sigma P dZ \quad (2)$$

⁸(American Petroleum Institute) defines gravity as the density of petroleum liquids. Low gravity is defined as heavy crude and vice versa for high gravity (over 30)

In the diffusion model by Merton, variance follows a Wiener process while w represents the speed of mean-reversion, α is the instantaneous rate of return, C is consumption and r is the risk free rate of interest. Vasicek (1977) also assumes that the market price of risk is constant therefore making the spot rate follow a stochastic OU-process described with equation (3), where $\alpha(y)$ is the instantaneous drift rate of the interest rate.

$$dr = \alpha(y)dt + \rho dz \quad (3)$$

Dempster, Medova et al. (2008) used an OU-model to describe the spread movements, see equation (4), where λ represents the rate of mean reversion, y the long-run price process (deviation from long-run mean), t the seasonality, θ the mean level and x the mean reverting spread level.

$$dx_t = \lambda(\theta + \phi(t) + y_t - x_t)dt + \sigma dW \quad (4)$$

3.3 First-time hitting density

First-hitting density models extend to a wide spectrum of financial applications, with Linetsky (2004) applying the use of a mean reversion in OU and CIR models. The implications of mean reverting processes extend over a wide range of financial applications from several forms of stochastic processes, including volatility and for modeling interest rates, credit spreads and commodity convenience yields. Linetsky (2004) provided a mathematical framework, approximations, for probabilities of mean reversion in large portfolios by using rate of mean reversion and volatility that will be included in the method.

3.4 Spread position from hedging

As shown, there are several types of spreads, whereby the most commonly taken spread position for the purpose of hedging is the crack spread. A crack spread position can be undertaken in order to hedge risk associated with the physical commodity that may not be traded on an exchange⁹. Companies that are exposed to the physical price of oil through business operations have an inherent short position when consuming oil products and a long position when producing oil products. To hedge the inherent position, companies turn to the financial market. An oil consuming company will hedge the inherent short position with a long position in the financial market. In the opposite strategy, oil product producers hedge their long position with a short position in the financial market. Not all oil products are available in the form of financial contracts to trade on the financial market so therefore the hedger needs to take a position in contracts with similar characteristics as that of the physical product, forming a crack position. Recently these hedgers have accumulated

⁹For example hedging the price of jet fuel using heating oil futures

the smallest net short position in 6 years, though this is only in light of WTI trading¹⁰. When accounting for Brent into this calculation, there is no decrease in the net short position. The decrease of positions taken in WTI is most probably a result of WTI's lower level of market power and decreased ability to follow market prices. Mismatches (spread positions) in hedging positions can generate losses and high levels of risk. Metallgesellschaft (MG) AG's US oil subsidiary known formally as MG refining and Marketing (MGRM) experienced heavy losses due to a hedging mishap when taking positions in futures with different maturities (calendar spread) to hedge for price changes. The strategy involved speculating on the short and long term relationship of futures prices. The strategy of MGRM to hedge long-dated short positions was to go long in near-month futures contracts. MGRM's large position in open interest meant resulted in losses from margin calls, resulting from other market participants hedging their positions in reaction to the rollovers resulting in contango rather than backwardation¹¹ (Krapels 2001).

3.5 Spread Trading

Spread trading was first introduced by Working (1949) that studied the inter-temporal price relationship resulting from costs of storage. The study, made by Working, implied that traders could profit from abnormal prices. Melamed (1981) reiterates the point that opportunities exist for both hedgers and spread traders when taking positions within spreads. Factors accruing to the price differentials of WTI and Brent include seasonality, transportation costs, convenience yield, supply/demand divergences and volatility of the underlying cash commodity (Milonas and Henker 2001). However, seasonality has a much higher effect on fuel and heating oil than on crude oil, therefore the effects from location and transportation on differences in spreads are more related factors for changes in the location spread between WTI and Brent. Spread levels will not be affected by changes in these factors as long as the changes influence both of the underlying assets to the same extent. The same can be concluded from the demand and supply factors where spread levels are not influenced by changes in demand or supply as long as both assets efficiently face the same demand and supply. A change in an asset's ability to meet the same demand and supply as the other corresponding asset in the spread can dramatically change the spread level. The disability of the assets to face the same demand and supply on the market is one of the main reasons to the changes in spread levels presented in the event study in this thesis.

There are several strategies available that can be applied to create arbitrage opportunities in spreads. Butterworth and Holmes (2002) employ the use of FTSE 100 and FTSE mid 250 contracts to study arbitrage opportunities in inter-market spread trading, involving a simultaneous long and short position respectively in identical futures contracts on two different exchanges. Subsequently for an arbitrage to be realized the spread mispricing relationship has to be predicted for the duration

¹⁰FT alphaville

¹¹Backwardation is the relationship of the Futures price and spot price of the underlying whereby the futures price lies below that of the spot price and vice versa for contango

of the contracts maturity i.e. throughout the life of the contract. However, although opportunities in the spread exist throughout the maturity of the contracts, the profitability of this trading strategy was reduced by the inability of traders to escape positions before adverse movements among the two contracts arise (Butterworth and Holmes 2002). This Arbitrage from a mean reverting process is generated if a position is taken when the spread level has diverged from its long-run mean level. A spread trader will gain arbitrage when the spread reverts towards its long-run mean level. By studying the spread process, the trader estimates the probabilities of convergence given a specific rate of reversion, given a specific volatility and spread level.

However, spread trading is considered to be a risky operation, with highly leveraged trades, though this may be considered both as a positive and a negative attribute. The downside of trading spreads include the transaction costs of holding two positions (Dunis, Laws et al. 2006). Milonas and Henker (2001) on the other hand show that the trading risks are associated with the volatility of the convenience yield. The Spread trading strategies discussed in the literature show arbitrage opportunities often involve different types of threshold filters to limit the risk associated with mean diversion. When taking a spread position at a specific spread level, the trader is exposed to the risk that the spread will diverge from its long-run mean instead of reverting. The risk of diversion increases with a decreasing rate of mean reversion. Losses can be enormous if the spread process is subjected to structural breaks, thereby changing the mean level, volatility or rate of mean reversion. Threshold filters are therefore used to limit the possible losses, where the filter is simply a level at which the position is closed to limit losses. Divergence from the mean level can also increase return. Return from divergence can be generated if a position is taken and the spread reverts to its long-run mean, but thereafter continues in the same direction diverting from its long-run mean level. Therefore, trading strategies need a threshold for closing the position when reverting towards the mean level. Risk within spread trading will therefore be highly connected to the rate of mean reversion, but the risk can be limited with threshold filters.

The characteristics of the spread process are dependent on the characteristics of the underlying assets. Spread processes can differ substantially in characteristics; a comparative example can be the differences in the characteristics of the WTI-Brent spread and the calendar spread in WTI. A visual examination of Figure 3 (location spread WTI-Brent), Appendix A, Figure 2 (calendar spread between two and one month WTI futures) and Appendix A, Figure 3 (calendar spread between four and one month WTI futures) indicates that a lower level of mean reversion exist in the calendar spread compared to the location spread in this case. One can therefore argue that the calendar spread is more sensitive to a decreasing rate of mean reversion, and the risk of trading within the calendar spread in comparison to location spread would therefore probably be higher.

3.6 Relation between assets underlying a spread position

Oil consuming organisations aiming to hedge their inherent short position take a long position in the financial market in a closely related oil commodity and vice versa with an inherent long position in oil. The oil input/output price in an organisation needs to be highly correlated or possibly co-integrated with their financial counterpart contract in order to create a position close to a perfect hedge without risk of uncertainty. The level of co-integration and correlation are dependent on the similarities and characteristics between physical products and traded oil products. Hedgers are limited in the number of products to hedge with, since not every grade of oil is available on the exchange to trade. A mismatch in properties between two oil products generates a lower level of correlation and co-integration, therefore creating a gap that cannot be hedged with futures contracts. The gap is simply the spread position. An example of this predicament, when an organisation is forced to take a position in a spread appears when airlines must hedge jet fuel with other oil products, often heating oil. Airlines in the US or Europe cannot hedge their inherent short position with a long position in jet fuel, since jet fuel is not traded on the NYMEX or ICE. The correlation between jet fuel prices in the NYH (New York Harbour) and futures contracts of WTI/Brent/heating oil traded on NYMEX are presented in Figure 4. As one can expect, the highest correlation is between jet fuel and heating oil. The obvious choice, with correlation in mind, would be to hedge with heating oil if you represent an airline and must choose one of the three contracts. The crack spread position taken when hedging jet fuel with heating oil is presented in Figure 5.

Differences in the properties of quality, location etc. in the underlying assets, increase the level of uncertainty in price relations. The uncertainty and fluctuation from a long-run relationship are used by spread traders to explore arbitrage opportunities. The fundamental assumption when analysing spreads is the existence of a dynamic relationship and a high level of correlation between the underlying assets. The theory of spread trading relies on the existence of co-integration, a long-run relationship. The ideas of variables sharing an equilibrium relation was formalized by Engle and Granger (1987) and the term co-integration was used to explain the co-movements and the long term equilibrium level of two assets. The law of one price implies that prices are co-integrated within markets with equivalent products, as in the markets of commodities. The price may not be the same in different areas due to trading costs, transportation costs etc. but the price relation cannot substantially be mispriced without allowing for arbitrage opportunities (Samuelson 1964) and will therefore be co-integrated.

The theory of co-integration says that a combination of two non-stationary time series is co-integrated if a serial combination of them is stationary. For two assets to be co-integrated they need to consist of a shared trend, and each asset therefore needs to be tested for the unit root to confirm the presence of a trend. If two time series individually have of unit root, but a combination of the series (spread process) do not, there then exists a co-integration relationship.

3.7 Structural breaks

Unexpected changes in an otherwise constant trend over a series of data, is under a general consensus often referred to as a structural break. Structural breaks are often associated with unexpected changes in mean levels. In this thesis, spread levels are assumed to follow a mean reverting process and a rapid change in mean levels will change the value of predetermined factors produced by the mean reverting model. Prices of similar commodities are assumed to be highly co-integrated with a stable long-run equilibrium relationship that represents the mean spread level, which implies high level of stationarity within spreads. A structural break will in this thesis therefore be defined as a period of non-stationary spread levels or lower level of co-integrated futures prices. Structural breaks increase uncertainty and the risk associated with spread positions. The implications of structural breaks are not accounted for in Dempster, Medova et al. (2008), which the event study of changes in the WTI-Brent spread in this thesis intends to show.

3.8 The event

In accordance with the efficient market hypothesis, that is the price of the same two goods should follow the same pattern, the widening location spread has been due to the issues of a WTI crude inventory build-up in Cushing, Oklahoma located in PADD 2 (Appendix A, Figure 6). The lack of pipeline available to transport oil from Cushing to the gulf coast results in the stock pile building up followed by the drop in WTI prices. The spread contrary to market beliefs was not short-lived, with a strong drop in demand for WTI futures and a shift in the pricing of oil to Brent. Shown in Appendix A, Figure 7 the level of production in the mid-west has grown accompanied by a constant level of storage forming the fundamental reason for a structural break in the WTI-Brent relationship.

4 Data

For the study of the spread characteristics, daily historical prices of one month to maturity (roll over prices of contracts closest to maturity) futures are used for both WTI and Brent from NYMEX. The futures prices are expressed as dollar per barrel and prices for WTI and Brent futures was collected from Bloomberg. One month futures contracts are chosen due to their inherent liquidity and their representation of spot prices. Data chosen for the study of spread characteristic stretches over a 20 year period from 1992-03-09 to 2012-03-09, the price development of WTI and Brent futures is presented in Figure 1. In order to portray the significance of structural breaks in the context of results obtained in spread studies by Dempster, Medova et al. (2008) the years post 2010 are important. The two year period, 2010-03-09 to 2012-03-09 represents the period for the event study in this thesis. The divergence of the price relation of WTI and Brent during the period of the event is presented in Figure 2. The location spread for WTI-Brent was conducted by subtracting one month futures prices of Brent from WTI. The historical development for spread levels during the 20 year period from 1992-03-09 to 2012-03-09 is presented in Figure 3.

