NORWEGIAN SCHOOL OF ECONOMICS AND
BUSINESS ADMINISRATION (NHH)
Bergen, Spring 2009

THE IMPACT OF SPECULATORS' ACTIVITY ON CRUDE OIL FUTURES PRICES:

empirical evidence of crude oil market efficiency and causal relationship between traders' positions and market returns

Master Profile:
Financial Economics

Executor:
Irina Morari

Supervisor:
Rögnvaldur Hannesson

This thesis was written as a part of the Master of Science programme in Economics and Business Administration. Neither the institution, nor the advisor is responsible for the theories and methods used, or the results and conclusions drawn through the approval of this thesis.

Abstract

Recently the speculators activity in crude oil futures markets has received a great deal of scrutiny – for instance, OPEC argued that speculators have been a critical factor to push up oil prices.

The objective of this paper is to assess the claims regarding speculative influences on oil price. More specifically, I intend to investigate whether speculation/arbitrage opportunities exist in crude oil futures market by testing for market efficiency; and, whether speculators do affect crude oil futures prices by testing for causality between traders' futures positions and market returns.

For this purpose I apply Johansen's co-integration methodology (1988), Engle-Granger's error-correction methodology (1987), Granger causality framework (1969) and Cumby and Modest's market timing framework (1987).

For the first test, findings show that the oil futures market is "long-term" efficient but does undergo "short-term" deviations; and that futures prices lead spot prices; furthermore, it was confirmed that futures prices on contract of longer maturity lead futures contracts with shorter maturity.

For the second test, it was found that traders' positions do not generally lead market returns; and that extreme levels of traders' positions have no impact on market returns.

These results have implications for various players in the oil market like international organizations, oil companies and governments when making investments decisions and policy recommendations.

Acknowledgements

I would like to thank Professor Rögnvaldur Hannesson, and Professor Ole Gjølberg for there contribution and helpful comments.

Contents

TABLES AND GRAPHS

INTRODUCTION

Through the 1990's oil price remained relatively stable (the first Gulf conflict left aside). However, from 2002 the oil prices began an increasing trend, with the rise in prices accelerating in 2004. Limited supply capacity and robust demand in emerging economies have been the driving factors used to justify much of the upward trend in oil prices during this period. However, in a year period from July 2007 to July 2008 oil price almost doubled, surpassing its previous peak in 1979 by over 10% in real terms; and since July (to December), it has fallen by more than 70% to four-years low. This surge in price has intensified the debate around the drivers of oil prices focusing the attention on the role of speculators in the oil market. There is some agreement among practitioners[1] that this precipitous rise is all the result of speculation, unsupported by the rudiments of supply and demand. Despite these allegations, there is no empirical evidence of speculators having systematically driven up prices.

The international policy reports and energy publications have been advancing several reasons to explain this price spike. The sustained rise in oil prices over several years has in part been self-reinforcing as growing market liquidity and financial innovation enabled more market participants to enter the market. This sustained growth has also attracted a new breed of investors in search of positive returns against a backdrop of falling equity, bond and credit markets. This has been supported by the fact that oil as a commodity has increasingly been viewed as a way of diversifying portfolio risk and hedging against future inflation and exchange rate fluctuations. The new breed of investors includes financial institutions and commodity index funds *inter alia*. They are believed to be interested only in riding a price trend by trading in futures contracts. Hence, they are not financing new oil wells that could boost global crude supplies; they are just reaping price gains in the commodity markets. Not surprisingly, the interest for how speculators affect the stability, price discovery and liquidity of asset markets is growing and calling for empirical evidence.

[1] At the London Energy Ministers' Meeting, December 2008, A.S. El-Badri, OPEC Secretary General, asserted that "in the summer, prices were driven to record highs by unlimited speculation, as the dollar weakened and investors sought cover in commodity market". (http://www.opec.org/opecna/Speeches/2008/LondonEnergy.htm)
Gillian, president of Petroleum Markets Association asserted that "Approximately 60 to 70 percent of the oil contracts in the futures markets are now held by speculative entities. (…) oil prices seemed to disconnect from the basic fundamentals of supply and demand". (http://www.cbsnews.com/stories/2009/01/08/60minutes/main4707770.shtml)

On the opposite side of the debate, there are academics like Krugman[2] and Weiner[3], *inter alia* asserting that speculation was not prevalent in the market when the price surged last year. Krugman (2008) argued that the only way speculation can increase oil prices is through hoarding or increasing private inventories of crude. He maintains that through the period of the alleged “oil bubble”, inventories have remained at more or less “normal” levels, which implies that the rise in oil prices is not the result of runaway speculation, but the consequence of decreasing supply and the rapid growth of emerging economies like China and India.

Furthermore, Weiner (2002) contended that “there is no reason to believe that speculation would result in an average level of commodity prices either higher or lower than would occur in its absence; rather, it is average volatility that would rise with speculative activity”. Justifying this statement he claims that “even if speculators can raise prices by buying up futures contracts, they cannot unload these positions at the higher price without a change in market fundamentals. The very action of unwinding their large positions will cause prices to fall. Therefore, the widely observed correlation between the size of speculative/ noncommercial positions cannot tell us anything about the profitability of such positions, nor whether speculators are making the market more or less efficient”.

Moreover, analysis by the Commodity Futures Trading Commission (CFTC) finds no causality in terms of various groups of traders (commercial and non-commercial) at NYMEX changing their positions in advance of changes in price. If anything, the analysis indicates causality in the opposite direction – many trader groups adjusting their positions in response to price changes[4]. Neither does a recent analysis by the International Monetary Fund find evidence of increased financialisation of commodities since 2003 having had a significant impact on the futures price level, futures price volatility, or on co-movement in futures prices across a range of commodities[4].

These controversial allegations warrant and motivate for further empirical research about the impact of speculators’ activity on oil prices. The research has implications for various players in the oil market like international organizations, oil companies and governments when making investments decisions and policy recommendations.

[2] Professor of economics and international affairs at Princeton University, a centenary professor at the London School of Economics, and an op-ed columnist for The New York Times. (http://www.nytimes.com/2008/05/12/opinion/12krugman.html?_r=1)

[3] George Washington University, Global Management Research, & Groupe de Recherche en Économie de l’Énergie et des Ressources Naturelles, Université Laval (http://www.sciencedirect.com/science/article/B6W5X-45MW05T-D/2/3fe7531fbbfdb448be452c5d1e7b6107)

[4] Cabinet Office UK, Global Energy team, ”The rise and fall o foil prices; analysis of fundamental and financial drivers”, (December, 2008)

In this paper I use a quantitative approach to investigate the assertions about the impact of speculators' activity on crude oil futures prices.
Generally, the test of the impact of speculators on market returns can become very elaborate and be limited by availability or quality of the data. Taking this into account I reduce my research to a market efficiency test – investigating whether there are any possibilities for speculative/arbitrage activity in the oil market in the long- and short-run and whether they are statistically significant; and price discovery test – investigating the information flow across spot and futures markets. This is supplemented with a test of causality relationship between traders' positions and market returns and a test of the impact of extreme traders' positions on the market. Though these tests do not provide explicit conclusions on whether noncommercial traders can influence oil prices, it provides insights into the price discovery process and the impact of all traders on the oil prices.

I begin the research using the traditional regression approach developed by Fama (1984). This is also called weak-form-market or speculative efficiency test. In the context of this approach, market efficiency requires that futures price should be unbiased predicator of futures spot price. Specifically, this approach tests whether the basis contains information about future spot price and about risk premium at the expiration of the future contract.
Further, I proceed with a more elaborate and more reliable test involving the cost-of-carry model. This test comes in two variants: test of long-term and short-term market efficiency as given by the pricing relation and test for price leadership. More specifically, I study (1) the long-run equilibrium relationship between the futures price and the spot price; (2) the long- and short-run efficiency of the futures market as an unbiased predictor of spot prices; and (3) lead-lag relationship between spot and futures prices. For this purpose I employ Johansen's co-integration methodology (1988), Engle-Granger error-correction methodology (1987) and Granger Causality framework (1969).
Lastly, I investigate (4) the causality relationships between net traders' futures positions and market returns and (5) the impact of extreme traders' positions on market returns. In particular, the question investigated is: ''Are (extreme levels of) trader positions useful for predicting market returns?''. This is not meant to be a test of trading profitability; rather, I attempt to investigate the informational content of these data in a general sense. Methodologically this is carried out by using Granger causality framework (1969) and the market timing framework similar to that of Cumby and Modest (1987).

The remainder of the study is organized as follows.
Section two sets out the theoretical framework and the hypotheses tested throughout the paper. The theoretical framework includes the price formation theories for commodity futures prices, such as theory of expectation and theory of storage for commodity; and the hypotheses tested are those of market efficiency and unbiasedness. Along with this I describe briefly the role of the futures market and the role of the market regulator CFTC.

Section two presents the methodology applied regarding the implementation of different behavioural hypotheses tests. Fama's regression approach (1984), Johansen's co-integration methodology (1988), Engle-Granger error-correction methodology (1987), and Granger causality framework (1969) are explained in this section.

Section three provides the data description and the statistical summary of data properties. As well I present a review of recent literature regarding the market efficiency and the impact of speculators.

Section four dwells on the description of the implementation process, empirical analysis and discussion of the results. The econometric evaluation of market efficiency and unbiasedness hypotheses is carried out within a co-integration and error correction framework, given that the time series for spot and futures are nonstationary variables. To increase the power of these tests and to make a more general statement about efficiency and unbiasedness I test separately and jointly the hypotheses of market efficiency and unbiasedness by imposing restriction on the parameters of the co-integration relation and on the parameters of the error-correction representation. Further the analysis of price discovery and information flow across spot and futures markets has been carried out by employing the Granger causality test and by imposing restriction on the relevant coefficients. The empirical evidence of the causality relationships between traders' futures positions and market returns was undertaken by using the Granger causality framework. This is supplemented by the test on the impact of extreme traders' positions on market returns. This test was carried out using the market timing framework proposed by Cumby and Modest.