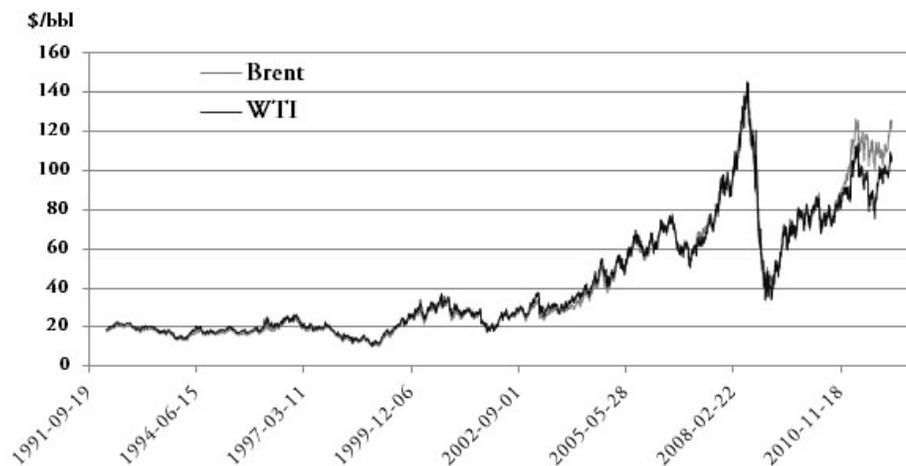


Figure 1: shows the daily roll-over prices closest to maturity for futures contracts of Brent and WTI from 1992-03-09 to 2012-03-09 traded on the NYMEX retrieved from Bloomberg. This period covers the existence of both WTI and Brent together

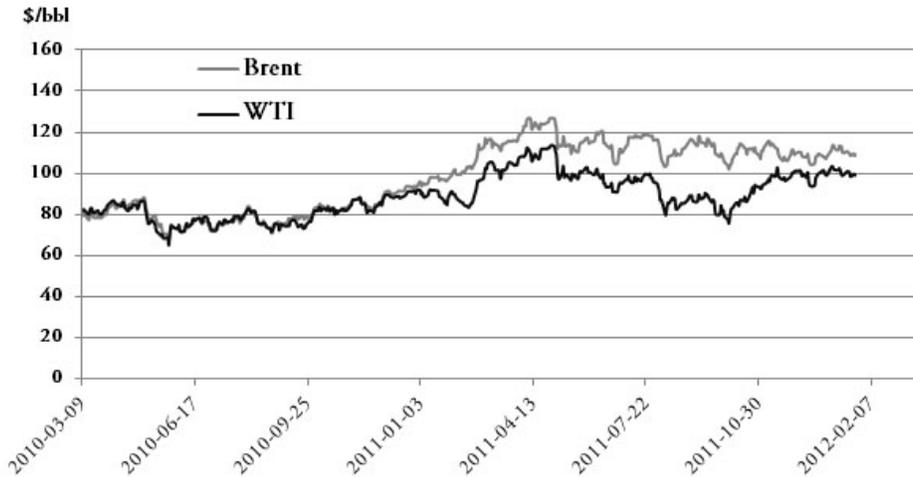


Figure 2: shows the daily roll-over prices closest to maturity for futures contracts of Brent and WTI from 2010-03-09 to 2012-03-09 traded on the NYMEX retrieved from Bloomberg. This period is also chosen to represent the event in this thesis

The dataset is divided into 10 sub-groups with each group representing a period of two years, to better describe the changes of characteristics in the data. Sub group 10 represent the two last years (2010-03-09 to 2012-03-09), which is the period chosen for the event study.

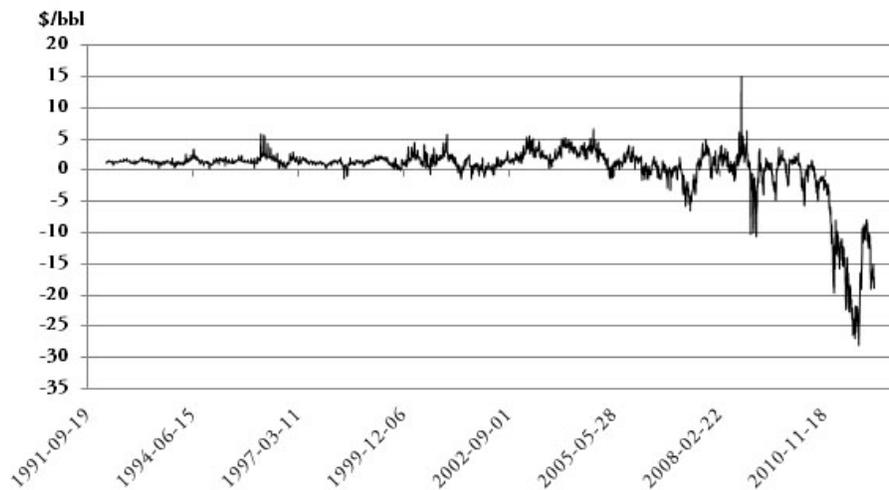


Figure 3: location spread for daily roll-over prices closest to maturity for one month futures contracts of Brent and WTI from 1992-03-09 to 2012-03-09 traded on the NYMEX. The spread is calculated by subtracting the daily price of Brent from prices of WTI. Source: Bloomberg. The negative development of spread values post 2010 is a result of the divergence of WTI and Brent prices that are illustrated in Figure 2

Two calendar spreads was created using historical daily prices for one, two and four month fu-

tures contracts of WTI traded on NYMEX from 1992-03-09 to 2012-03-09, data collected from Bloomberg. The first calendar spread was constructed by subtracting one month futures prices from two month futures prices, see Appendix A, Figure 2. The second calendar spread was constructed by subtracting one month futures prices from four month futures prices, see Appendix A, Figure 3. Prices are expressed as dollar per barrel.

In the example when hedgers takes a position in spreads (crack spread), jet fuel spot prices from NYH(New York Harbor) and one month futures contracts on heating oil are considered. Prices were collected from Datastream and converted from cents per gallon to dollar per barrel. The 30-days roll over correlation between jet fuel and WTI/Brent/Heating oil are presented in Figure 4. The crack spread was created by subtracting historical daily futures prices of heating oil from prices of jet fuel, see Figure 5.

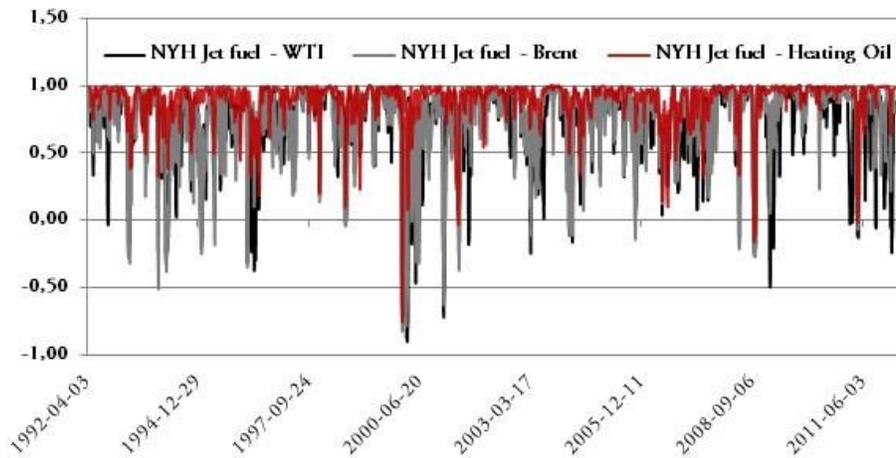


Figure 4: roll over correlations for periods of 30 days between jet fuel prices from NYH New York Harbor and one month futures contracts on WTI/Brent/Heating Oil traded on NYMEX. Data for period 1992-04-03 to 2012-03-09. Source: Datastream

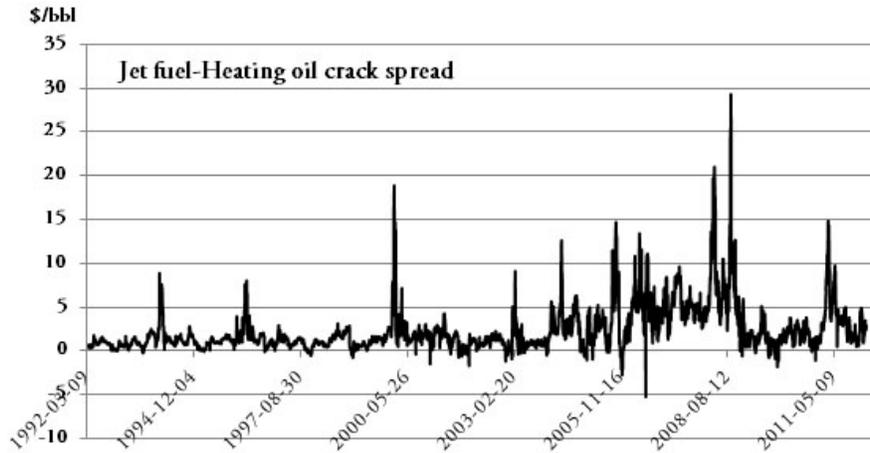


Figure 5: crack spread between daily roll-over prices closest to maturity for one month futures contracts on heating oil traded on the NYMEX and daily jet fuel prices from New York Harbor, from 1992-03-09 to 2012-03-09 The spread is calculated by subtracting the price of jet fuel from one month futures prices of heating oil. Source: Datastream

To show the divergence of WTI from other crude oils, prices for the Dubai/Omani benchmark for the period 1992-03-09 to 2012-03-09 are collected from Datastream (see Appendix A, Figure 4).

5 Method

5.1 Test of unit root

One month futures prices for WTI and Brent are tested for unit roots to detect the presence of trends using the one tail ADF (Augmented Dickey-Fuller)-test. The ADF-test relies on the OLS (Ordinary Least Squares) auto regression in equation (5).

$$S_t - S_{t-1} = \delta_0 + \delta_1 * S_{t-1} + \sum_{i=1}^p \delta_{i+1} (S_{t-1} - S_{t-i-1}) + n_t \quad (5)$$

S_t represents the futures price at time t and S_{t-1} the price lagged one period of the futures price at time $t-1$. δ_1 and $i=0\dots p$, are constants, while n_t is a Gaussian disturbance term. If δ_1 is negative and exceeds the critical value proposed by Dicky and Fuller (1979), the null hypothesis can be rejected, that the time series do not consist of a unit root. The test for unit root is performed using the statistical software package STATA 11.

5.2 Spread calculations

The WTI-Brent spread will be calculated from one month (closest to maturity) daily futures prices by subtracting prices of Brent from WTI. Spread levels for calendar spreads will be calculated by subtracting 1+i month daily futures prices from one month daily futures prices. The spread level for crack spreads between jet fuel and Brent-WTI will be calculated from one month daily futures prices by subtracting prices of Brent/WTI. The spread level for crack spreads between jet fuel and heating oil is calculated by subtracting one month daily futures prices of heating oil from daily market prices of jet oil in NYH (New York Harbour). Unit roots are tested for both in the whole sample period and in the two years sub groups to find if a trend is present in every period.

5.3 Test of co-integration

If a unit root is present in WTI and Brent prices, but the spread is a stationary process, the WTI prices and Brent prices can be assumed to be co-integrated. The presence of co-integration between WTI-Brent will be detected for by using ADF-test for the location spread of WTI-Brent. The ADF-test for spreads will be tested for using STATA 11. Co-integration will be tested for both in the whole sample period and in the two years sub groups to find if the data is stationary in spreads for all periods.

5.4 Analysis of spread characteristics

The spread process will be assumed to follow a mean reverting OU-process. Expected return from spreads will be calculated by applying a simple trading strategy and a first hitting time density approach. A large portfolio approximation of the first-time hitting density of OU-processes,

developed by Linetsky (2004), will be applied in this thesis. By using the first-time hitting density, expected return are calculated by finding the probability that a specific spread level converges to a mean level during a fixed time period. The density of mean reverting processes can be calculated by assuming a stochastic behaviour for the spread level towards a specified mean level in large portfolios. By knowing the probability of ending up in a specific value in a mean reverting OU-process, one can calculate the expected return by assuming an initial spread level. Expected return is calculated using historical data, for every consecutive period of 30 trading days. The historical spread characteristics are used to predict the outcome in an event.