Last section outlines the conclusions that emerge from the tests undertaken. As well, I discuss possible limitations that would alter the conclusions and possible implications that the study yields. Lastly, I advance some suggestions that would motivate an eventual further research.

1 THEORETICAL FRAMEWORK

1.1. Role of Futures Markets

Market participants would agree that the main benefits from futures markets are their price discovery and risk management functions; moreover, that futures markets exist for commodities subject to very high levels of uncertainty about future spot prices.

While futures markets fulfill their price discovery function, in that they provide forecasts of future spot prices, risk management refers to hedging by controlling the risk associated with spot price fluctuations. Other advantages of futures markets are, among others, the reversibility of futures contracts, the voluntary participation in markets, the continuing operation of markets, the inter-temporal allocation of resources, and the lay-off risk instrument – the transfer of the risk associated with random fluctuations of spot prices form hedgers to speculators.

Whereas commercial hedgers - typically those who have an underlying commercial interest in the commodity, such as oil producers, refineries, and airlines, want to avoid an exposure to adverse movements in the price of an asset, speculators such as commodity index funds wish to take a position in the market betting on the price movements. In this way speculators assume the price risk that commercial hedgers wish to unload. Yet another type of investor, the arbitrageur, is in the futures market to take advantage of disparities between prices in two different markets. If they see the futures price of an asset getting out of equilibrium with the spot price, they will take offsetting positions in the two markets to lock in a profit.

All of them are interested in the efficiency of futures markets. Hedgers are interested in futures market efficiency since they base their investment decisions –as investment in new oil fields, refinery and storage capacity – on expectation of price development. Speculators and arbitrageurs are interested in the efficiency of futures markets since they capitalize on arbitrage opportunities whenever short-run prices depart from the long-run equilibrium.

Testing the efficiency of futures markets is an important research issue for both market participants and supervisors. If futures markets are efficient, that is, if the futures price is the best unbiased predictor of the subsequent spot price, implying that the current futures price incorporates all relevant information, agents are able to mitigate potential losses by using appropriate hedging tools. Alternatively, if futures markets are inefficient, they may introduce an extra cost to hedgers, such as losses caused by the price volatility.

1.2. The CFTC’s Reporting System

Market efficiency is also of interest for the Commodity Futures Trading Commission (CFTC), the futures market supervisor, since its primary task is to ensure market integrity and customer protection.

The CFTC is responsible for monitoring and regulating futures and options trading in order to ensure that the markets are free from manipulative influences or other price distortions. One of the measures used to achieve this goal is the CFTC’s market surveillance program, known as the Large Trader Reporting (LTR) system, meant to “determine when a trader’s position in a futures market becomes so large relative to other factors that it is capable of causing prices to no longer accurately reflect legitimate supply and demand conditions” (CFTC, No. 5-92).

The Commodity Futures Trading Commission (CFTC) collects data from futures commission merchants (FCM), clearing members, and foreign brokers on the composition of open interest for all futures contracts[5] (Sanders, et al., 2004).

The open interest includes reporting and non-reporting traders’ positions, where reporting traders hold positions in excess of CFTC reporting levels[6]. Reporting traders are further classified as commercials or non-commercials. Commercials typically include those who have an underlying commercial interest in the commodity upon which the futures contract is based, and are also referred to as “hedgers”. For instance, a bank using the futures contracts in order to transfer its risk exposure to rising interest rates, or an oil refiner to lock in the price of its heating oil and gasoline output. Non-commercials are not involved in an underlying cash business and they are referred to as speculators. These are commodity index investors and financial institutions, among others (Sanders, et al., 2004). The non-reporting traders are those that do not hold positions in excess of CFTC reporting levels.

Overall, the positions are broadly discussed in terms of hedgers (reporting commercials), funds (reporting non-commercials), and small speculators (non-reporting traders).

Within the CFTC’s LTR system each futures account is identified with an “owner” and a “trader.” The “trader” is an entity that makes trading decisions or has material financial interest. For example, a large corporation may have refining, exploration, and retail

[5] A subset of this data is released to the public through the CFTC's Commitments of Traders (COT) report.

[6] The reportable level is on a futures-equivalent or delta-adjusted basis. That is, option and spread positions are adjusted to reflect their sensitivity (delta) to the underlying futures price. So, a trader may hold contracts in excess of the reportable level, but if the position is delta-neutral, then it is not a reportable position (Sanders, et al., 2004).

departments. The overall corporation is the account "owner," but each department may be considered a separate "trader." A "trader" may have accounts with a number of FCMs. Positions are aggregated across accounts controlled by the same entity and those in which the entity has a 10% or greater financial interest. Thus, within the context of the CFTC reports, a "trader" is any entity that directly controls trading (i.e., an authorized trader) or has at least a 10% financial interest in an account. A trader's position is aggregated across all such accounts (Sanders, et al., 2004).

When an account has a reportable position FCM sends out CFTC Form 102. Form 102 provides information regarding financial interests and the commercial nature of the account. The account trader is requested to complete a CFTC Form 40 within 10 days of obtaining a reportable position. With Form 40 detailed information on the controlling interest in the account is collected and the trader is asked to self-identify as a commercial or noncommercial, where a commercial is "engaged in business activities hedged by use of the futures and option markets… this would include production, merchandising, or processing of a cash commodity, asset/liability risk management, security portfolio risk management, etc." (CFTC Form 40). In addition, more detailed data are collected about the trader's incentives. For instance, non-commercials are asked to identify themselves as commodity trading advisors (CTAs), commodity pool operators (CPOs), or floor brokers. Likewise, commercials are required to identify the cash markets in which they have underlying risk and the nature of their commercial business (e.g., producer, processor, merchandiser, or end-user). Form 40 is updated every two years or upon special calls by the CFTC (Sanders, et al., 2004).

Having outlined what CFTC regulation involves, it is important to note the limits of that regulation. The CFTC's mandate does not include imposing limits on market risk, leverage parameters, capital requirements, risk assessment procedures or the instruments that may be traded. Moreover, the coverage of existing data does not extend to off-exchange/over the-counter (OTC) activities. This is a significant drawback, since financial investments in commodity markets occur largely off-exchange through swap dealers who only hedge residual risk. Thus, a full picture of the activities of financial investors requires more detailed information on the actions of swap dealers in the OTC markets.

Hedgers, funds, and speculators in the crude oil futures market

The following relation presents how the market’s total open interest (TOI) is disaggregated (Sanders, et al., 2004):

$$\underbrace{\overbrace{[NCL + NCS + 2(NCSP)]}^{\text{Noncommercial}} + \overbrace{[CL + CS]}^{\text{Commercial}}}_{\text{Reporting}} + \underbrace{[NRL + NRS]}_{\text{Nonreporting}} = 2(TOI)$$

where, NCL, NCS, and $NCSP$ are noncommercial long, short, and spreading positions, respectively. CL and CS are commercial long and short positions, and NRL and NRS are long and short positions held by non-reporting traders. Reporting and non-reporting positions must sum to the market's TOI, and the number of long positions must equal the number of short positions.

Sanders, et al., (2004) point to the weaknesses of these definitions and classifications. While they agree on the fact that the basic classification of reporting versus non-reporting is relatively clear across traders and that large measurement errors with respect to position size are unlikely, they stress the fact that this description tells nothing about the incentives of non-reporting traders - they may be hedgers, speculators, or market makers. Furthermore, they stress the fact that the disaggregation of reporting traders into commercial versus non-commercial market participants has potential sources of error - commercial traders may not always be hedgers, and hedgers may not always be hedging. For instance, because of the speculative position limits placed on non-commercials, there is some incentive for traders to classify themselves as commercials. Also, since cash positions for true commercials are unknown, their positions may be speculative in nature. (Sanders et.al., 2004).

In summary, the trader’s labels of “funds,” “hedgers,” and “small speculators” placed on the CFTC trader classifications of reporting non-commercials, reporting commercials, and non-reporting traders, respectively, are somewhat tenuous. First, there is no information about the incentives of non-reporting traders. It is only known that they do not hold positions in excess of CFTC reporting levels. Second, pure hedge positions are a subset of the reporting commercial classification, and reporting commercial positions likely reflect a diverse set of incentives in aggregate. Finally, the “funds” or reporting non-commercials are probably the most precise classification, effectively capturing the positions of a subset of speculators (i.e., managed funds) (Sanders et al., 2004).

1.3. Theory of Futures Pricing

Generally, there are two views on the price formation process for commodity futures prices (Fama, French, 1987):

1. Theory of expectation, which implies that the futures price contains a forecast of the future spot price and an expected risk premium.
2. Theory of storage for commodity, which implies that the futures price of a commodity will be the same as the cost of borrowing funds, purchasing the commodity in the spot market and storing it over the borrowing period.

Both theories infer that there should be a long run stable relationship between spot and futures prices.

1.3.1. Expectation Hypothesis

In financial literature, it is common to consider prices as following a random walk and every change as being unpredictable and both independently and identically distributed. The spot price is formed continuously based on the available information and will change only if new information flows in the market. Hence, the spot prices of an underlying asset can be thought as the best available predictor of the expected future spot rate:

$$S_t = E[S_{t+T}] \qquad (1)$$

where S_t is the spot price at time t and $E[S_{t+T}]$ is the expected future spot price at time $t+T$. However, it is possible that expectations about the future spot price will deviate from the price that finally is going to prevail by some random error e_{t+T}:

$$S_{t+T} = E[S_{t+T}] + e_{t+T} \quad (2)$$

Within equation (1) profit opportunities will still exist. The risk- neutral agents will try to make a profit, by buying an asset at a discount from the spot market and then selling it at a premium in the future market, whenever the level of future price diverges from their expectations about the spot price at a certain moment in future. As consequence of buying and selling future contracts, the price will change until it equals the expected spot price:

$$F(T)_t = E[S_{t+T}] \qquad (3)$$

where $F(T)_t$ is the futures price at time t on a futures contract that expires at time T.

Combining equations (2) and (3), it results that:

$$e_{t+T} = S_{t+T} - F(T)_t \qquad (4)$$

or

$$S_{t+T} = F(T)_t + e_{t+T} \qquad (5)$$

and the equation (5) is the algebraic representation of the unbiasedness.

Based on equation (3), the expectation theorem argues that the difference between the futures price and the current spot price can be expressed as the sum of an expected premium and an expected change in the spot price (Fama, French, 1987):

$$F(T)_t - \mathrm{S_t} = E_t[P(T)_t] + E_t[S_{t+T} - S_t] \qquad (6)$$

where $F(T)_t$ is the delivery futures price at time t on a futures contract that expires at time T, and S_t is the spot price at time t.

The expected premium, $E_t[P(T)_t]$, is the bias of the futures price, $F(T)_t$, as a forecast of the future spot price, S_{t+T}:

$$E_t[P(T)_t] = F(T)_t - E_t[S_{t+T}] \qquad (7)$$

where $E_t[S_{t+T}]$ is the rational forecast, conditional on all information available at t.

When $E_t[P(T)_t] = 0$, we have $F(T)_t = E_t[S_{t+T}]$. This pricing relationship is called unbiasedness hypothesis and implies that the market is efficient and that futures price is an unbiased estimator of future spot price, or that the futures price should lead the spot price (Fama, 1983).

1.3.2. Spot-Futures Parity for a Commodity

The storage theorem, also called cost of carry theorem, describes the inter-temporal relationship between spot prices and futures prices of continuously storable commodities and provides the starting point when modelling the market efficiency.

In the cost-of-carry model the futures price is represented as (Hull, 2005):

$$F(T)_t = S_t e^{(r_t + w)(T-t)} \qquad (8)$$

where t is the current date, T is the futures contract maturity date,
w a storage cost (rent of storage space, insurance, physical deterioration or wastage),
r is the continuously compounded riskless rate of interest at t.

Intuitively, equation (8) suggests that arbitrage ensures that the future price of a commodity will be the same as the cost of borrowing funds, purchasing the commodity in the spot market and storing it over the borrowing period. Since interest rates and storage costs (together the 'cost-of-carry') are positive, this parity implies that the future price of commodity should be above the spot price. In this case, the market is in 'contango' (the futures curve slopes upward).

However, it happens that futures price is observed below the spot price (the futures curve slopes downward). In this case, the market is in 'backwardation', implying that the cost-of-carry of commodity is not the only determinant of the price of the future. An explanation which is often used to account for backwardation involves the notion of a 'convenience yield'. Convenience yield arises because holding the commodity in inventory can have productive value. Thus, convenience yield reflects the benefits that arise to the owner of a commodity but not to the owner of a contract for future delivery of the commodity (Hull, 2005).

If the convenience yield of holding the commodity is modelled as a premium which is included in the spot price, then equation above may be written as (Hull, 2005):

$$F(T)_t = S_t e^{(r_t + w - \varphi)(T-t)} \qquad (8')$$

For a sufficiently high value of the convenience yield, φ, it is clear that the futures price may lie below the spot price (backwardation).

The convenience yield is also thought of as the reflection of the market's expectations concerning the future availability of the commodity - the greater the probability for shortages, the higher the convenience yield and vice versa. Thus, the size of the convenience yield in the market is related to the level of inventories - when inventory levels are low (implying increased chance for shortages in the near future), the convenience yield will be higher than the cost of carry, and the basis, $F(T)_t - \mathrm{S_t}$, will be negative. On the other hand, when inventory levels are high (implying little chance of shortage), the convenience yield will approach zero, and the basis will be positive, having an upper limit on the cost of carry.

Fama and French (1987) argue that the theory of storage in $(8')$ and theory of expectations in (6) are alternative not competitive views, and that the variation in the expected premium or the expected change in the spot price in (6) translates into variation in the interest rate, the marginal storage cost, or the marginal convenience yield in $(8')$.

Taking into account the fact that for most futures, several contracts are traded at the same time, for a trader in the market, buying a futures contract with expiration at time T is similar as to buying a futures contract that expires at time $T - i$ and then store the commodity from $T - i$ to T. Hence, a similar relationship as in $(1')$ also holds for two futures contracts with different time to maturity (Asche, Guttormsen, 2002):

$$F(T)_t \geq F(T-i)_t e^{(r_t+w-\varphi)(T-i)} \qquad (9)$$

This condition would verify whether there are any long run relationships between futures prices of different maturities. If the futures price on longest contract forecasts the future spot price, it should also be the case that this will forecast the futures price of any other shorter contract.

1.4.Theory of Market Efficiency and Unbiasedness

If futures markets are to fulfill their price discovery function and provide forecasts of future spot prices, it is required that the markets are efficient and the risk premium is absent.
Financial markets are defined as efficient if prices fully reflect all available information including agents' expectations about the price movements in a way that no profit opportunities are left unexhausted. This implies that all emerging information should be immediately impounded into the expectations about future prices. Based on these expectations, agents would swiftly arbitrage away any deviations of the expected returns consistent with abnormal profits. Thus, no investor can earn extraordinary profits by predicting future prices on the basis of available information. This is known as the efficient market hypothesis (EMH).

The term 'all available information' suggests three version of the EMH: the weak, semi-strong and strong forms of the hypothesis. The weak-form hypothesis asserts that stock prices already reflect all information contained in the history of past prices, trading volume or short interest (Bodie, Kane, Markus, 2008). The semi-strong-form hypothesis goes further by stating that stock prices already reflect all publicly available information regarding the prospects of a firm. Such information includes, in addition to past prices, fundamental data on the firm's product line, quality of management, balance sheet composition, patents held, earning forecasts, and accounting practices (Bodie, Kane, Markus, 2008). The strong-form hypothesis stipulates that stock prices reflect all relevant information including insider information. (Bodie, Kane, Markus, 2008).

It is common to assume the equilibrium of prices to be characterized by a rational use of information. The advantage of this is that systematic errors in expectations are impossible. In this light, the EMH can be regarded as a joint hypothesis of rational expectations and risk neutrality. Thus, assuming risk-neutral and rational actors, the futures price close to delivery should represent the expected spot price when deliveries actually happen.

Fama (1970, 1991) contends that market efficiency per se is not testable and it must be tested jointly with some assets pricing model. According to the financial literature, the model establishing the idea that futures prices are unbiased estimators of future spot prices are the appropriate framework to test efficiency. Using this model (exposed in equations (6) and $(8')$), efficiency will necessarily imply that the market price fully reflects all available information and so there exists no strategy that traders can speculate in the futures market on the future levels of the spot price exploiting profits consistently.

Using the framework in equation (6) one can test for price forecasts in futures prices and time-varying expected premiums. The relevant test consists in running the regressions of the change in the spot price and the premium on the basis (Fama, 1984):

$$S_{t+T} - S_t \quad = \propto_1 + \beta_1 \, [F(T)_t - S_t] + e^c_{t+T} \qquad (10)$$

$$F(T)_t - S_{t+T} = \propto_2 + \beta_2 \, [F(T)_t - S_t] + e^P_{t+T} \quad (11)$$

where $[F(T)_t - S_t]$ is the basis at time t, and $[F(T)_t - S_{t+T}]$ is the premium; S_{t+T} is the observed spot price at time $t + T$ and $F(T)_t$ is the futures price at time t for a contract expiring at T; finally e^c_{t+T} and e^P_{t+T} are residual terms.

The evidence that β_1 is significant means the basis observed at t contains information about the change in the spot price from t to $t + T$. That is, the futures price has power to forecast the future spot price. Evidence that β_2 is significant means the basis observed at t contains information about the premium to be realized at $t + T$. That is, predictable variation in realized premiums is evidence of time-varying expected premiums.

The cost-of-carry model in $(8')$ provides another starting point when modelling market efficiency. However, in practice, it is difficult to test the arbitrage relationship embodied in $(8')$ due to the unobservable nature of convenience yields in the oil markets. Hence, most studies have employed the Fama's (1970) weak form market efficiency, also called speculative market efficiency tests of the form:

$$S_{t+T} = \propto + \beta \, F(T)_t + \varepsilon_{t+T} \quad (12)$$

In this approach, market efficiency requires that futures prices should be unbiased predictors of future spot prices. Otherwise, risk-neutral speculators could make consistent profits on long or short futures positions through time. In this specification, market efficiency, in the absence of a risk premium, requires that the constant term to be zero and the slope coefficient to be unity. Thus, simple empirical tests of the speculative efficiency hypothesis are based on tests of the joint hypothesis $\propto = 0$ and $\beta = 1$ in (12).

2 METHODOLOGY FRAMEWORK

2.1. Unit Root Test

Before running any time series regression, one should examine the properties of the variables. It is empirically crucial to test the data for unit roots; that is, whether a data sequence has time dependent mean, variance, and co-variance (Wooldridge, 2006). When a system is non-stationary, the shocks to the system are persistent and will not die away over time. According to Hendry and Juselius (1999), unit-root process is a sort of stochastic non-stationary process induced by persistent combinations of past effects. Possible reasons why variables may contain unit roots are technical progress, political turmoil, policy regime changes, *inter alia*. These often lead to structural breaks in the time-series providing hazardous analysis with meaningless conclusions (Hendry, Juselius, 1999).