5.5 Calibrating the Ornstein-Uhlenbeck (Vasicek) process

The OU-process is calibrated using OLS (ordinary least squares), a linear regression, to obtain estimates of μ (long run spread level), λ (rate of mean reversion) and σ (volatility parameter). The OLS regression was preferred to other methods such as Maximum likelihood since estimates was similar independent on estimation method. Least squares estimates are computed using Matlab, where the linear relationship in spreads are estimated for historical spread levels of 30 trading days. Each period is calibrated with least squares assuming a general linear relation, see equation (6), in the WTI-Brent spread. The relationship among spread levels in time can easier be expressed using equation (7).

$$(P_{i+1,WTI} - P_{i+1,Brent}) = a + b(P_{i,WTI} - P_{i,Brent}) + \varepsilon \quad (6)$$

$$S_{i+1} = a + b(S_i) + \varepsilon \quad (7)$$

Notation P in equation (6) represents WTI and Brent prices at time i and $i+1$. S_i in equation (7) represents the spread level at time i , while S_{i+1} represents the lagged spread level, a and b represent the linear relationship between spread levels among two periods of time, which in this case is one trading day. OLS regression is used to find a and b , the relationship of spread levels. The OLS estimates and residuals are computed in Matlab using equations (8) and (9), the code for which is in appendix B.

$$a = \frac{\sum_{i=1}^n S_i - b \sum_{i=1}^n S_{i-1}}{n} \quad b = \frac{\sum_{i=1}^n S_i S_{i-1} - \frac{\sum_{i=1}^n S_i \sum_{i=1}^n S_{i-1}}{n}}{\sum_{i=1}^n S_{i-1}^2 - \frac{(\sum_{i=1}^n S_{i-1})^2}{n}} \quad (8)$$

$$sd(\varepsilon) = \frac{\sum_{i=1}^n S_i^2 - \frac{(\sum_{i=1}^n S_i)^2}{n} - b \sum_{i=1}^n S_i S_{i-1} - \frac{\sum_{i=1}^n S_i \sum_{i=1}^n S_{i-1}}{n}}{n(n-1)} \quad (9)$$

2)

Using the parameters from the linear regression, λ , μ and σ can be defined for the OU-model, equation (10).

$$dS_t = \lambda(\mu - S_t)dt + \sigma dW_t \quad (10)$$

The rate of mean reversion parameter λ is computed using the negative of a fraction with the log of the numerator (least squares relation coefficient b), see equation (11). δ is the number of trading days (in this case 30 trading days), to obtain a value that represents the daily rate of mean reversion.

$$\lambda = -\frac{\ln b}{\delta} \quad (11)$$

The long-run mean calibration, equation (12), is a fraction of the intercept a over the 1 minus the relationship coefficient b .

$$\mu = \frac{a}{1 - b} \quad (12)$$

The corresponding volatility to be implemented in the OU-process is calculated using equation (13) and standard error of residuals.

$$\sigma = sd(\varepsilon) \frac{\sqrt{-2 \ln b}}{\delta(1 - b^2)} \quad (13)$$

The description of the OU-model in equation (14) is created by combining equation (10), (11), (12) and (13).

$$dS_t = \left(-\frac{\ln b}{\delta} \frac{a}{1 - b} - S_t \right) dt + sd(\varepsilon) \frac{\sqrt{-2 \ln b}}{\delta(1 - b^2)} dW_t \quad (14)$$

5.6 Calculation of expected return using first-time hitting density

Total expected return, denoted $E[r]$, from a spread position is calculated for a fixed time period T using a simple trading strategy. To follow the reasoning behind the calculation of expected return, see Figure 6. The spread level will be assumed to be a distance of x at the initial time point $\tau = 0$. Spread level x is 2σ to maintain the same probability of initial distance from mean level μ in every time period studied. The spread position is closed if the spread converges to a specified mean level μ during any time point τ before the fixed time T . If the spread level does not converge at any τ

until time T , the position is closed at time point T and return is calculated from spread level y

at time T . The spread levels are assumed to follow a stochastic process moving towards a mean level, and by assuming a large portfolio one can calculate the density and probability of converting before T by using first-time hitting density of Linetsky (2004). 30 days trading period was chosen due to the high level of liquidity within futures contracts with shorter time to maturity, so T will be set to 30.

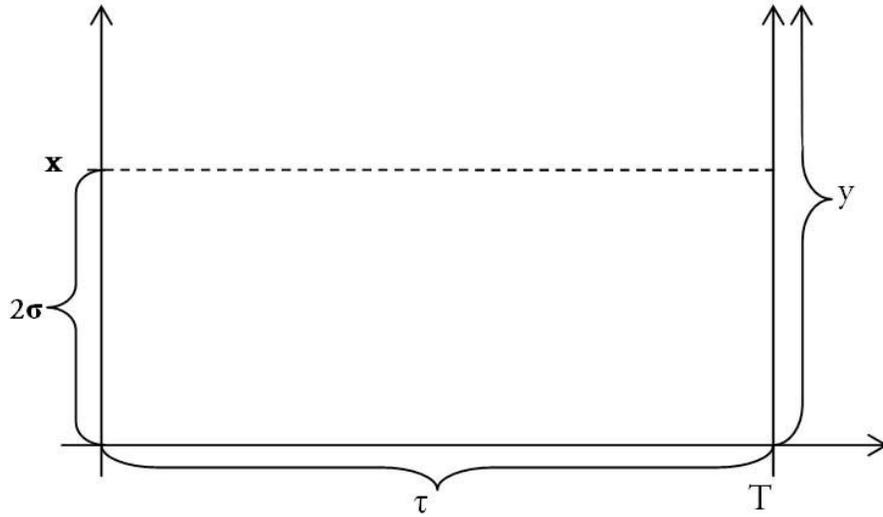


Figure 6: Illustration of trading strategy; expected return using first-time hitting density

$$E[r] = E[r_c + r_{nc}] \quad (15)$$

$$E[r_{nc}] = E[r_{ncA} + r_{ncB}] \rightarrow E[r] = E[r_c + (r_{ncA} + r_{ncB})] \quad (16)$$

The $E[r]$, using first hitting density can be described as expected return from converging spread levels, $E[r_c]$, and non converging spread levels, $E[r_{nc}]$, see equation(15). Expected return from non convergence can also be described as the sum of two different scenarios, $E[r_{ncA}]$ and $E[r_{ncB}]$, see equation (16). The total expected return, $E[r]$, can therefore be described as expected return from three different scenarios ($E[r_c]$, $E[r_{ncA}]$ and $E[r_{ncB}]$), described as:

First scenario, the spread level x converges and hits the mean level μ at any time point $\tau < T$. The expected return $E[r_c]$ is the product of the spread level x and the probability of convergence, P_c before time T , see equation (23). Increasing rate of mean reversion will increase expected return from convergence and decreasing rate of mean reversion will decrease expected return. The $E[r_c]$ is limited to 2σ and can only take on positive values.

Second scenario, the spread level x does not converge to the mean level μ at any time point $\tau < T$ generating expected return $E[r_{ncA}]$. A spread level failing to convert to μ will instead take the value of $y > \mu$. The initial spread level x can converge towards μ and take a value of $x > y$, but will not be included in $E[r_{ncA}]$ in the case of hitting mean level μ before T . The $E[r_{ncA}]$ is calculated as the integral (with respect to y) of the payoff $x-y$ and the probability of arriving at each specific y . The $E[r_{ncA}]$ from spread levels at time T with $y < x$ will be positive, while $E[r_{ncA}]$ from $y > x$ will be negative. Therefore can $E[r_{ncA}]$ be both positive and negative dependent on the level of mean reversion. Increasing rate of mean reversion will increase $E[r_{ncA}]$ and decreasing rate of mean reversion will decrease $E[r_{ncA}]$. The positive $E[r_{ncA}]$ is limited by the level of x , but the negative $E[r_{ncA}]$ does not have any limitations except for the stochastic properties of the spread process.

Third scenario, there is the probability that the spread level converges, later reverting, taking on a value of y . The convergence to mean level μ is already captured in $E[r_c]$ and return from this behaviour is also captured in $E[r_{ncA}]$ because of the taken value of y at time T . The extra amount of $E[r_{ncB}]$, as a result of the taken value y , must therefore be subtracted from $E[r_{ncA}]$ creating $E[r_{nc}]$.

The first-time hitting density for convergence in OU-processes can be described by equation (17). The notations when calculating the first-time hitting density of the three different scenarios will follow those of proposition 2 by Linetsky (2004), except that Linetsky uses notation κ instead of λ for the rate of mean reversion. The first-time hitting density of convergence (the first scenario) will be denoted as a integral of P_c , first-time hitting density of non convergence will be denoted as a integral of P_{ncA} in the second scenario and as an integral of P_{ncB} in the third scenario.

$$f_{T \rightarrow 0}(t) = \sum_{n=1}^{\infty} c_n \lambda_n e^{-\lambda_n t} \quad (17)$$

The WTI-Brent spread process is assumed to have asymmetric mean reverting properties, i.e. the same properties independent of the spread level being above or below mean level μ . Expected return are therefore calculated using only the down hitting density calculation by Linetsky (2004), calculation for a process reverting downwards to its mean level. To calculate the first-time hitting density, eigenvalues λ_n and coefficients c_n needed to calculate following that of Linetsky (2004), see equation (18).

$$k_n = n - \frac{1}{4} + \frac{\bar{y}^2}{\pi^2} + \frac{\bar{y} \sqrt{2}}{\pi} \quad n - \frac{1}{4} + \frac{\bar{y}^2}{2\pi^2}$$

$$\lambda_n = \lambda \sqrt{2k_n - 1} \quad (18)$$

$$c_n = \frac{(-1)^{n+1} 2^{\frac{\sqrt{k_n}}{2}}}{(2k_n - 1)^{\frac{1}{2}} k_n - 2^{\frac{1}{2}} \bar{y}} e^{\frac{1}{4}(\bar{x}^2 - \bar{y}^2)} \cos \bar{x} \sqrt{2k_n} - \pi \lambda_n + \frac{\pi}{4}$$

Calculations for normalized mean values of x and y respectively are represented in equations (19) and (20).

$$\bar{x} = \frac{\sqrt{2\lambda}}{\sigma} (x - \mu) \quad (19)$$

$$\bar{y} = \frac{\sqrt{2\lambda}}{\sigma} (y - \mu) \quad (20)$$

The first-time hitting density of convergence can be calculated by taking the integral of equation (17) and letting t change from zero to T . In large portfolios, large n , equation (17) converges to equation (21) and the probability of convergence, $P_c(T)$, are not an integral of time t , but of the length of the trading period T . The value of n is found by the relationship presented in equation (22), of which the matlab code is presented in appendix B.

$$P_c(T) = \sum_{n=1}^{\infty} c_n (1 - e^{-\lambda_n T}) \quad (21)$$

$$c_n \lambda_n e^{-\lambda_n T} \sim A e^{Bn} \quad A = \frac{2\lambda}{\pi} e^{0,25(\bar{x}^2 - \bar{y}^2)} \text{ and } B = 2\lambda \quad (22)$$

To find the $E[r_c]$, the initial spread level is simply multiplied with the probability of convergence see equation (23).

$$E[r_c] = x P_c(T) = x \sum_{n=1}^{\infty} c_n (1 - e^{-\lambda_n T}) \quad (23)$$

The probability of non-convergence in the first scenario, i.e. the probability that the spread level will take on a value of y , is calculated using equation (24).

$$P_{nc.A}(y; x, T) = \frac{1}{\sigma_s(T)} e^{-\frac{1}{2} \frac{(y - \mu_s(x, T))^2}{\sigma_s(T)^2}} \quad (24)$$

$$\mu_s(x, T) = x e^{-\lambda T} + \mu (1 - e^{-\lambda T}) \text{ and } \sigma_s(T) = \frac{\sigma^2}{2\lambda} (1 - e^{-2\lambda T})$$

Calculations of the return when not converting $E[r_{ncA}]$ is a little more complex since the expected return must be calculated for every possible value of $x-y$. $E[r_{ncA}]$ when the spread level takes on a value of y , is the P_c multiplied with return for every given $x-y$. The $E[r_{ncA}]$ will not be a function of time since the spread level is y at time T when the position is forced to be closed. The $E[r_{ncA}]$ is calculated as shown in equation (25). The integral for y will change from zero (since the position is closed when converting to μ) to infinity (since no upper trading filter is applied). Infinity will be set to 10^5 in Matlab.

$$E[r_{nc.A}] = \int_0^{\infty} (x - y) P_{nc.A}(y; x, T) dy \quad (25)$$

The $E[r_{ncB}]$ depend on both the probability of convergence to τ at any time period $t < T$ and the probability that the spread level takes a value of y . The $E[r_{ncB}]$ is calculated using equation (26), equivalent to equation (27). The $E[r_{nc}]$ is calculated by subtracting $E[r_{ncB}]$ from $E[r_{ncA}]$, which can be expressed as a combination of equation (25) and (27) when in equation (28).

$$E[r_{nc.B}] = \int_0^{\infty} (x - y) \int_0^T P_c(t) * P_{nc.A}(y; 0, T - t) dt dy \quad (26)$$

$$E[r_{nc.B}] = \int_0^{\infty} (x - y) \int_0^T \sum_{n=1}^{\infty} c_n \lambda_n e^{-\lambda_n t} P_{nc.A}(y; 0, T - t) dt dy \quad (27)$$

$$E[r_{nc}] = \int_0^{\infty} (x - y) \int_0^T \sum_{n=1}^{\infty} c_n \lambda_n e^{-\lambda_n t} P_{nc.A}(y; 0, T - t) dt dy \quad (28)$$

Total expected return, $E[r]$, is calculated using Matlab (code in appendix B). The calculation for expected return will not be possible to calculate when using T , since matlab only handle real numbers. Equation (24) will not be a real number then $t=T$ in the integral. To escape this problem, T will be set to approximately $T - (1/\text{infinity})$. $E[r]$ as function of λ and σ is visualized in

Figure 7.