If stationarity is not a realistic characterization of data, then any emerging regression results will be spurious; that is, apparently significant regression will result from unrelated data (Engle, Granger, 1987). Yet, there is an exception, meaning that non-stationary data can be used in regression and still get meaningful results.

Usually, if y and x are non-stationary, I(1), variables, then it will be expected that any difference, or any linear combination of them, like $e_t = y_t - \alpha - \beta x_t$ are non-stationary I(1) as well. However, there is an important case when the unit roots in y and x 'cancel each other out'; and e or their linear combination is stationary I(0). In this case, y and x are cointegrated and the spurious regression problem disappears. This is because co-integration implies that y and x share a similar stochastic trend; and, since the difference e_t is stationary, they never diverge too far from each other. From this point of view, unit root test is a part of co-integration test which is the core of the market efficiency test.

The most used method for unit root testing is the Augmented Dickey-Fuller (ADF) which is based on the regression:

$$\Delta y_t = \propto + \lambda t + \beta y_{t-1} + \sum_{i=1}^{lag} c_i \, \Delta y_{t-i} + v_t \quad (13)$$

The equation tests the null hypothesis of a unit root $(H_o: \beta = 1)$ against a stationary alternative $(H_1: \beta < 1)$.

One important underlying assumption is that error term has:

- zero-mean: E $[\varepsilon_t]$
- constant variance: Var $[\varepsilon_t] = \sigma^2$
- uncorrelated residuals: Cov $[\varepsilon_t; \varepsilon_{t-1}] = 0$
- normally distributed residuals: $\varepsilon \sim N(\mu, \sigma^2)$.

One critical decision is the lag length. If the model does not have enough lag-terms to capture full dynamics in the process, error autocorrelation is likely to occur. In order to 'whiten' the residuals, that is, reduce their autocorrelation, additional lags have to be introduced (Wooldridge, 2006; Doornik and Hendry, 2007). The inclusion of too many lags reduces the power of the test - more lags are introduced, more of the initial observations are lost. The inclusion of too few lags will result in a misspecification problem (Wooldridge, 2006). A simple method to determine the number of lags is the examination of the autocorrelation function of the residuals or the significance of the estimated lag coefficients c_i. Another methods for deciding on the lag length are Akaike information criterion (AIC) and Schwartz Bayesian criterion (SBC) (Enders, 1995):

$$AIC = T\ ln\ (residual\ sum\ of\ squares) + 2n$$

$$SBC = T\ ln\ (residual\ sum\ of\ squares) + n\ ln(T)$$

n – number of parameter estimated

T- number of usable observations

The idea is to choose between models with different lag length over the same sample period the model with the smallest information criterion. Increasing the number of regressors increases n, and has the effect of reducing the residual sum of squares. Thus, if a regressor has no explanatory power, being added to the model will cause both the AIC and SBC to increase.

Another critical decision is the inclusion of the intercept and the trend. Including a constant and a trend ensures that the test will have the correct rejection frequency under the null hypothesis (Wooldridge, 2006). The inclusion of the constant and the trend is justified by economic judgment (or common sense) and statistical significance.

ADF tests have very low power to discriminate between alternative hypotheses, and are not valid when the data have jumps or structural breaks in the data generation process. The errors are assumed to i.i.d., which very often not the case (Alexander, 2008). A price jump increases the probability of a type I error, i.e. that a true null hypothesis will be rejected.

The limitation of the ADF is that it assumes a linear deterministic trend to account for the upward trend in the economic variables. However, economic time series often exhibit changes in the trend when major economic events such as oil crises or financial crunches occur. Perron (1989) claimed that one will often conclude non-stationarity of time series in a model which ignores the breaks in the time trend even though, in reality, this follows a stationary stochastic process around a trend with a break.

The most important implication of the unit root tests is that under the null hypothesis random shocks have a permanent effect on the system. This runs counter to the general belief that business cycles are transitory fluctuations around a more or less stable trend path (Perron, 1989). His main conclusion is that most macroeconomic time series are not characterized by the presence of unit root and that fluctuations are indeed transitory.

I base my test on Perron's method (1989) for testing unit root, considering the null hypothesis that a time series has a unit root with possibly nonzero drift against the alternative that the process is "trend-stationary" and allowing under both null and alternative hypotheses for the presence of a one-time change in the level and/or in the slope of the trend function.

Perron (1989) asserted that only certain "big shocks" have had permanent effects on the various macroeconomic time series and that these shocks were exogenous – that is, not a realization of the underlying time-invariant stochastic process (Perron, 1989; Serletis, 2007). Modeling such shocks as exogenous, removes the influence of these shocks from the null hypothesis of a unit root. Therefore, the null should be tested against the trend-stationary alternative by allowing, under both the null and the alternative hypotheses, for the presence of a one-time break (at a known point in time) in the intercept and/or in the slope of the trend function (Perron, 1989; Serletis, 2007).

If the shocks/breaks in the series are known, then it is relatively simple to adjust the ADF test by including dummy variables to ensure there are as many deterministic regressors as there are deterministic components (Harris, R., Sollis, R. 2003). The critical values for unit root tests involving changes in the intercept and/or trend are the ones found in Perron's articles (1989, 1990). However, it is unlikely that the date of the break will be known a priori, as assumed by Perron (1989).

Three different models may be considered:

Model A - "crash model" - allows for a shift in the intercept of the deterministic trend function. The null hypothesis of a unit root is characterized by a dummy variable which takes the value one at TB and zero otherwise. Under the alternative hypothesis of a "trend-stationary" system, model A allows for a one-time change in the intercept of the trend function.

Model B - "changing growth model" - allows a shift in the slope of the trend function. Under the alternative hypothesis, a change in the slope of the trend function without any sudden change in the level at the time of the break is allowed. Under the null hypothesis, the model specifies that the drift parameter μ changes from μ_1 to μ_2 at the time TB.

Model C - "crash/changing growth model" - allows for both effects to take place simultaneously, i.e., a change in the level followed by a different growth path, such as productivity slowdown.

Using the nomenclature of Perron (1989), the null hypotheses are parameterized as follows:

Model (A) $$y_t = \mu + y_{t-1} + dDTB_t + e_t \quad (14.a)$$

Model (B) $$y_t = \mu_1 + y_{t-1} + (\mu_2 - \mu_1)DU_t + e_t \quad (14.b)$$

Model (C) $$y_t = \mu_1 + y_{t-1} + dDTB_t + (\mu_2 - \mu_1)DU_t + e_t \quad (14.c)$$

Where $DTB_t = 1$ if $t = TB + 1$, 0 otherwise

$DU_t = 1$ if $t > TB + 1$, 0 otherwise

And the alternative hypotheses are Perron (1989):

Model (A) $$y_t = \mu_1 + \beta t + (\mu_2 - \mu_1)DU_t + e_t \quad (15.a)$$

Model (B) $$y_t = \mu + \beta_1 t + (\beta_2 - \beta_1)DT_t^* + e_t \quad (15.b)$$

Model (C) $$y_t = \mu_1 + \beta_1 t + (\mu_2 - \mu_1)DU_t + (\beta_2 - \beta_1)DT_t + e_t \quad (15.c)$$

Where $DT_t^* = t - TB$ if $t > TB$, 0 otherwise

$DT_t = t$ if $t > TB$, 0 otherwise

TB refers to the time of break, TB $(1 < TB < T)$, i.e. the period at which the change in the parameters of the trend function occurs. Note that the dummy variables are linked through $DTB_t = \Delta DU_t = \Delta^2 DT_t$ (Perron 1989).

The null and alternative hypotheses can be nested as in the following regression equations:

Model (A)

$$\Delta y_t = \mu^A + \theta^A DU_t + \beta^A t + d^A D(T_B)_t + \alpha^A y_{t-1} + \sum_{i=1}^{k} c_i \, \Delta y_{t-i} + e_t \qquad (16.a)$$

Model (B)

$$\Delta y_t = \mu^B + \theta^B DU_t + \ \beta^B t + \gamma^B DT_t^* + \alpha^B y_{t-1} + \sum_{i=1}^{k} c_i \, \Delta y_{t-i} + e_t \qquad (16.b)$$

Model (C)

$$\Delta y_t = \mu^C + \theta^C DU_t + \beta^C t + \gamma^C DT_t + d^C D(T_B)_t + \ \alpha^C y_{t-1} + \sum_{i=1}^{k} c_i \, \Delta y_{t-i} + e_t \ (16.c\ )$$

The null hypothesis imposes the following restrictions on the true parameters of each model:
Model (A): $\alpha^A = 1, \theta^A = 0,\ \beta^A = 0$
Model (B): $\alpha^B = 1, \gamma^B = 0, \beta^B = 0$
Model (C): $\alpha^C = 1, \gamma^C = 0, \beta^C = 0$

Under the alternative hypothesis of a trend stationary process, it is expected that
$\alpha^A, \alpha^B, \alpha^C < 1;\ \beta^A,\ \beta^B,\ \beta^C \neq 0;\ \ \gamma^A,\ \gamma^B,\ \gamma^C \neq 0$
And, d^A, d^C and θ^B should be close to zero.

2.2. Co-Integration Test

As mentioned in the above section a system of variables, y and x, is defined as co-integrated if a linear combination of them, like $e_t = y_t - \alpha - \beta x_t$, is stationary, I(0). In this case, the non-stationary time series, y and x, share similar stochastic trends and never diverge too far from each other over time. That is also to say there exists some influences (market forces) implying that these series are bound by some long-term relationship. Hence, a co-integrating relationship may also be thought of as an equilibrium relationship where the co-integrating variables may deviate from their relationship in the short run, but they will always have a constant mean they steadily return to in the long run.

Thus, the goal of co-integration analysis is to test whether there are any common stochastic trends or any equilibrium relationship between non-stationary variables - if there is a common trend in a set of variables they must have a long term equilibrium relationship.