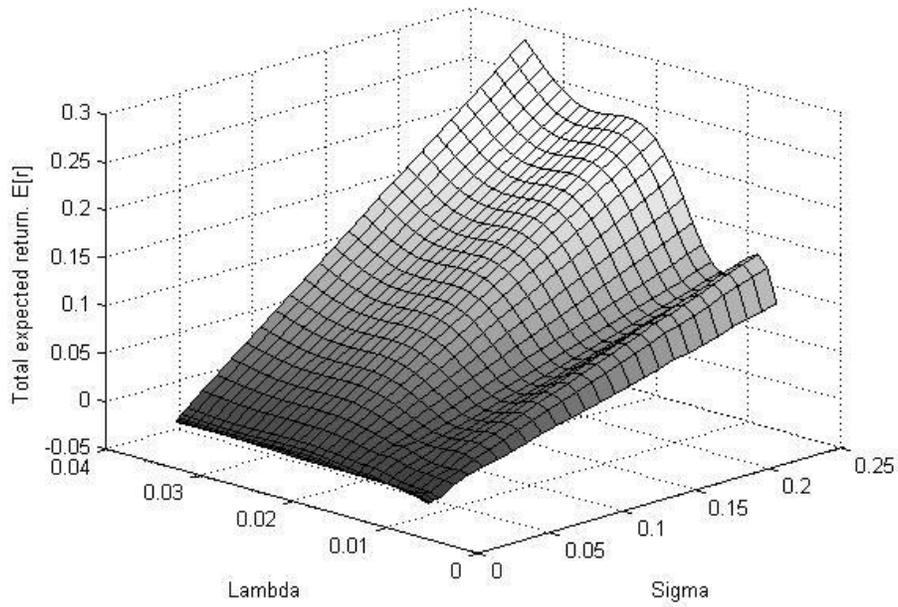


Figure 7: $E[r]$ as a function of λ and σ calculated using first-time hitting density

$E[r]$ and the characteristics of the spread processes is calculated using 5220 observations of daily spread levels from 1992-03-09 to 2012-03-09. The data of spread levels are divided in periods of 30 trading days, 157 periods of which 17 observations in the event study.

6 Results

A summary of mean, max, min and standard deviation for Brent, WTI and the WTI-Brent spread of one month futures contracts for period 1 to 10, 1 to 9 and for each two year period is presented in Appendix A, Table 1. Standard deviation increased substantially for both WTI and Brent prices during period 9 (increased from 11.791 in period 8 to 27.984 in period 9 for Brent and from 12.695 to 28.872 for WTI). Despite the change within WTI and Brent prices in period 9, the spread level remained steady with a standard deviation increase from 1.982 in period 8 to 2.418 in period 9. Period 10 (event period) saw the WTI and Brent standard deviation in prices decrease relative to period 9, from 27.984 to 16.440 for Brent and 28.872 to 10.565 for WTI. In contrast, standard deviation in the spread increased from 2.418 in period 9 to 8.293 in period 10. The spread mean, pre-event is 1.256.

Period	Correlation
t	ρ
1-10	0.992
1-9	0.998
1	0.996
2	0.949
3	0.959
4	0.994
5	0.971
6	0.949
7	0.981
8	0.99
9	0.997
10	0.899

Table 1: correlation in each 2 year period

Observing Table 1, the WTI-Brent correlation is significantly high through all periods though period (10) shows a relatively lower level of correlation.

The result when testing for unit roots using the ADF-test on Brent and WTI prices are presented in Table 2. The null hypothesis cannot be rejected at any level of significance except for period 2 for prices of Brent and WTI since the test statistics are not lower than the critical value for any period. The ADF-test on first differences for prices reject the null hypothesis in all periods, since the test statistics are lower than critical value at the 10, 5 and 1% level.

The test for co-integration between Brent and WTI, or stationarity of spread levels, using the ADF-test is presented in Table 3. All periods are stationary at a 10% level except period 8 and 10 where no trend is included in the regression. However, period 8 can almost be classified as stationary and co-integrated at a 10% level with a p-value of 0,1006. When allowing for trend in the regression,

Period	Prices in levels		Prices in first-difference		critical values		
	Brent	WTI	Brent	WTI	1%	5%	10%
1-9	1.055	-0.386	-58.565* *	-58.262* *	-3.43	-2.86	-2.57
1-10	0.565	-0.634	-55.520* *	-55.307* *	-3.43	-2.86	-2.57
1	-0.525	-0.824	-18.498* *	-19.373* *	-3.43	-2.86	-2.57
2	-3.932*	-3.221*	-16.455* *	-15.477* *	-3.43	-2.86	-2.57
3	-0.705	-1.674	-16.171* *	-16.410* *	-3.43	-2.86	-2.57
4	-0.484	-0.095	-22.587* *	-19.886* *	-3.43	-2.86	-2.57
5	-2.530	-2.341	-17.525* *	-17.097* *	-3.43	-2.86	-2.57
6	-1.550	-1.381	-18.262* *	-18.067* *	-3.43	-2.86	-2.57
7	-0.712	-0.838	-11.163* *	-10.475* *	-3.43	-2.86	-2.57
8	0.238	0.146	-25.741* *	-25.752* *	-3.43	-2.86	-2.57
9	-1.001	-1.142	-18.039* *	-18.069* *	-3.43	-2.86	-2.57
10	-0.780	-0.771	-18.514* *	-18.345* *	-3.43	-2.86	-2.57

Table 2: Test statistics and critical values for ADF-test, 1 period is 2 years

all periods are stationary except for period 10 which has a p-value of 0,2618. Every period except period 10 indicates high levels of co-integration between Brent and WTI.

Period	Obs	Test Statistic	P-value	
			No trend	Trend
1-9	4697	-2.582	0.0967	0.0049
1-10	5220	-10.545	0.0000	0.0000
1	523	-5.606	0.0000	0.0000
2	524	-6.671	0.0000	0.0000
3	521	-5.895	0.0000	0.0000
4	524	-4.7200	0.0001	0.0000
5	523	-4.677	0.0001	0.0000
6	522	-4.152	0.0008	0.0003
7	523	-3.523	0.0074	0.0013
8	523	-2.564	0.1006	0.0387
9	522	-5.583	0.0000	0.0000
10	524	-0.638	0.8621	0.2618

Table 3: ADF test for spreads with and without trend

Calibration of the OU-model resulted in estimates of μ , λ and σ for each of the 30-day trading periods. The estimated values of μ are presented in Figure 8, λ in Figure 9 and σ in Figure 10.

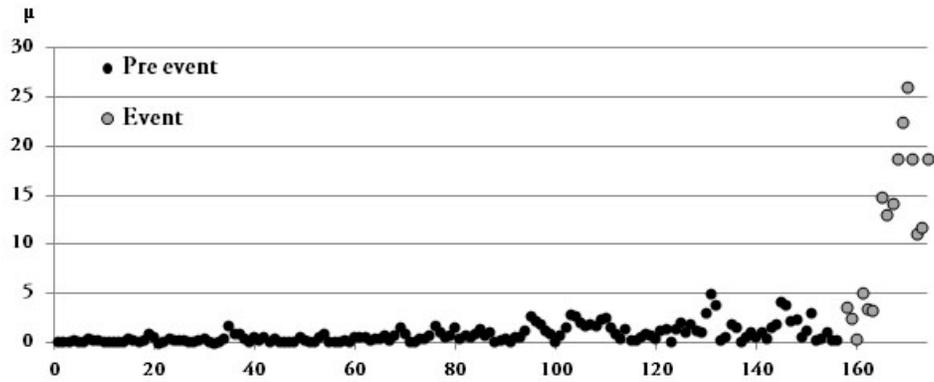


Figure 8: μ (long run spread level) for each 30-days trading period from calibrating the OU-model. Values of μ are adjusted to a mean of zero by subtracting the mean value of the whole period (1.256) from each 30 trading day μ . The Figure displays the absolute values of μ .

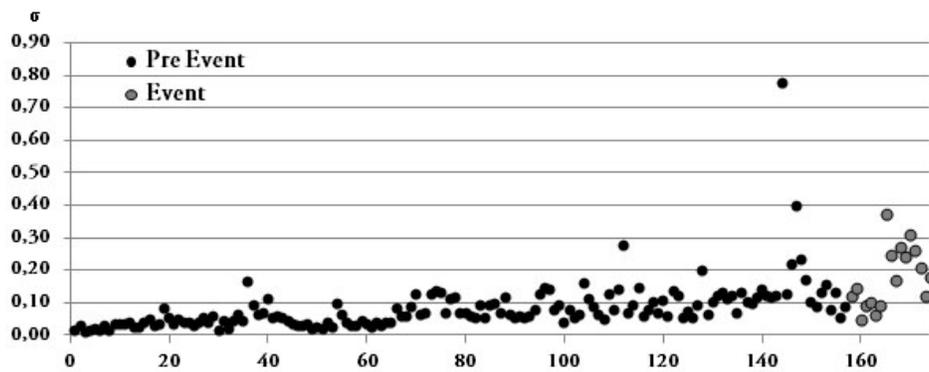


Figure 9: λ (rate of mean reversion) for each 30-days trading period from calibrating the OU-model

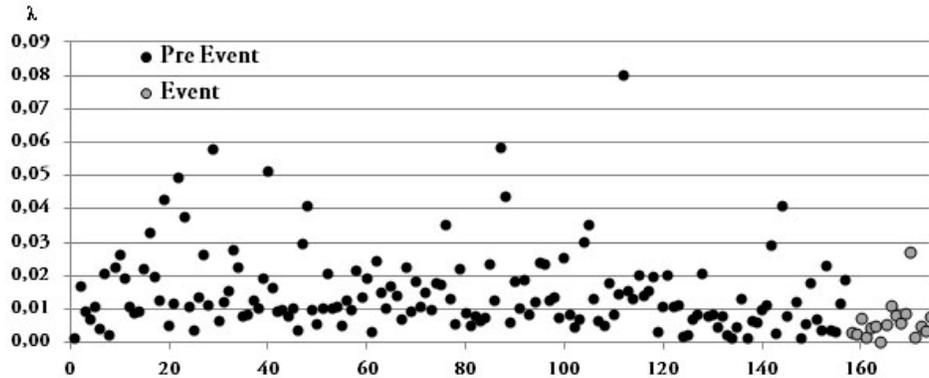


Figure 10: μ (long run spread level) for each 30-days trading period from calibrating the OU-model. Values of μ are adjusted to a mean of zero by subtracting the mean value of the whole period (1.256) from each 30 trading day μ . The Figure displays the absolute values of μ .

Histograms for both pre-event and event estimated λ are presented in Figure 11. 95% of all pre-event estimates of λ take on a value between 0-0.04, while during the event 95% of all estimates are located between 0-0.01. The σ estimates consist of a significant trend over time. To get a better insight of the distribution, the estimates of σ can be de-trended. Pre-event, 95% of all estimates of de-trended σ take on a value between 0.103-0.054, while during the event 95% of all estimates are located between -0.1-0.2 . The histogram for all σ are presented in Figure 12.

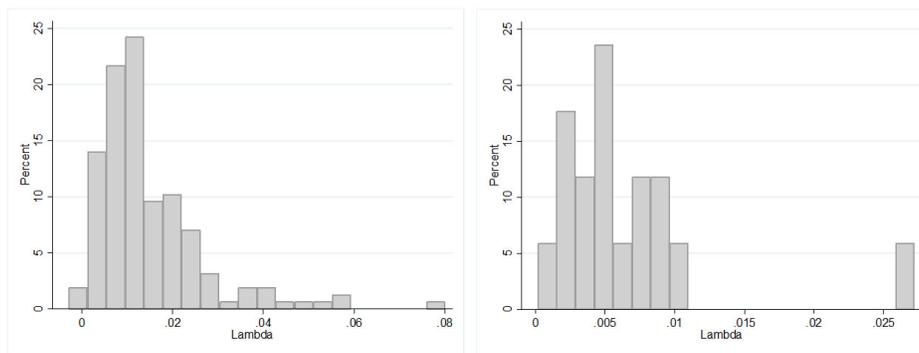


Figure 11: histogram for λ . The histogram to the left represent pre-event estimated λ , while the histogram to the right represent estimated λ for the event.

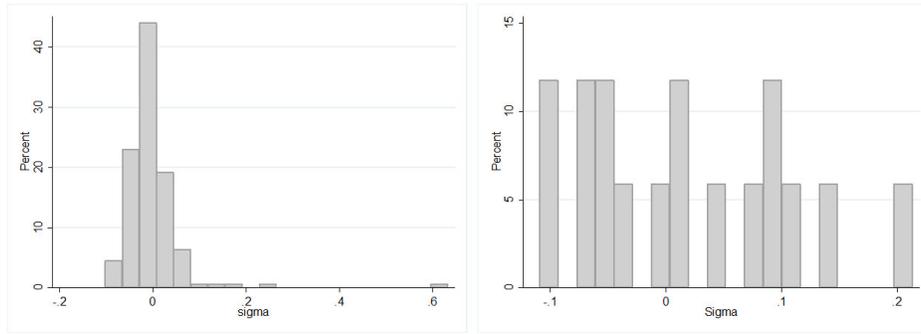


Figure 12: histogram for σ . The histogram to the left represent pre-event estimated σ , while the histogram to the right represent estimated σ for the event.

By using STATA, it was possible to find a significant relationship between σ , λ and μ using historical data for the pre-event period. Observing Figure 13, it can be shown that there is an inverse relationship between λ and σ with the increased distance (\$) from the long-run pre-event μ (\$1.256). This linear relationship can be used to predict the outcome of a structural break as in the event chosen for this thesis. A comparison between the predicted values and the real values of λ and σ in the event is presented in Figure 18.

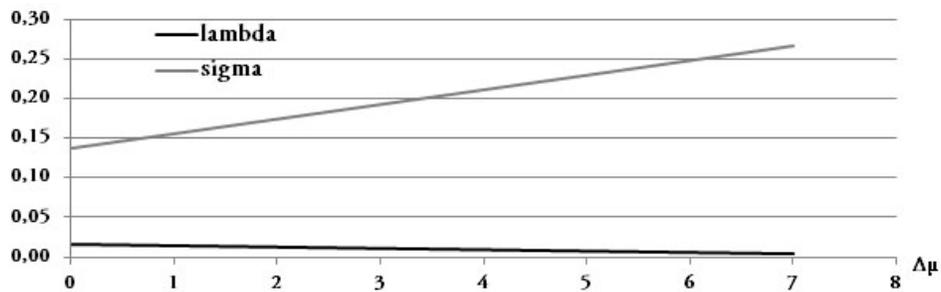


Figure 13: the linear relationship between λ , σ and μ , expressed as functions of changes in the mean level, $\Delta\mu$

OU-model parameters for the last 30-day period before the event are; $\lambda=0.0185$, $\mu=1.4794$ and $\sigma=0.0887$. By fixing σ to 0.0887 and changing λ , $E[r]$ is calculated using the first-time hitting approach. The outcomes of $E[r]$, $E[r_c]$ and $E[r_{nc}]$, are presented in Figure 14. Changes in expected return when fixing λ to 0.0185 are presented in Figure 15.

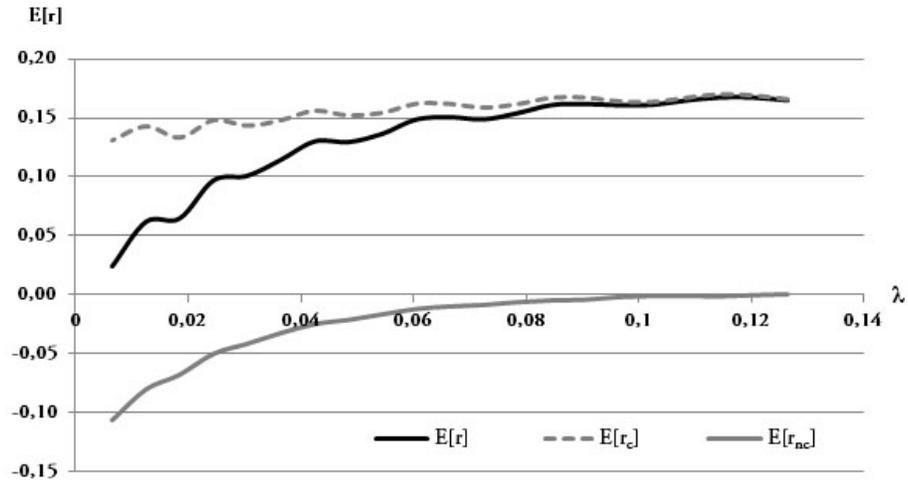


Figure 14: $E[r]$ when fixing σ at 0.0887 and changing λ

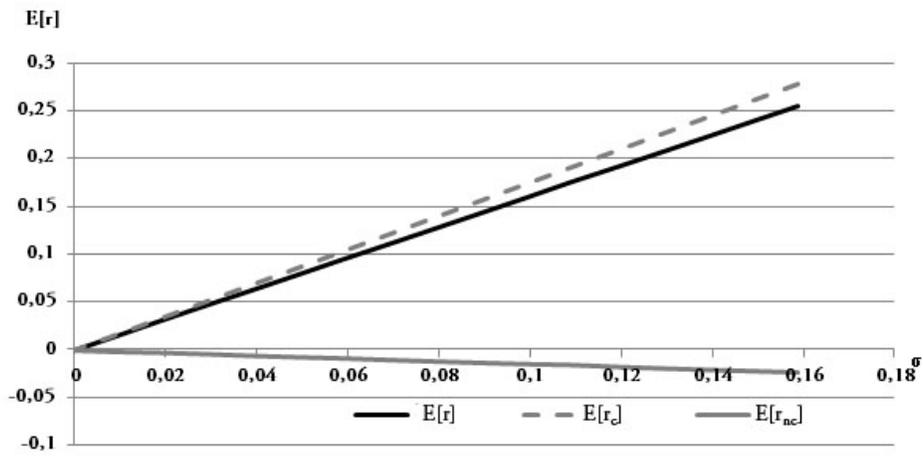


Figure 15: $E[r]$ when fixing λ to 0.0185 and changing σ

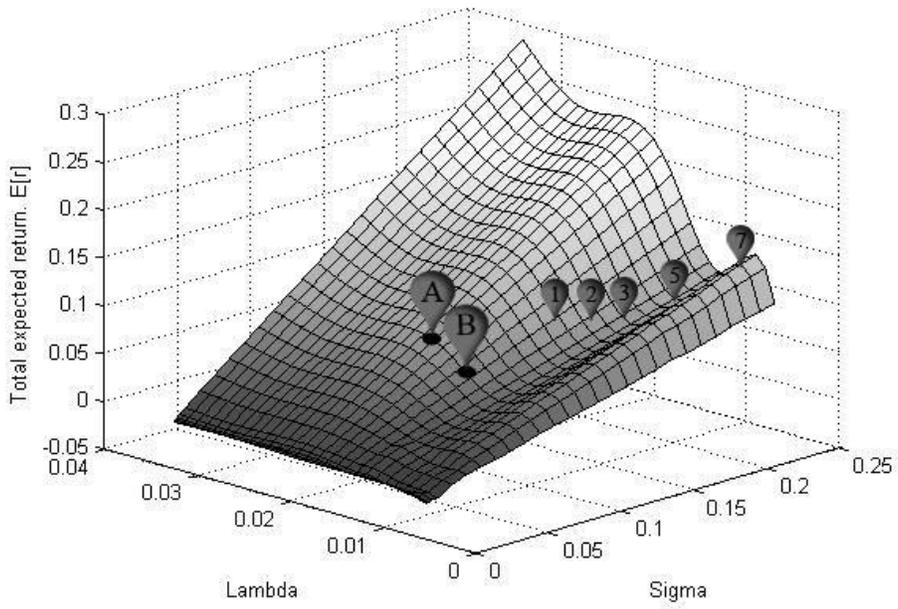


Figure 16: $E[r]$ (\$) using first-time hitting density

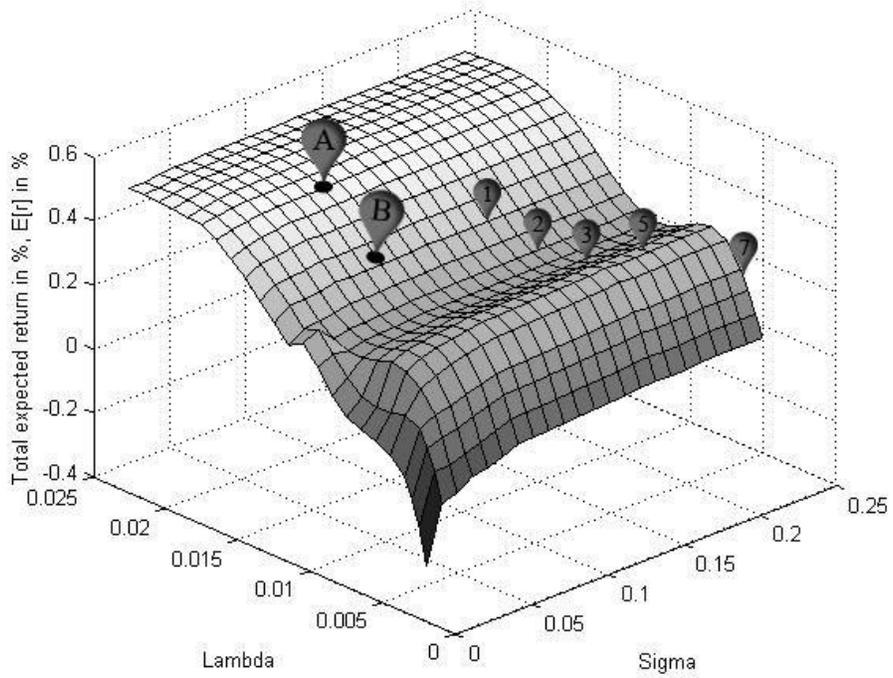


Figure 17: $E[r]$ as a percentage using first-time hitting density

The $E[r]$ in dollars for a trading period is visualized in Figure 16 and marker A represents the $E[r]$ for the last 30 trading day period in the pre-event, which is realized at 0.0533. Marker A in Figure 17 displays the $E[r]$ as a percentage for last 30 trading day period in the pre-event is 8.24%. Marker B represents the $E[r]$ for the average λ and σ in the pre event period. Marker B does not display the actual return in the pre-event since the distribution of λ and σ is not considered. The average $E[r]$ pre-event using the average λ and σ is 5.28% while the actual return was 27.02%, see Table 4.

Marker A in Figure 16 represents the $E[r]$ for the last 30 trading day period in the pre-event, which is realized at 0.0533. Figure 17 displays the $E[r]$ as an percentage and marker A represents the $E[r]$ for the last 30 trading day period in the pre-event, realized at 8.24%. Marker B represent the $E[r]$ for the average λ and σ in the pre-event. Marker B do not display the actual return in the pre-event since the distribution of λ and σ is not considered. The average $E[r]$, pre-event, using the average λ and σ was 5.28% while the actual return was 27.02%, see Table 4. Marker 1, 2, 3, 5 and 7 in Figure 16 and Figure 17 represents estimations of λ and σ for 1, 2, 3, 5 and 7 dollars distance from the pre-event long run mean level \$1.256 using the relationship between λ and σ visualized in Figure 13. Predicted values for λ and σ in case of a structural break are presented in Figure 18 together with real λ and σ in the event. Only λ and σ for μ smaller than 9 are presented.

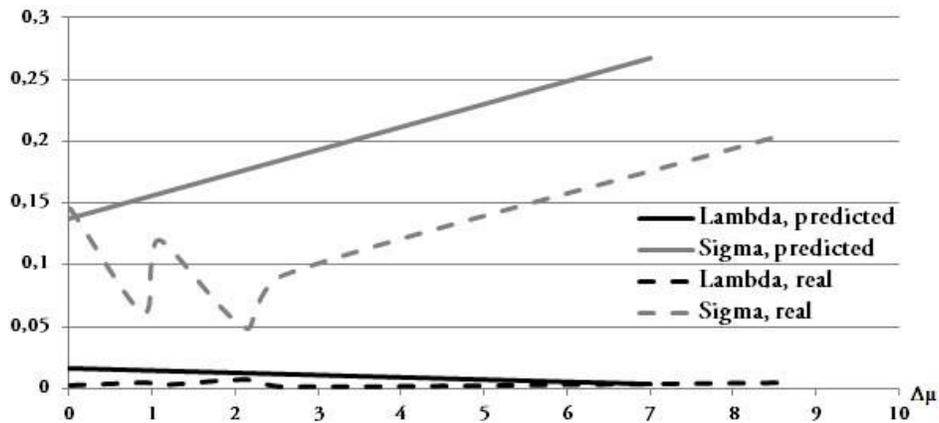


Figure 18: predicted and real values of λ and σ for the event

	Lambda λ	Sigma σ	E[r]	%r	E[r [*]]
WTI-Brent					
Mean, Pre-event	0.0149	0.0826	0.0533	27.02%	5.28%
Mean, Event	0.0064	0.1793	0.0531	1.74%	8.98%
WTI2-WTI1					
Mean	0.0092	0.1022	0.0385	1.85%	4.96%
WTI4-WTI1					
Mean	0.0089	0.0545	0.0217	9.2%	2.48%

Table 4: Return with different spreads * is return calculated with mean values of σ and λ

Average λ for the jet fuel-heating oil crack spread was 0.0092 and the σ was 0.1022 during the period 1992-03-09 to 2012-03-09. The average E[r] from the Jet fuel-heating oil spread was 1.85%, see Table 4, when using the same simple trading strategy as in the case of location spread.