There are different empirical techniques for co-integration analysis. The most common co-integration methodologies exposed and applied in this paper are the Engle-Granger (1987) and Johansen (1988, 1991) methodologies. The first one is based on an OLS regression, while the second is based on characteristics roots (eigenvalues) analysis of a certain matrix. In this paper I applied the bivariate Johansen methodology, while the Engle-Granger methodology is exposed only to provide an intuitive understanding of the co-integration analysis.

Engle- Granger Methodology (EGM)

EGM is based on an OLS regression and applies a unit root test to the residuals of the regression. EGM is the only case when it is legitimate to perform an OLS analysis on non-stationary and co-integrated variables and get consistent estimators. If non-stationary dependent and independent variables are not co-integrated then OLS will provide inconsistent estimates and spurious conclusions.

The test of co-integration involves two steps:

1. Establish a long-term relationship between variables by running the regression:

$$y_t = \alpha + \beta x_t + e_t \qquad (17)$$

The equilibrium error, $Z_t(= e_t \ = y_t - \beta x_t - \alpha)$, captures the random deviations from the long run equilibrium. The co-integrating vector is the vector of coefficients in Z_t. So in this case the co-integrating vector is $(1, -\beta, -\alpha)$.

2. Co-integration test consists in an ADF unit root test on the residual:

$$\Delta\hat{e}_t = \ \phi\hat{e}_{t-1} + \sum_{i=1}^{p-1} \phi_i \Delta\hat{e}_{t-i} + v_t \qquad (18)$$

where $v_t \sim IID(0, \sigma^2)$.
The null hypothesis for co-integration test is whether the linear combination is non-stationary $e_t \sim I(1)$, against the alternative hypothesis $e_t \sim I(0)$.

The question of the inclusion of constant and/or trend terms in equation (18) depends on whether a constant and/or trend term appears in (17) That is, deterministic components can be added to either (17) or (18), but not both (R. Harris, R. Sollis, 2003).

Engle and Granger's (1987) two-step co-integration procedure has several limitations. First, no strong statistical inference can be drawn on the OLS coefficients α and β. Second, the co-integrating vector is assumed to be unique. However, when there are more than two variables, the uniqueness of the co-integration vector cannot be assured using the two-step co-integration procedure (Enders, 2004). Another limitation is that the single equation ECMs are only valid given an exogeneity assumption (Enders, 2004). However, this is what one might want to test.
Another important limitation is that the EGM co-integration two step procedure cannot be used to test restrictions on coefficients, as the test procedure does not have well defined limiting distributions.
Johansen's maximum likelihood method provides solutions to these problems.

Johansen Methodology (JM)
JM investigates co-integration in a multivariate system where there are at least two integrated variables. It is a maximum likelihood test for the presence of multiple co-integrating vectors and allows the testing of a restricted version of the co-integrating vectors.

Let X_t be denoted as an $n \times 1$ vector of the I(1) variables, for instance, a set of (log) prices. The underlying hypothesis is that X_t follows an unrestricted vector auto-regression (VAR) in the levels of the variables (Alexander, 2008):

$$X_t = \Phi \mathrm{D_t} + \mathrm{B}X_{t-1} + \varepsilon_t \qquad (19)$$

Where $\mathrm{D_t}$ contains deterministic terms (constant, trend, dummies), $X = \{X_1, \dots X_n\}$, and $\varepsilon_t \sim IID(0, \sigma^2)$, and B is a vector of slope coefficients.

Or, equivalently, subtracting X_{t-1} from both sides, the VAR system can be expressed as (Alexander, 2008):

$$\Delta X_t = \Phi \mathrm{D_t} + \Pi X_{t-1} + \varepsilon_t \qquad (20)$$

where $\Pi = \mathrm{B} - \mathrm{I}$ and I is the $n \times n$ identity matrix.

This may be augmented with sufficient lagged dependent variables to remove the autocorrelation in residuals (Alexander, 2008):

$$\Delta X_t = \Phi \mathrm{D_t} + +\Pi X_{t-1} + \Gamma_1 \Delta X_{t-1} + \ \dots + \Gamma_q \Delta X_{t-q} + \varepsilon_t \qquad (21)$$

where $\Gamma_i = -(I - \Pi_1 - \cdots - \Pi_i)(i = 1, \dots k - 1)$ and they are the short run impact matrices, while Π is the long run impact matrix that contains information about the long-run relationship between the series and lends itself to hypothesis testing.The rank of Π, r, determines how many linear combinations of X_t are stationary.

ΔX_t and its lags are I(0). The term ΠX_{t-1} is the only one which might include I(1) variables and for ΔX_t to be I(0) it must be the case that ΠX_{t-1} is also I(0) (Alexander, 2008). Therefore, ΠX_{t-1} must contain the co-integrating relationships if they exist.

The condition that ΠX_{t-1} must be stationary implies nothing about the relationships between $\{X_1, \dots X_n\}$, if the rank of the matrix is zero. That is to say, if $r = 0$ so that $\Pi = 0$, $X_t$is I(1) and not co-integrated, thus none of the linear combinations are stationary.

If $r = n$, all variables in levels (or logs) are stationary.

In the intermediate case where $0 < r < n$, there exist r linearly independent co-integrating vectors (or, r stationary linear combinations of$\{X_1, \dots X_n\}$), and $n - r$ common stochastic trends (unit roots).

Thus, the test for co-integration is a test on the rank of Π, and the rank of Π is the number of co-integrating vectors.

In this case where $0 < r < n$, one can factor Π as $\Pi = AB'$. Both A and B are $n \times r$ coefficient matrices, where the rows of B' contain the cointegrating vectors (the error correcting mechanism in the system); and the elements of A contains the factor loadings that distribute the impact of the co-integrating vectors to the evolution of ΔX_t, or, more straightforward they measure the speed of convergence to the long-run steady state.

JM suggests two tests for the number of co-integration vectors in the system- maximum characteristic roots (eigenvalues) test and the trace test (Enders, 2004):

$$\lambda_{trace}(r) = -T \sum_{i=r+1}^{n} \ln\,(1 - \hat{\lambda}_i) \qquad (22.a)$$

$$\lambda_{max}(r, r+1) = -T \ln\big(1 - \hat{\lambda}_{r+1}\big) \qquad (22.b)$$

where, $\hat{\lambda}_i$ is the estimated value of the characteristic roots obtained from the estimated Π matrix.

Both tests have null hypothesis that there are at most r cointegration vectors. For the maximum eigenvalue test, the alternative hypothesis is that there are exactly $r + 1$ cointegartion vectors, while the alternative hypothesis in the trace test is that there exist more than r co-integrating vectors (Enders, 2004).

Critical values of these statistics are given in Johansen and Juselius (1990). They depend on number of lags in (21) and whether the model includes a constant and /or a trend.

Deterministic terms play a crucial role in both data behaviour, distributions of estimators and tests in an integrated process (Hendry, Juselius, 2000). Depending on their presence or absence, the system may manifest drift or linear trends in co-integrating vectors. An appropriate formulation of the model is important to ensure that co-integrating rank tests are not too dependent on 'nuisance parameters' related to the deterministic terms (Doornik, Hendry, 2007).

Johansen's maximum likelihood (ML) estimation has some advantages relative to the EG two-step procedure. First, the JM provides more efficient estimates of the co-integrating relationship and therefore more detailed inference can be drawn from parameters A and B. Second, JM allows a wide range of hypothesis/restrictions on the coefficients A and B, using likelihood ratio tests. Third, the procedure allows all the distinct co-integrating vectors to be identified and does not impose a specific number of co-integration relationships a priori. This implies that tests of the number of co-integration relationships are carried out simultaneously (Enders, 2004). Lastly, JM is shown to be fairly robust to the presence of non-normality and heteroscedasticity disturbances.

Though JM has advantages relative to the EGM for testing for co-integration, limitations may also arise. One inherent problem with JM is the inability to test for or assess the short-term dynamics. A solution to the problem associated with the JM is offered by error-correction model (ECM). It affords co-integration theory to reconcile the long run equilibrium with short run dynamics in a system of variables.

2.3. Error Correction Model

So far, the co-integration was referred as to the idea of I(1) variables trending together or bearing a long run equilibrium relationship to each other. However, it may be insightful to understand the short-run behaviour.

Statistically, in a (bivariate) co-integrated system with Y_t and X_t variables, $Z_t = Y_t - X_t$ cannot deviate too far from the trend line, or the long-term equilibrium, but this does not exclude short-term deviations. For example, if the futures price for a commodity moves "too far" from the equilibrium level, buyers and sellers may engage in arbitrage so that futures price will return to its long-term equilibrium.

Engle-Granger Methodology (EGM)

Next step in the Engle-Granger (EG) approach after establishing the long run equilibrium relationship of integrated variables is to assess how short term deviations from equilibrium are corrected. For this purpose error correction model (ECM) is used, which is a representation of vector auto-regression model (VAR).

$$\Delta Y_t = \Phi_1 \mathrm{D_t} + \theta_1 Z_{t-1} + \sum_{i=1}^{m} \gamma_{11}^{i} \Delta X_{t-i} + \sum_{i=1}^{m} \lambda_{12}^{i} \Delta Y_{t-i} + \varepsilon_{1t} \quad (23)$$

$$\Delta X_t = \Phi_2 \mathrm{D_t} + \theta_2 Z_{t-1} + \sum_{i=1}^{m} \gamma_{21}^{i} \Delta X_{t-i} + \sum_{i=1}^{m} \lambda_{22}^{i} \Delta Y_{t-i} + \varepsilon_{2t} \quad (24)$$

Thus, ECM captures information flows through two channels: the lagged error-correction term, $Z_{t-1} = Y_{t-1} - \beta X_{t-1} - \alpha$, and the lagged difference terms. The Z_{t-1}term is embedding partly the long run properties through the long-run multiplier, β; and, partly the short run properties as being the equilibrium error term. Further properties of short-run behavior are captured by the inclusion of lagged explanatory variable, implying that if X changes, the equilibrium value of Y will also change.

To estimate the ECM we can apply OLS to each equation separately (Alexander, 2008). And, at least one of the coefficients θ_1 and θ_2 must be significant, otherwise the variables would not be co-integrated. The magnitude of the coefficient estimates θ_1 and θ_2 determines the speed of adjustment back to the long term equilibrium following an exogenous shock. When these coefficients are large, adjustment is quick so Z will be highly stationary and reversion to the long term equilibrium determined by $E(Z)$ will be rapid (Alexander, 2008).

In case of $\beta > 0$ in $Z_{t-1} = Y_{t-1} - \beta X_{t-1} - \alpha$, the model will have an error correction mechanism only if $\theta_1 < 0$ and $\theta_2 > 0$. That is so because in this case the error correction term will constrain deviations from the long term equilibrium in such a way that errors will be corrected. If Z is large and positive, then Y will decrease because $\theta_1 < 0$ and X will increase because $\theta_2 > 0$, both have the effect of reducing Z and this way error are corrected. If Z is large and negative, then Y will increase because $\theta_1 < 0$ and X will decrease because $\theta_2 > 0$, both have the effect of increasing Z correcting the error (Koop, 2008).
In case of $\beta < 0$, the model will have an error correction mechanism only if $\theta_1 < 0$ and $\theta_2 < 0$.

Granger Causality
Once the ECM is specified it may be used to model the lead-lag behaviour between variables in a system of co-integrated variables as a way of inferring price dominance. The test of lead-lag relationship between variables can be referred to as Granger causality test. Specifically, one can say that X *Granger causes* Y if lagged values of X help to predict current and future values of Y better than just lagged values of Y alone (Alexander, 2008). Used in this way they are not meant to imply causality in its true sense, but rather to indicate temporal relations between variables.

The test for Granger causality from X to Y is a test for the joint significance of all the variables containing lagged X in equation (1), and the test for Granger causality from Y to X is a test for the joint significance of all the variables containing lagged X in equation (2). That is, (Alexander, 2008):
X Granger causes Y when $H_0: \beta_{12}^1 = \beta_{12}^2 = \cdots = \beta_{12}^m = \theta_1 = 0$ is rejected
Y Granger causes X when $H_0: \beta_{21}^1 = \beta_{21}^2 = \cdots = \beta_{21}^m = \theta_2 = 0$ is rejected
Thus, the null hypothesis is that Granger causality does not occur (H_0: $\beta = 0$).
The parametersβ_{12} and β_{21} provide information about the flow of information between two variables and the parameters $\theta_1$and θ_2, contain information about exogeneity.

3 DATA DISCRIPTION

3.1. Data Construction

As the purpose of this paper is twofold, two sets of variables are needed. For the first task – analysis of the efficiency of crude oil futures markets – spot and futures time series are required – namely, data on spot price and futures prices. For the second task – analysis of the relationship between traders' positions and futures prices – data on the crude oil traders' futures positions are needed.

The first set of time series – WTI spot and futures prices – are in monthly format and spans over a period of about 18 years: January 1991- October 2008[7] (214 observations). The spot and futures closing prices were obtained from Datastream. While the spot price time series is easy to obtain, futures price time series needs to be constructed. To construct futures price I used the method suggested by Gjølberg. Given that there are 12 futures contracts opened each year; over the last 18 years we have about 216 (18×12) futures contracts with different maturity dates to base the main futures price time series. A contract is often open for several months, and the subsequent daily futures prices reflect the changing market expectation of what the spot price will be on the last day of trading. The matching futures prices were sampled from a specific day (21th every month[8]), less than one (three, six) month(s) from the last day of trading. Then the futures price is selected by working backward from 20th to 21st every month for every contract[9]. For instance, if a futures contract expires in January 2009, the relevant prices for one-month futures price span over the period 21.10.2008-20.11.2008, for a three-month futures price span over the period 21.08.2008- 19.09.2008, for a for a six-month futures price span over the period 21.05.2008-20.06.2008. This is also illustrated in exhibit1 for contracts expiring February, March and April 2009.

Exhibit 1 Illustration of Futures Price Construction

Contract expiration date	One-month futures price	Three-month futures price	Six-month futures price
January 2009	21.10.2008-20.11.2008	21.08.2008- 19.09.2008	21.05.2008-20.06.2008
February 2009	21.11.2008-19.12.2008	22.09.2008-20.10.2008	23.06.2008-18.07.2008
March 2009	22.12.2008-20.01.2009	21.10.2008-20.11.2008	21.01.2008-20.08.2008
April 2009	21.01.2009-20.02.2009	21.11.2008-19.12.2008	21.08.2008-19.09.2008

If matching futures prices are not sampled as described above, the time series analysis will suffer from autocorrelation problems because of informational overlap (Hansen, Hodrick,

[7] The data is until October 2008, due to availability.
[8] The 20th every month is the last trading day for a crude oil futures contract
[9] 22nd or 23rd is used if 21st not available and 18th or 19th is used if 20th is not available

1980). Consequently, autocorrelation in the errors of the usual regression equation for testing efficiency might induce the appearance of inefficiency even in efficient markets.

Another set of variables – crude oil traders' futures positions – are in weekly format (as of Tuesday's close) and spans over January 1993 –October 2008 (826 observations). These were collected from COT reports available on CFTC web site[10]. A matching set of futures returns is calculated for one, three and six-month futures prices[11].

In relating traders' positions to market returns, there are two relative measures of position size. The first is simply the percent of the total open interest held by each CFTC trader classification. This measure is the sum of the long and short positions held by the trader class divided by twice the market's total open interest (Sanders, 2004):

$$Reporting\ non{-}commercials'\ percent\ of\ TOI_t = \frac{NCL_t + NCS_t + (2NCSP_t)}{2(TOI_t)}$$

$$Reporting\ commercials'\ percent\ of\ TOI_t = \frac{CL_t + CS_t}{2(TOI_t)}$$

$$Non{-}reporting\ traders = \frac{NRL_t + NRS_t}{2(TOI_t)}$$

The second measure captures the net position of the average trader in a CFTC classification. The percent net long (PNL) position is calculated as the long position minus the short position divided by their sum (De Roon et al., 2000):

$$Reporting\ non{-}commercials\ PNL = \frac{NCL_t - NCS_t}{NCL_t + NCS_t + (2NCSP_t)}$$

refered to as "speculative pressure".

$$Reporting\ commercials\ PNL = \frac{CL_t - CS_t}{CL_t + CS_t}$$

refered to as "hedging pressure".

$$Non{-}reporting\ traders = \frac{NRL_t - NRS_t}{NRL_t + NRS_t}$$

refered to as "small trader pressure".

Thus, the PNL for each CFTC classification represents the net position held by the group normalized by its total size.

[10] http://www.cftc.gov/marketreports/commitmentsoftraders/cot_historical.html

[11] Same principle as above described was used here

3.2. Summary Statistics

Exhibit 2 displays the summary statistics for spot and futures crude oil geometric returns including the mean, standard deviation, skewness, excess kurtosis coefficients and correlation coefficients. From these we can infer the nature of the returns distributions.

Exhibit 2 Summary Statistics - Spot and Futures Prices – 1991 (1) – 2008 (10)

	Spot Price	One-month Futures Price	Three-month Futures Price	Six-month Futures Price
Mean	5.78 %	6.00 %	6.46 %	6.98 %
Std.Devn.	33.21 %	31.52 %	27.83 %	23.91 %
Minimum	-34.72 %	-42.77 %	-38.87 %	-30.55 %
Maximum	25.07 %	21.54 %	21.31 %	20.39 %
Skewness	-0.597	-0.958	-1.029	-0.891
Excess Kurtosis	0.648	2.154	2.814	2.586
Normality test:	12.11[0.002]**	26.1 [0.00]**	29.0 [0.00]**	25.11 [0.00]**

NOTE: Sample period: 1991(1) - 2008(10). Monthly observations: 214
Summary Statistics is calculated for monthly log returns and then annualized:

$$lnR\ (S) = ln(S_t/S_{t-1})\ ,\ lnR(Fn) = ln(Fn_t/Fn_{t-1})\ ;$$
$$lnR(ann.) = lnR(mont) * 12$$
$$St.Dev\ R\ (ann.) = St.Dev.R\ (mnth) * \sqrt{12}$$

The normality test is a Jarque-Bera test with the null hypothesis of normal distribution.

The annual mean returns are within the range of 5.78% and 6.98%. And the annual standard deviation is greatest for spot 33.21% and lowest for six-month futures, 23.91%. The excess kurtosis coefficient for spot prices is slightly greater than zero, indicating on slightly fatter tails distribution than normal distribution would suggest. For futures prices the excess kurtosis is greater than that of spot, as expected. The skewness of the distribution is negative though small in magnitude, revealing a distribution skewed to the left. The Jarque-Bera normality test confirms the evidence provided by excess kurtosis and skewness; that is, the null hypothesis of normal distribution is rejected at all conventional levels.

The summary statistics for the second set of variables – traders' market positions, as measured by the percent of the total open interest - is presented in exhibit 3.