Average λ for the WTI (two month contract)-WTI (one month contract) calendar spread was 0.0089 and the σ was 0.0545 during the period 1992-03-09 to 2012-03-09. The average E[r] was 9.20%. Average λ for the WTI (four month contract)-WTI (one month contract) calendar spread is 0.0121 and the σ was 0.0488 during the period 1992-03-09 to 2012-03-09. The E[r] was 5.43%.

7 Discussion

Changes in the WTI-Brent location spread, during the period 2010-03-09 to 2012-03-09 are a result of a divergence in the price of WTI from the WTI-Brent equilibrium relationship and other types of crude oil. The underlying cause in the divergence of prices from equilibrium is location specific, whereby WTI oil stored in Cushing, Oklahoma is experiencing difficulties in transporting the oil to the gulf coast for global distribution. The significance of Brent in describing the real price of oil is reflected by the volume of Brent and WTI futures contracts traded (see Appendix A, Figure 1). Brent has over-taken WTI in the race to represent global oil prices.

The WTI-Brent relationship is reflected by the correlation, whereby WTI-Brent futures prices experience a drop in the correlation during the event chosen in this thesis, from the average in period 1-9 (pre-event) of 0.998 to 0.899 in period 10 (the event). Prices of WTI and Brent experienced an increase in the standard deviation during period 9, mainly as a result of the drop in oil prices during the financial crisis. Despite the drop in prices, the standard deviation of the spread level remained relatively stable. This indicates a low level of dependency between the effect of factors affecting the individual price series and the actual spread levels themselves. The same low level of dependency was observed from period 9 to 10 where standard deviation for prices decreased while standard deviation for spread levels increased substantially. Despite the low level of dependency between asset prices and spread levels, dependency can be argued to exist. One can argue that a divergence from the long-run mean in spread levels is related to uncertainty of asset prices. Increased uncertainty of asset prices would increase variation in spread levels, i.e. there would be a relation between changes in spread levels and volatility of asset prices. The relationship was detected in this thesis and used to predict the outcome of a structural break.

The relationship is further clarified through the test for co-integration, whereby the unit root test indicates that there is a significant trend present in every two year period except for period 2 for Brent and WTI prices. Since total expected return is calculated for 30 trading day periods, period 2 was examined to find the presents of trends within the 30 days trading periods, see Appendix A, Table 2. It was possible to find a significant trend in every 30 day trading period, despite the two year period not showing a trend. The ADF-test for spreads concluded a high level of co-integration between Brent and WTI in all periods, excluding period 10, which confirms the presence of a structural break at the event.

Pre-event data was used to find the relationship between μ , λ and σ in the 30 day trading period. The dynamics of the model analyzing via the parameters above help to predict the outcome/results in the case of structural breaks and large changes in mean levels μ . Only 10.20% of all observations in the 30 day trading period had mean levels μ further than \$2 from the pre-event long run mean of \$1.256. An estimated value of λ and σ for the purpose of using the model in trades would therefore be less trustworthy with increasing level of μ as the case is with structural breaks. Predictions for

λ and σ are only calculated for μ \$7 away from the pre-event long run mean of \$1.256 since one can argue that a prediction of λ and σ will not be appropriate for a larger μ when the relationship is based on a situation where 89.8% of all observations have μ smaller than \$2 from the pre-event long run mean. The actual λ during the event was as predicted lower than the average pre-event λ . The same stands for σ , which as expected was on average larger during the event than in the pre-event period. Average λ and σ values do not represent the average expected return since λ does not have a linear relationship with expected return. An increase in λ will not increase expected return to the same extent as a decrease of λ decreases expected return, see Figure 7. To predict the $E[r]$ of a structural break, the distribution of λ and σ must be considered. All possible combinations of λ and σ with respect to its distribution must be used to find an estimate of $E[r]$. The technique used in this thesis to calculate $E[r]$ cannot be used to find appropriate estimates of $E[r]$, since the technique is relatively time-consuming. The Matlab code must therefore be upgraded to be significant faster. A solution to this problem can also be to use another mathematical program.

The expected return was calculated using the simple trading strategy and the assumptions that the trader in the initial state of the 30 day trading period knows the actual mean level, rate of mean reversion and volatility for the trading period. Since a trader in real life does not know the exact values for these parameters, the actual return a trader will get is therefore dependent on how precise estimates of these parameters the trader can find. The most important observation is instead the actual change in total expected return from 27.02% to 1.74% when a spread of WTI-Brent was subjected to a structural break. The variance of total expected return in the pre-event was 0.075, which must be considered to be relatively high, resulting in a high level of risk. Hedgers with inherent position in oil WTI will experience a significant amount of risk exposure with the structural break at hand. The green area in Appendix A, Figure 8 shows a decrease the net short position for WTI related products taken by producers and consumers. The category represented by producers and consumers has historically taken a net short position since companies who sell oil tend to take larger positions than those who buy oil. This thesis has shown the risk of creating a spread position with WTI, the decrease in the net short position can be a response to the increasing level of risk connected to spread positions created with WTI.

Jet fuel-heating oil crack spread has a lower rate of mean reversion though a higher level of volatility relative to the WTI-Brent location spread during the pre-event period. The total expected return was lower for the Jet fuel-heating oil than for the WTI-Brent location spread for the pre-event period. One can argue that the Jet fuel-heating oil spread is more sensitive to structural breaks since total expected return will decrease more in the Jet fuel-heating oil spread than in the WTI-Brent location spread if rate of mean reversion decreases with one unit. The model and framework presented in this thesis can be used to evaluate the level of risk associated with hedging positions, such as positions taken by airlines. This thesis has shown the risk related to spread positions taken when hedging for changes in actual prices. The high level of volatility within the spreads chosen in

this thesis result in a high level of risk if spread positions are taken for shorter periods of time to hedge upcoming changes in prices. Thus substantial losses can be generated due to the high level of volatility connected to the spread position.

The calendar spread exhibits a similar average level of mean reversion as the Jet fuel-heating oil crack spread and the pre-event WTI-Brent location spread. The volatility within the calendar spread was lower than both the fuel-heating oil crack spread and the pre-event WTI-Brent location spread, but had a higher average total expected return. The higher level of total expected return is possible due to the differences in distribution of parameters.

The sensitivity of returns to changes in the level of mean reversion is extremely high, reflecting the risk inherent in spread trading whether it is for a hedger or arbitrageur. Although spreads remain insensitive to drastic market changes i.e. volatility, the change in the oil specific characteristics not shared by the counterpart in the spread exposes the spread taker to a high level of risk. The structural break incurred in 2011 gives rise to arbitrage opportunities though losses could be endless. The change in the WTI-Brent spread remains yet to be seen in the airline industry's hedging of jet fuel strategy. The majority of airlines do not hedge the whole jet fuel price exposure and with the portion that is hedged, swaps and heating oil futures are mostly consistent within the derivative portfolio.

Previous research indicates possibilities for arbitrage in location spreads for WTI and Brent and no evidence to the contrary is found within this thesis. Although market efficiency implies that any abnormal gaps in the spread would be dissipated by market participants, contrary to findings in this paper. The reason for a sustained period of arbitrage possibilities might be the risk associated with buying spread positions. Although both types of crude originate from different locations, their inherent physical properties allows for their association as two of the same product. The manner of which the spread is not dissipating begs one to wonder how issues regarding the transportation of WTI oil will affect will affect spread takers in the near future.

The model presented in this thesis to calculate total expected return indicates the possibility of a profiteering, but also the risk associated with changes in properties of the spread process. The chance of getting large gains exist, though are dwarfed by the risk of changes in properties within the spread process.

According to the model presented in this thesis, risk in the presence of structural breaks in mean reverting processes increases mainly as a result of a decrease in the rate of mean reversion. In reality, risk will increase mainly as a result of uncertainty of properties (mean level, rate of mean reversion and volatility) of the spread process. Therefore, structural breaks must be considered as a risky occurrence for arbitrageurs and hedgers.

8 Conclusion

In this thesis a model for calculating expected return using a first-time hitting density function for Ornstein-Uhlenbeck mean reverting processes in the WTI-Brent spread is formed. In order to understand the significance of a structural break in a spread, the relatively recent divergence in the WTI-Brent spread provides good reason to look at the mean-reverting process model in these two commodities specifically. The results tend to the possibility of arbitrage, though this is certainly not risk-free. The period constituting a structural break in the long-run equilibrium mean shows a significant increase in the volatility and a decrease in the rate of speed reversion.

The results concur with that of Dempster, Medova et al. (2008) in terms of opportunity for arbitrage. The risk implied by structural breaks in a mean reverting process are shown by the changes in expected return therefore contributing to the pre-existing literature on spread trading. The standard mean reverting model is taken a step further by including the first-time hitting density. This provides expected returns based on probabilities for reaching and not reaching a pre-specified value.

A significant relationship between distance from the long run mean and the parameters λ and σ was found, further used in the calculation for expected return. Limited divergence from the long-run mean level during the pre-event limits the accuracy of the prediction for the event period. The method of using first-time hitting density to calculate expected return amplifies the relationship between the expected outcome and changes in spread characteristics.

The results are derived from this strategy, which unfortunately remains unrealistic in the sense that hedgers undergo the same trading strategy as arbitrageurs because they cannot close their positions earlier, which would expose them to the risk of price changes.

Assuming an OU-process using a first-time hitting density approximation, the expected return diminishes significantly provided the event in 2010. The model shows that returns are sensitive to changes in the rate of mean reversion, reflecting the risk inherent in spread trading whether it is for a hedger or arbitrageur. Although spreads remain insensitive to drastic market changes i.e. volatility in prices, the change in the oil specific characteristics not shared by the counterpart induce a high level of uncertainty.

Suggestions on further research include looking at other commodities, exchange rates and anything that follows a mean reverting process. Another prospective contribution would be to compute the expected return with respect to the distributions in λ and σ

9 References

Butterworth, D. and P. Holmes (2002). "Inter-market spread trading: evidence from UK index futures markets." *Applied Financial Economics* 12: 783-790.

Cortazar, G. and E. S. Schwartz (1994). "The Valuation of Commodity-Contingent Claims." *The Journal of Derivatives* 1(4): 27-39.

Dempster, M. A. H., E. Medova, et al. (2008). "Long term spread option valuation and hedging." *Journal of Banking and Finance* 32: 2530-2540.

Dicky, D. and W. Fuller (1979). "Distribution of the estimates for autoregressive time series with a unit root." *Journal of the American Statistical Association* 74: 427-429.

Dunis, C. L., J. Laws, et al. (2006). "Trading futures spreads: an application of correlation and threshold filters." *Applied Financial Economics* 16: 903-914.

Engle, R. and C. Granger (1987). "Cointegration and error correction representation, estimation and testing." *Econometrics* 55: 251-276.

Fattouh, B. (2008). "The Dynamics of Crude Oil Price Differentials." *Oxford Institute for Energy Studies*.

Fennell, A. (2010). "Trading Crude Oil Spreads: WTI and Brent." Retrieved 2012.03.12, 2012, from <http://www.letstalkfutures.com/2010/04/21/trading-crude-oil-spreads-wti-and-brent/>.

Gibson, S. (1990). "Convenience Yield and the Pricing of oil contingent claims." *The Journal of Finance* 45(3): 959-976.

Giorno, V., A. G. Nobile, et al. (1986). "Some Remarks on the Rayleigh Process." *Journal of Applied Probability* 23(2): 398-408.

Girma, P. B. and A. S. Paulson (1999). "Risk Arbitrage Opportunities in Petroleum Futures Spreads." *The Journal of Futures Markets* 19(8): 931-955.

Johansen, S. (1991). "Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models." *Econometrica* 59(6): 1551-1580.

Jurek, J. W. and H. Yang (2006). "Profiting From Mean-Reversion: Optimal Strategies in the Presence of Horizon and Divergence Risk." *Working Paper Harvard*

Krapels, E. (2001). "Re-examining the Metallgesellschaft affair and its implication for oil traders." *Oil and Gas Journal*.

Linetsky, V. (2004). "Computing hitting time densities for CIR and OU diffusions; applications to mean-reverting models." *Journal of Computational Finance* 7(June): 1-22.

Ma, W. C. (1989). "Forecasting Efficiency of Energy Futures Prices." *Journal of Futures Markets* 9: 393-419.

Melamed, L. (1981). "The Futures Market: Liquidity and the Technique of Spreading." *The Journal of Futures Markets* 1(3): 405-411.

Merton, R. C. (1971). "On The Pricing Of Contingent Claims and Modigliani-Miller Theorem." *Journal of Financial Economics* 5(2): 241-249.

Milonas, T. N. and T. Henker (2001). "Price spread and convenience yield behaviour in the international oil market." *Applied Financial Economics* 11: 23-26.

Samuelson, P. (1964). "Theoretical notes on trade problems." *Review of Economics and Statistics* 46: 145-154.