Exhibit 3 Summary Statistics - Traders' Percent of Total Open Interest

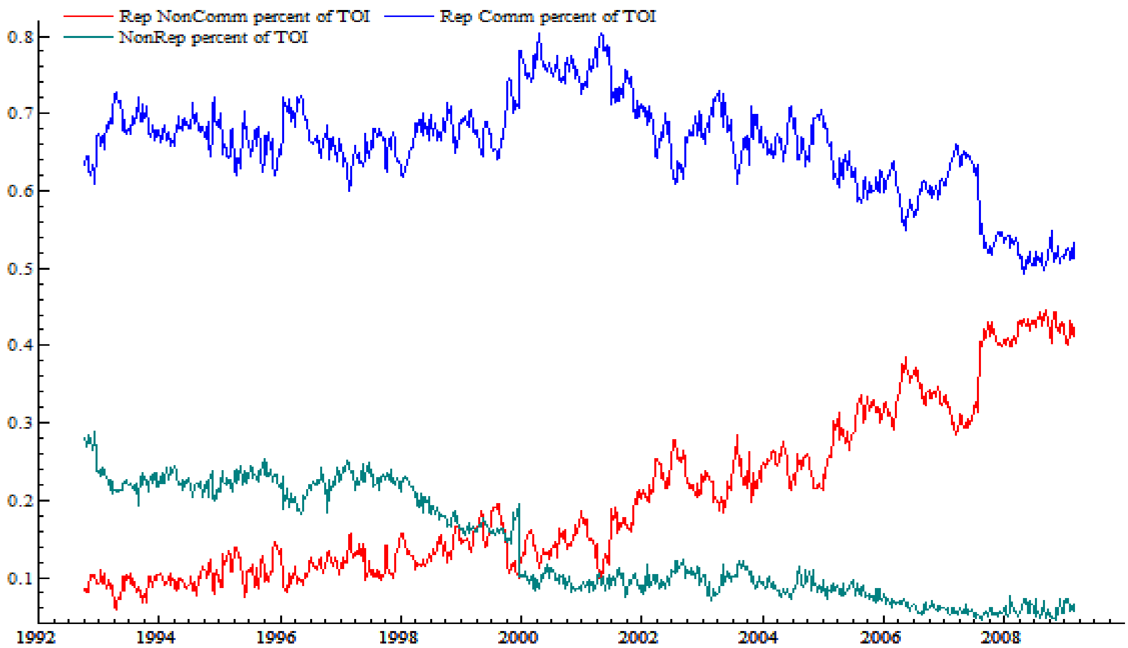
	Rep.NonComm	Rep. Comm	Nonrep.		Rep. NonComm	Rep.Comm	Nonrep.
Mean	19.92 %	66.19 %	13.89 %	Mean	11.66 %	67.29 %	21.05 %
Std.Devn.	10.07 %	6.00 %	6.74 %	Std.Devn.	2.62 %	2.56 %	2.62 %
Minimum	5.90 %	49.34 %	4.55 %	Minimum	5.90 %	60.06 %	10.22 %
Maximum	44.65 %	80.35 %	25.27 %	Maximum	19.57 %	78.08 %	25.27 %

NOTE: Sample period: 1993(1) - 2008(10)
Weekly observations: 826
All series are stationary at the 5% level

NOTE: Sample period: 1993 (01) - 1999(12)
Weekly observations: 364

From exhibit 3 it can be concluded that reporting commercial traders are the largest position holders in the crude oil futures markets, holding 66.2% of the open interest. The next largest group for the whole sample period is reporting non-commercials, holding about 20% of the open interest. The smallest group is the non-reporting traders, holding about 14% of the open interest. Last year open interest has fallen by about 17 % as financial investors have been forced to liquidate positions in response to deepening global economic and financial problems. A better illustration of the results in exhibit 3 is displayed in exhibit 4.

Exhibit 4 Traders' Positions as a Percent of Total Open Interest - 1993(1) - 2008(10)



NOTE: Algebra code: TOI - Total Open Interest

```
Rep.Comm. Percent of TOI = (CommLong+CommShort)/(2*TOI)
Rep.NonComm Percent of TOI = (NonCommLong+NonCommShort+2*NonCommSpread)/(2*TOI)
NonRept Percent of TOI = (NonReptLong+NonReptShort)/(2*TOI)
```

From the exhibit 4 it is obvious that the relative size of each trader category changes through time – after 2000 the reporting non-commercials increased their share of the market, holding about 40% of the open interest, compared to 11.6% prior to 2000. The non-reporting positions were holding about 20% (on average) prior 2000 compared to less than 10% (on average) after 2000. The reason for these trends could stem from increases in speculative limits during the late 1990s and the general growth in managed futures. The market share of the reporting commercials decreased from about 65% on average prior to 2000 to about 55% on average after 2000.

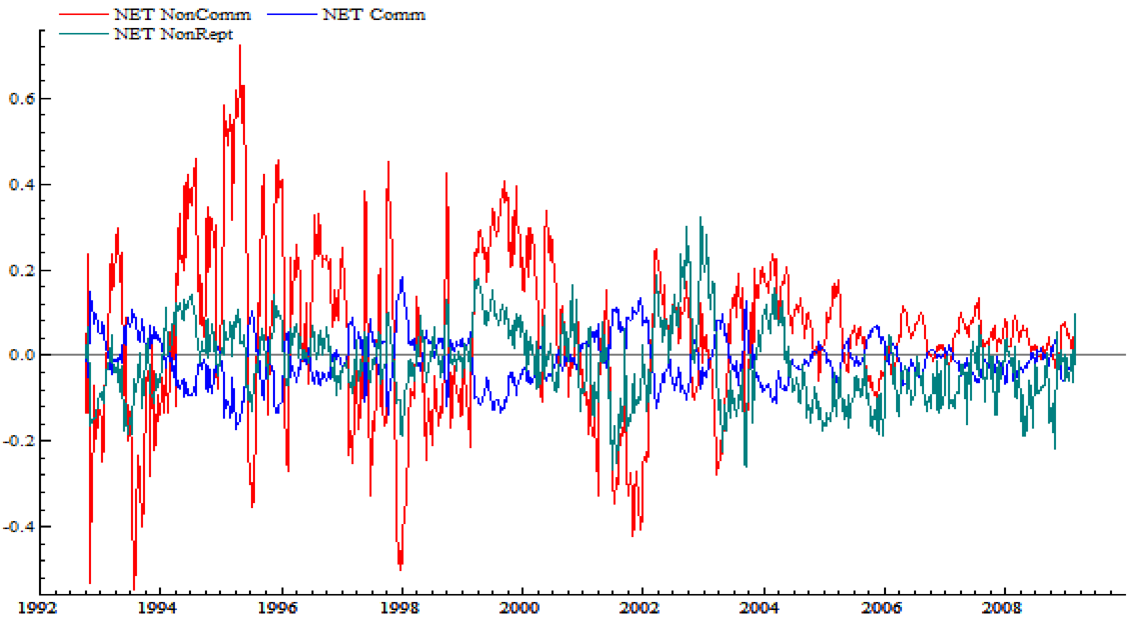
Lastly, the summary statistics for net market positions, as measured by the percent net long traders' positions is presented in exhibit 5.

Exhibit 5 Summary Statistics - Traders' Percent Net Long Positions

	Rep. NonComm	Rep. Comm	Nonrep.		Rep NonComm	RepComm	Nonrep.
Mean	5.02 %	-1.30 %	-1.66 %	Mean	8.25 %	-1.74 %	0.95 %
Std.Devn.	19.37 %	5.60 %	9.11 %	Std.Devn.	24.87 %	6.57 %	7.19 %
Minimum	-54.57 %	-17.09 %	-26.80 %	Minimum	-54.57 %	-17.09 %	-18.61 %
Maximum	72.57 %	18.20 %	32.21 %	Maximum	72.57 %	18.20 %	17.88 %
NOTE:	Sample period: 1993(1) - 2008(10) Weekly observations: 826 All series are stationary at the 5% level			**NOTE:**	Sample period: 1993 (01) - 1999(12) Weekly observations: 364		

From exhibit 5 it can be seen that the reporting commercials hold net short positions, while reporting commercials hold net long positions. This is also true for shorter sample periods. However, the non-reporting traders in the period 1993-2008 hold net short positions, while in period 1993 -2000 they hold long positions. These data suggest that crude oil hedgers are traditional short hedgers usually associated with production hedgers. Furthermore, the positions held by all categories are volatile, with each group shifting from net long to net short over the sample period. The most volatile group is the reporting non-commercials where the PNL can reach extremes greater than -50% (short) and greater than 70% (long). The volatility of the non-commercials' is clearly illustrated in the exhibit 6.

Exhibit 6 Traders' Percent Net Long Positions - 1993(1) - 2008(10)



NOTE: `Algebra code: (Percent Net Long (PNL)`

```
Rep.Comm. PNL=(CommLong-CommShort)/(CommLong+CommShort)
RepNonCommpercentPNL=(NonCommLong-NonCommShort)/(NonCommLong+NonCommShort+2*NonCommSpread)
NonRept PNL=(NonReptLong-NonReptShort)/(NonReptLong+NonReptShort)
```

From this exhibit 6, it is evident that non-commercials, though a relative small category of the total market, must be active traders who are changing from long to short positions over the week. The volatility of each category's net position indirectly reveals information about the diversity of motives within each group. It seems that the least diverse set of motives exists for noncommercial traders. In fact, the data suggest that traders in this group largely act in concert, relative to traders in other groups. Thus, it is not surprising that they are thought to influence the market. This proposition is explicitly tested in a later section.