Schwartz, E. S. (1997). "The Stochastic Behaviour of Commodity Prices: Implications for Valuation and Hedging." *The Journal of Finance* 52(3): 923-973.

Vasicek, O. (1977). "An Equilibrium Characterization of the Term Structure." *Journal of Financial Economics* 5: 177-188.

Working, H. (1949). "The Theory of Price of Storage." *The American Economic Review* 39(6): 1254-1262.

Appendix A

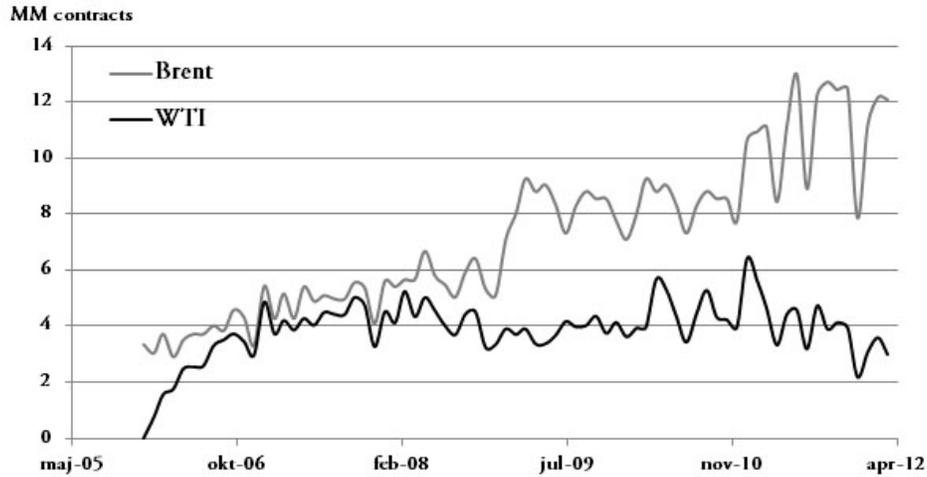


Figure 1: Volume of futures contracts traded on the ICE (Millions of contracts/month). Source: ICE Report Center. Following a close relation in demand for both types of crude futures, the appeal for Brent over WTI is clear-cut

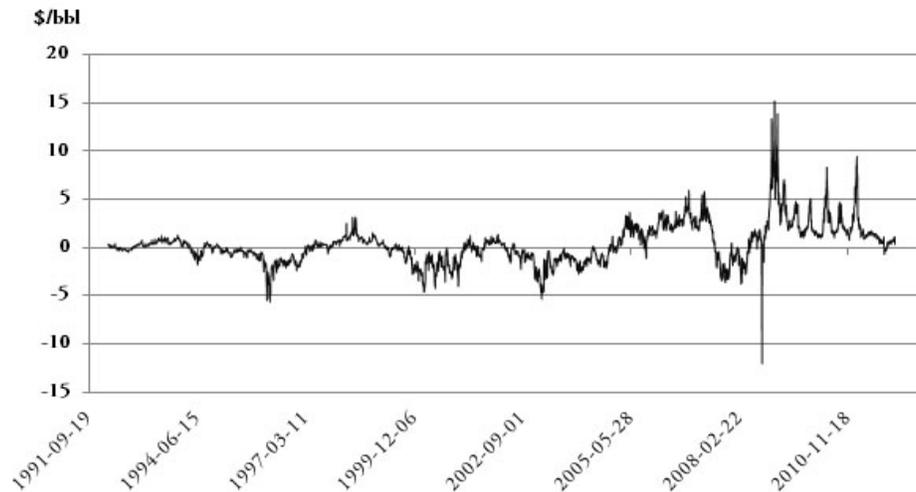


Figure 2: Calendar spread for daily roll-over prices (closest to given maturity) for one and two month futures contracts of WTI from 1992-03-09 to 2012-03-09 traded on the NYMEX. The spread is calculated by subtracting the daily price of one month futures contracts from two month futures prices of WTI. Source: Bloomberg

Period	Start date	End date	Obs	Brent				WTI				Spread			
				Mean	Max	Min	Std	Mean	Max	Min	Std.	Mean	Max	Min	Std
1-10	1992-03-09	2012-03-09	5220	42.49	9.64	146.08	31.30	42.64	145.29	10.72	28.97	0.15	14.88	-27.88	4.44
1-9	1992-03-09	2010-03-09	4697	36.16	146.08	9.64	25.67	37.42	145.29	10.72	25.44	1.26	14.88	-10.67	1.47
1	1992-03-09	1994-03-09	523	17.88	21.57	12.92	2.18	19.14	22.89	13.91	2.30	1.27	2.02	0.28	0.24
2	1994-03-09	1996-03-11	524	16.75	19.22	13.02	1.09	18.16	20.75	14.08	1.18	1.41	3.30	0.18	0.38
3	1996-03-11	1998-03-09	521	19.60	24.86	12.98	2.45	21.09	26.62	14.33	2.52	1.49	5.73	0.21	0.71
4	1998-03-09	2000-03-09	524	16.77	31.90	9.64	5.37	18.01	34.13	10.72	5.46	1.24	3.70	-1.42	0.62
5	2000-03-09	2002-03-11	523	26.07	34.59	17.68	4.04	27.33	37.20	17.45	4.46	1.26	5.66	-1.44	1.11
6	2002-03-11	2004-03-09	522	27.74	34.10	22.70	2.46	29.90	37.83	23.47	3.18	2.16	5.45	-0.82	1.15
7	2004-03-09	2006-03-09	523	49.49	64.72	30.21	10.10	51.68	69.81	34.27	9.38	2.19	6.53	-1.66	1.31
8	2006-03-09	2008-03-10	523	72.41	104.16	51.70	11.79	72.30	107.16	50.48	12.69	-0.11	4.87	-6.45	1.98
9	2008-03-10	2010-03-09	522	78.84	146.08	36.61	27.98	79.22	145.29	33.87	28.87	0.39	14.88	-10.67	2.41
10	2010-03-09	2012-03-09	524	99.33	126.65	69.55	16.44	89.61	113.93	68.01	10.65	-9.73	2.68	-27.88	8.29

Table 1: Summary statistics

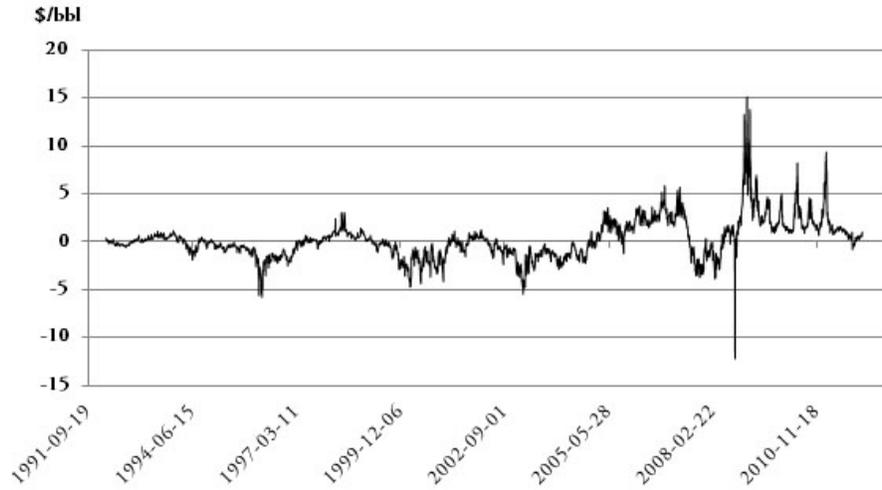


Figure 3: calendar spread for daily roll-over prices (closest to given maturity) for four and one month futures contracts of WTI from 1992-03-09 to 2012-03-09 traded on the NYMEX. The spread is calculated by subtracting the daily price of one month futures contracts from four month futures prices of WTI. Source: Bloomberg

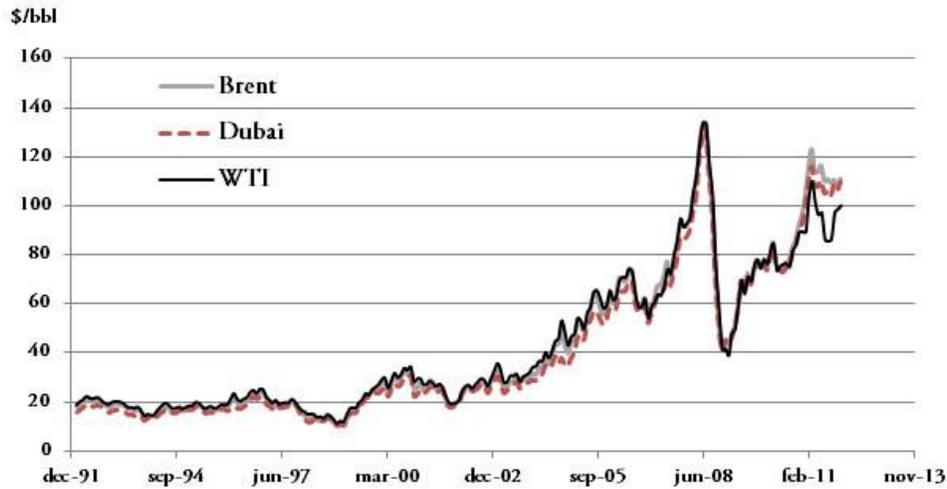


Figure 4: Monthly average of benchmark prices for WTI, Brent and Dubai. Prices retrieved from Datastream

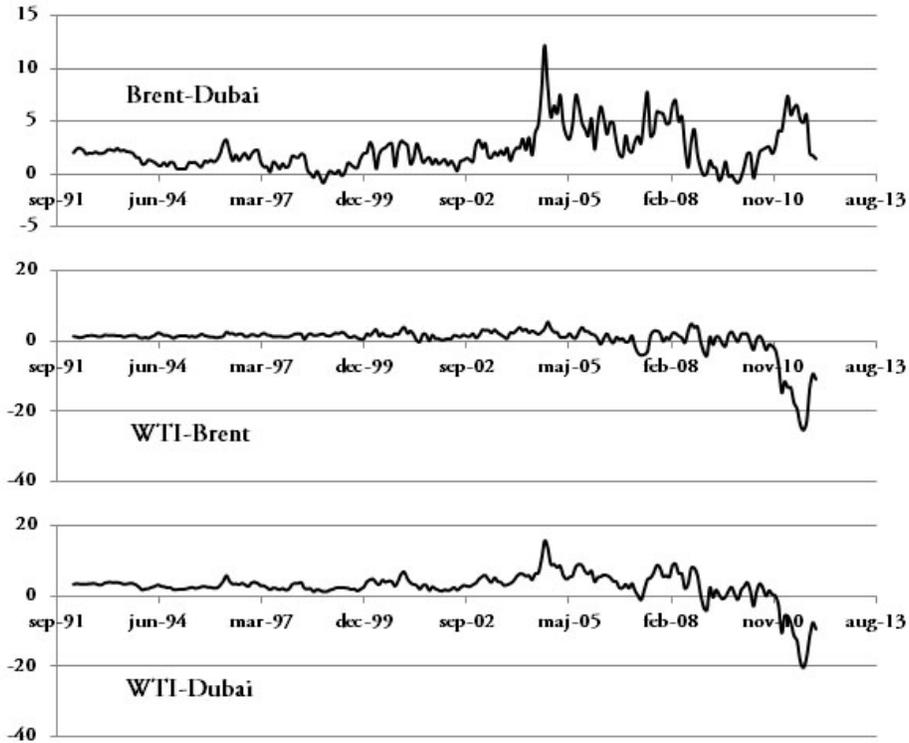


Figure 5: location spread for benchmarks of Brent-Dubai, WTI-Brent and WTI-Dubai)

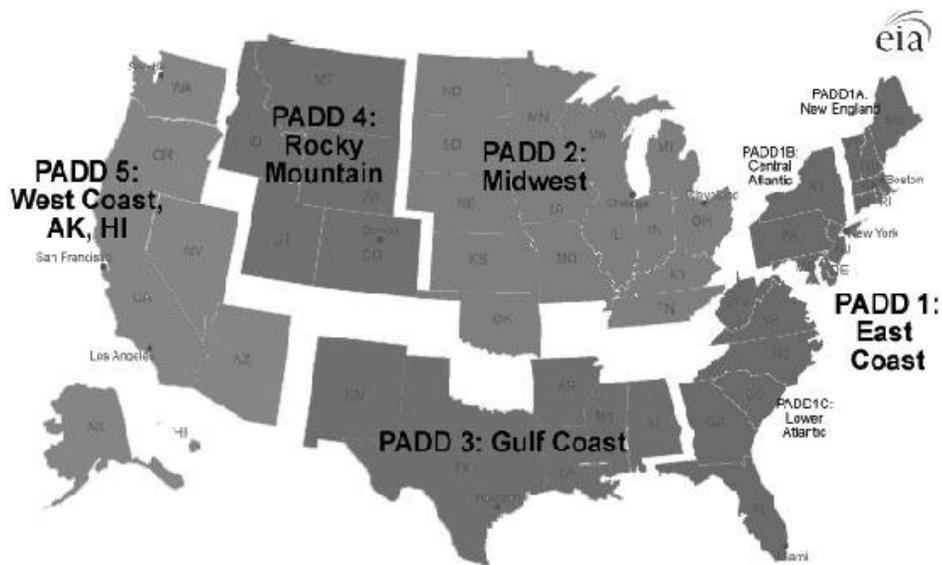


Figure 6: oil regions of the United States. Source: EIA (Energy Information Administration U.S)

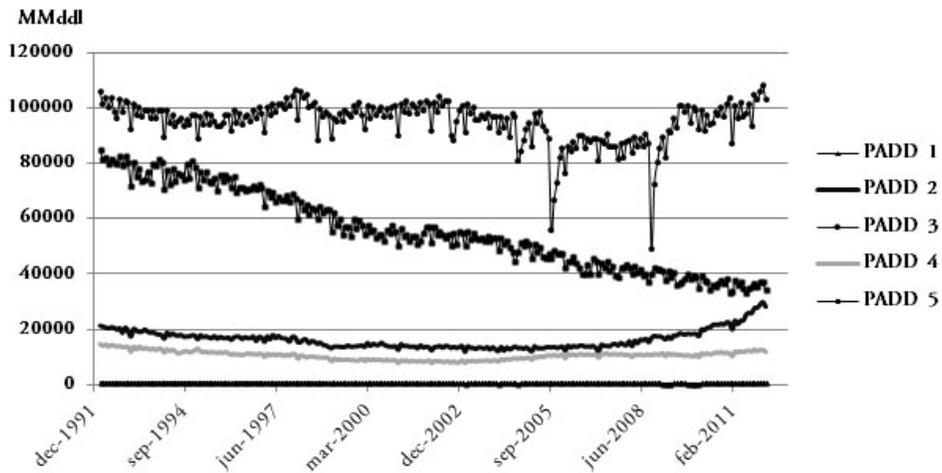


Figure 7: crude oil production according to region. Source: EIA

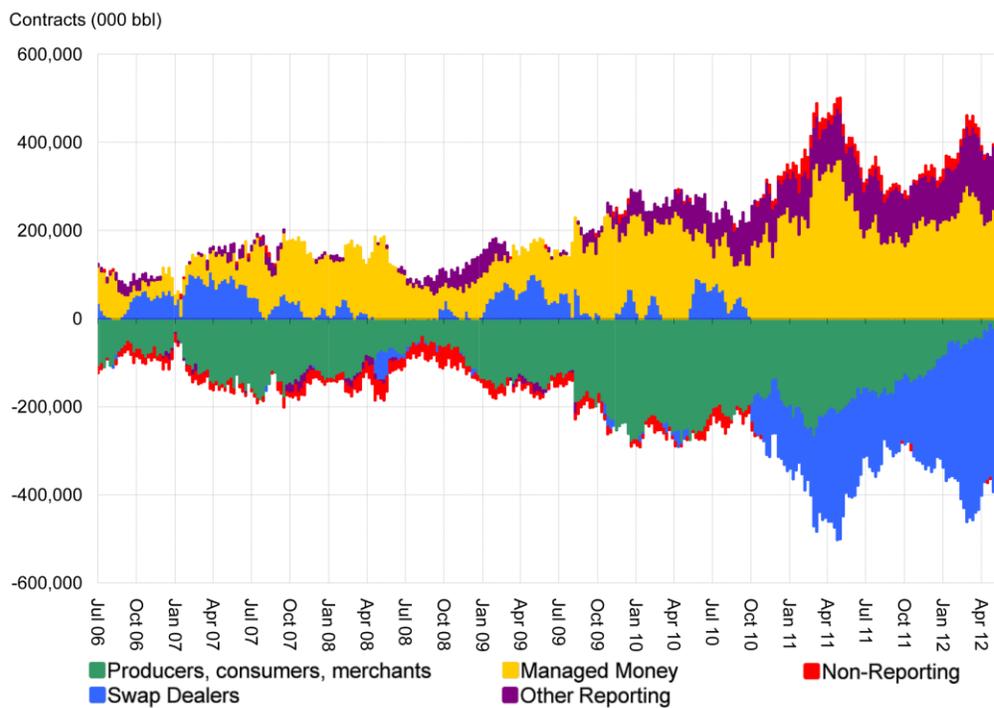


Figure 8: Net positions in WTI-related contracts traded on NYMEX, Source: Reuters

Start Date	Finish Date	Brent	WTI	Critical Values		
				10%	5%	10%
1994-03-09	1994-04-20	-1.541	-0.949	-3.75	-3.00	-2.63
1994-04-21	1994-06-02	-1.807	-0.188	-3.75	-3.00	-2.63
1994-06-03	1994-07-15	-0.348	-1.856	-3.75	-3.00	-2.63
1994-07-18	1994-08-29	-1.576	-0.861	-3.75	-3.00	-2.63
1994-08-30	1994-09-28	-0.875	-0.862	-3.75	-3.00	-2.63
1994-09-29	1994-11-23	-0.494	-0.577	-3.75	-3.00	-2.63
1994-11-24	1995-01-05	-2.315	-1.712	-3.75	-3.00	-2.63
1995-01-06	1995-02-17	-2.449	-0.552	-3.75	-3.00	-2.63
1995-02-17	1995-04-03	-2.181	-0.637	-3.75	-3.00	-2.63
1995-04-03	1995-05-16	-1.924	-1.667	-3.75	-3.00	-2.63
1995-05-16	1995-06-28	-1.818	-0.803	-3.75	-3.00	-2.63
1995-06-28	1995-08-10	-1.605	-1.714	-3.75	-3.00	-2.63
1995-08-10	1995-09-22	-1.899	-1.401	-3.75	-3.00	-2.63
1995-09-22	1995-11-06	-1.180	-1.288	-3.75	-3.00	-2.63
1995-11-06	1995-12-19	-0.831	-0.799	-3.75	-3.00	-2.63
1995-12-19	1996-01-31	-1.564	-0.864	-3.75	-3.00	-2.63
1996-01-31	1996-03-14	-2.335	-0.458	-3.75	-3.00	-2.63

Table 2: ADF test: Spread levels

Appendix B

```
%Matlab code for a least squares regression for calculating lambda (rate
%of mean-reversion), mu(the long-run spread level) and sigma(volatility
%parameter). A matrix "S" is constructed with spread prices.
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

A=(0:30:length(S)); %Dividing dataset of prices into periods of 30 trading
%days.

for i=1:length(A)
start=A(1,i); %Finding the starting point in dataset of spread prices.
endt=start+29; %End point in dataset is 29 observations forward, ending up
%with 30 observations.

delta=30; %Number of steps in the chosen data.
n = 29;
Sx = sum( S(start:endt-1) );
Sy = sum( S(start+1:endt) );
Sxx = sum( S(start:endt-1).^2 );
Sxy = sum( S(start:endt-1).*S(start+1:endt) );
Syy = sum( S(start+1:endt).^2 );

a = ( n*Sxy - Sx*Sy ) / ( n*Sxx -Sx^2 );
b = ( Sy - a*Sx ) / n;
sd = sqrt( (n*Syy - Sy^2 - a*(n*Sxy - Sx*Sy) )/n/(n-2) );

lambda = -log(a)/delta;
mu = b/(1-a);
sigma = sd * sqrt( -2*log(a)/delta/(1-a^2));
ANS(i,:)=[lambda, mu, sigma]
end
```

Figure 9: Matlab code for estimating μ , λ and σ

```

%Matlab code for calibrating expected return from prices following a
%Ornstein-Uhlenbeck (Vasicek) process, using an intensity based probability
%measure of conversion to mean.
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%
%Inputs to the intensity based probability measured return are lambda (rate
%of mean-reversion), mu(the long-run spread level) and sigma(volatility
%parameter), all calculated using a least squares regression method. The
%function E_return will calculate:
%
%r_pc = expected return from the probability of convergence before time T.
%r_pnc_A = expected return from the probability of not converting before T
% but taking the value of \emph{y} > $0 at time T.
%r_pnc_B = expected return from the probability of convergence but before
% - T take a value of \emph{y} > $0.
%r_pnc = r_pnc_A + r_pnc_B = Total probability of not converting at T.
%E_return = r_pc + r_pnc = total expected return.

function(r_pc, r_pnc_A, r_pnc_B, r_pnc,
E_return) = E_return(lambda, mu, sigma)

x = (2*sigma); %Initial spread level
N = (30); %Number of trading days
N = 1000; % Number of trials for large-N assumption to achieve desired error
%tolerance. Function N_estimate are used for this purpose.

n = (1:1:N)';

x_b = ((sqrt(2*lambda))/sigma) * (x - mu); %Normalized spread level
y_b = ((sqrt(2*lambda))/sigma) * (y - mu); %Normalized long-run mean level

%Using notations from Linetsky (2004) when estimating eigenvalues lambda_n
k_n = (n - (1/4) + ((y_b^2)/(pi^2))) + ((y_b * sqrt(2))/pi) * (sqrt(n -
(1/4) + ((y_b^2)/(2*pi^2))));
lambda_n = lambda * (2 * k_n - (1/2));

%Using notations from Linetsky (2004) when estimating coefficients c_n
c_nA = (((-1)^(n+1)) * (sqrt(k_n))) / ((2 * k_n - 0.5) * (pi * sqrt(k_n))) * ((2 * (y_b -
y_b)) * y_b));
c_nB = exp((1/4) * ((x_b^2) - (y_b^2)));
c_nC = cos((x_b * sqrt(2 * k_n)) - (pi * k_n) + (pi/4));
c_n = c_nB * c_nC * c_nA;

```

```

%Return from probability of conversion.
r_pc=x*(sum(c_n.*(1-(exp(- lambda_n.*T)))));

%Return from probability of no conversion(r_pnc_A) before T
sigma_s=sqrt(((s_qmA2)/(2-d lambda))*(1-(exp(-2
lambda_n.*T))))); u_s=((x*(exp(- lambda_n.*T))+u*(1-exp(-
lambda_n.*T)));

OO=10AS; %approximation for infinity
E=@(y)((x-y).*(1./(sqrt(2*pi))*s_gma_s)).*(exp(-0.5.*((
(y-u_s).^2)./(s_gma_sA2)))));
r::pnc_A=quad(E,0;OO);

%Return from probability of no conversion, but ending up in \emph{y} %after
conversion(r_pnc_A).

for i=1: length(n);
E=@(y,t)(x-y).*(c_n(i)-d lambda_n(i)).*(exp(-
lambda_n(i).*(1-t)).*(1./(sqrt(2*pi))*s_gma_s(i)).*(1-(exp(-2*k_n(i).*(T-
t)))))).*(exp(-0.5.*((y-(u*(1-exp(-k_n(i).*(T-
t))))).^2)./(((s_qmA2)/(2*k_n(i)).*(1-(exp(-2*k_n(i).*(T-t)))))))));
DB(i)=db_quad(E,0,OO,0,T);
End
r_pnc_B=sum(DB);
;

r_pnc=r_pnc_A-r_pnc_B; %Return from probability of not converting

E_return=r_pc+r_pnc; %Total return: return from converting and return from
%not converting;

%The 'N_estimate' function determines how many terms needed in the series to
%achieve a desired error tolerance, used by Linetsky (2004).
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

function [n,NN]=N_estimate(lambda,mu,sigma)

A=((2-d lambda)/pi).*(exp(0.25.*(((sqrt(2-d lambda))/s_gma).*(x-
mu))^2)-(((sqrt(2-d lambda))/s_gma).*(y-mu))^2));
B=2-d lambda;
n=[1:1:2000];

x_b=((sqrt(2*k_n))/s_gma).*(x-mu);
y_b=((sqrt(2*k_n))/s_gma).*(y-mu);

k_n=(n-(1/4)+((y_b^2)/(piA^2))+((y_b*sqrt(2))/pi).*(sqrt(n-
(1/4)+((y_b^2)/(2*piA^2)))));
lambda_n=k_n.*(2*k_n-(1/2));

c_nA=(((1-(1/4)).*(n+1)).*(sqrt(k_n))./(2*k_n-0.5)
-(((pi-sqrt(k_n)).*(2A-0.5)).*(y_b))));
c_nB=exp((1/4).*(x_b^2)-(y_b^2));
c_nC=cos((x_b*sqrt(2*k_n)).*(pi-k_n)+(pi/4));
c_n=c_nB.*c_nC.*c_nA;

NN=(dg((abs(c_n.-lambda_n.-exp(lambda_n)))/A)/B)
end

%The appropriate N are located using a visual approach.
plot(n,NN)

```

Figure 10: Matlab code for calculating $E[r]$ using first-time hitting density approximation.