The results of the data description are consistent with the fact that innovations in the derivatives markets in recent years have allowed financial investors easier access and exposure to oil markets (Cabinet Office-UK, 2008):

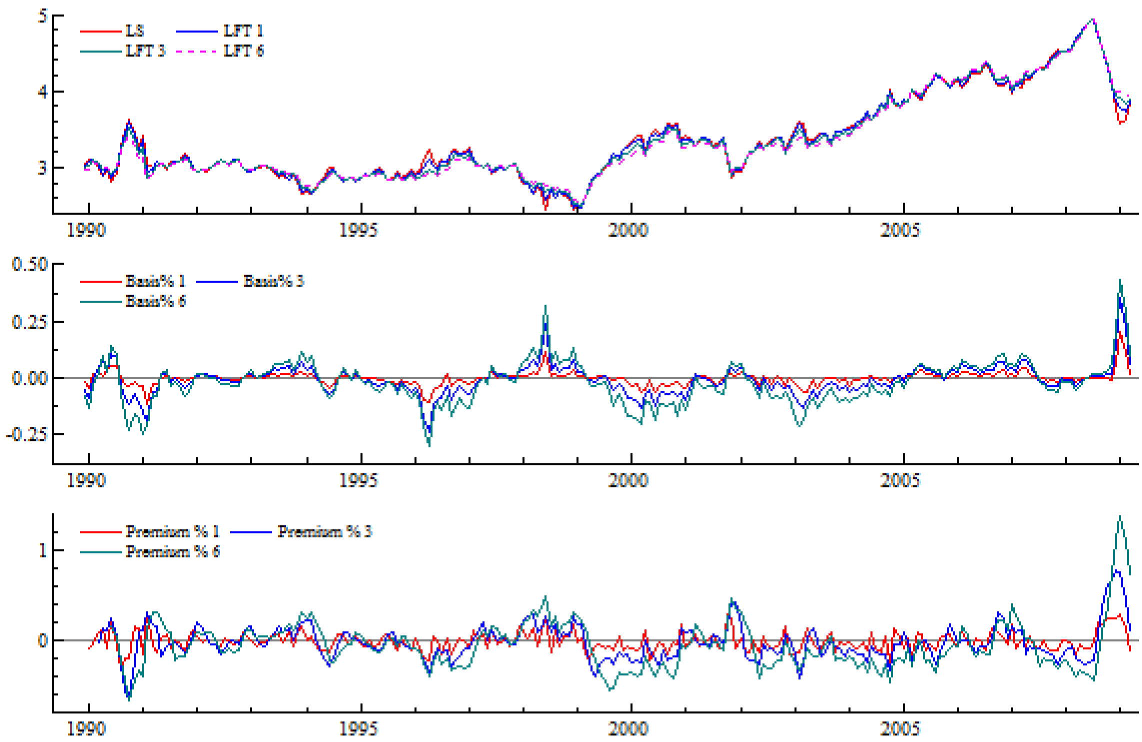
- as an inflationary hedge, against rising global prices and a depreciating US dollar.
- as a search for yield in a deteriorating global economic environment, especially against a backdrop of decreasing returns/increasing risk in other asset markets.
- as a portfolio diversification mechanism to reduce risk, as commodities have historically been counter-correlated with stocks and bonds.

Consequently, futures markets have reported a sharp increase in activity of non-commercial traders such as hedge funds and commercial traders such as commodity swap dealers, who often act as intermediaries for financial investors looking for greater exposure to commodity prices. According to the Cabinet Office-UK (2008), they changed the nature of traders in oil futures markets, away from traditional investors hedging against future price changes to financial investors with limited specialist knowledge of oil markets looking for sustained exposure to commodities.

3.3. WTI Spot and Futures Price Development

Evidence of the state of the WTI crude oil futures market in the period 1990 -2009 is provided in exhibit 7 this is a log representation[12] of the variables. This exhibit displays the WTI spot and futures prices in logarithm form (the top graph), the spread between the 1-, 3- and 6-month futures prices and WTI spot, or the basis (the middle graph), and the premium or, also called the forecasting error (the bottom graph).

Exhibit 7 Logarithm Representation of Spot and Futures Prices, Basis and Premium



NOTE: Sample period: Jan. 1990 – Feb 2009.

From exhibit 7 it is evident that the basis and the premium are volatile over time. As well, we can read when the oil market was in contango and when it was in backwardation – when prices fall the market finds itself in contango modus, and when prices increase the market shifts to backwardation modus. The market was in contango, for example, just before the Gulf War in 1990, and moved strongly into backwardation during the last stage of the war. The oil market was also in backwardation during the temporary oil price increase in 1996. With the beginning of the Asian crisis, august 1998, oil prices fell considerably and the oil market was

[12] The level representation is graphed in the exhibit 28 included in the appendix

in contango until the beginning of 1999, then market moved strongly into backwardation when spot oil prices were increasing through the period 1999 – 2005 (in the dotcom bobble year 2002 market shifted shortly into contango modus). During the temporary, but extreme oil price increase from mid 2007 to mid 2008, the oil market found itself in backwardation. And, it moved from backwardation to contango in August 2008 as oil prices plunged through the last four month of 2008.

Concerning the forecasting error, this was greatest the more uncertainty and volatility was dominating the market.

3.4. Literature Review

The empirical literature on efficiency and price discovery in energy markets is extensive and this section mainly focuses on current works. One of them is the study of Asche and Guttormsen (2002). They investigated the relationship between spot and futures gas oil prices. Their data set is in monthly format and extends over April 1981 to September 2001. Using co-integration and error correction models, they found that spot prices and futures prices with different time to maturity follow the same stochastic trend and move proportionally over time. Furthermore, they found that futures prices lead spot prices, and futures contracts with longer time to expiration lead contracts with shorter time to expiration.

Concerning the research on the impact of speculators' activity in the oil futures market, one of the most recent researches is the study of Sanders, Boris and Manfredo (2004). Their empirical analysis is focused on CFTC (weekly) traders' positions in crude oil, gasoline, heating oil, and natural gas futures from 1992 through 1999. After showing that the largest positions are held by reporting commercials and the smallest by reporting non-commercials, they found a positive correlation between returns and noncommercial positions, and a negative correlation between returns and commercial positions. Further, using Granger causality they found that traders' net positions do not lead market returns in general; and using the Cumby and Modest market timing framework they found little evidence concerning the impact of extreme traders' positions on market returns.

These findings are in part consistent with those of Gurrib (2006) who studied the impact of speculators' activity on U.S. Crude Oil futures prices. He used the weekly net long noncommercial positions from CFTC as a proxy for speculation, the data set ranging from June 1995 to June 2006. Applying a vector error-correction model, he found significant short-run causality running from speculators' trading activity to futures prices. However, the magnitude of this effect was small. While long-run causality was found to run from prices to speculators' activity and not vice versa, whenever spot prices, futures prices and speculators' net positions are co-integrated. He went a step further and, using the conditional standard deviation as a proxy to volatility, he found no significant relationship between large speculators trading activity and volatility in the futures market.

Another interesting empirical research on the role of traders in the energy futures markets is that of Haigh, Hranaiova and Overdahl (2005). Employing a well-disaggregated CFTC dataset consisting of trader positions in U.S. energy futures markets, they analyzed trading relationships between managed money traders (a category of speculative traders) and other groups of traders (e.g., floor brokers, swap dealers, producers, manufacturers). They found that on average managed money traders did not change their positions as frequently as other groups. Using causal techniques they determined that, on average, changes in managed money traders positions were triggered by position changes of other trader groups. Furthermore, they found that managed money traders are an important source of liquidity to the other participants and they rejected the hypothesis that managed money traders trading causes price volatility in U.S. futures markets.

A research that found opposite evidence on the impact of speculators is the research of Kaufmann and Ullman (2008). They attempted to determine where changes in the price of crude oil originate and how they spread by examining the causal relationships among prices for crude oils from North America, Europe, Africa, and the Middle East on both the spot market and in futures markets for both near month and far month contracts. Their data set is in weekly format and has different starting dates for different types of oil ending in March 2008. They analyzed the causal relationships between the prices of crude oils by using two techniques: a two step dynamic ordinary least squares error correction model procedure developed by Stock and Watson (1993) and a full information maximum likelihood estimate for a vector error correction model developed by Johansen and Juselius (1990). They found relatively weak links between spot and futures markets and they suggested that this might have allowed the long-run relationship between spot and future prices to change after September 2004. They concluded that fundamentals initiated a long-term increase in oil prices that was exacerbated by speculators.

A different study involving a set of oil price driving factors is the study of Stevans and Sessions (2008). They examined the relationship between the U.S. real price of oil and futures prices, the exchange value of the dollar, inventories, demand, and supply. They used weekly data for the period January 1988 – March 2008. All of the variables are treated as jointly endogenous and vector error correction model was used to test for co-integration among the variables. They found that for model specifications with short-term futures contracts, supply does indeed dominate price movements in the crude oil market. However, for specifications including longer-term contracts that are inherently more speculative, the real price of oil appears to be determined predominantly by the futures price. They stressed the implication of their results, namely that if regulators really wanted to limit speculation in the oil market, they should keep the shorter-term futures contracts and eliminate the more speculative six months futures contracts.

A more extensive and elaborate study on the recent surge in oil prices is the work of Krichene (2006). In this paper they analyzed the relationship between monetary policy and oil prices. With this they showed that an oil demand shock, resulting from record low interest rates, led to the substantial oil price increases of 2004 – 2005. They used monthly time series spanning 1970 – 2005 for crude oil prices, interest rates, and the US dollar nominal effective exchange rate. They found that monetary policy, manifested through changes in interest rates and monetary aggregates, has a significant effect on aggregate demand for goods and services as well as on asset prices such as exchange rates, housing prices, and stock prices. They suggested that the sustained pressure on oil prices observed in 2004 – 2005 could be explained by an excessively expansionary monetary policy, with interest rates falling to record levels in the context of an integrated international capital market. As a result of low interest rates and a depreciating US dollar, demand for oil expanded faster than its supply. Given the short-run price inelasticity of both oil demand and supply, equilibrium was attained through large increases in the price of oil. Based on this study, already by that time, they anticipated a runaway of energy prices becoming inflationary and resulting in a recession.

In sum, there is contradictory evidence on the impact of speculators activity on oil prices and on the reasons of the most recent price spike. This warrants a very elaborate analysis. However, in this paper I focus solely on examining the market efficiency and the role of speculators in price formation, though I recognize the wide list of oil price drivers that might be studied within different frameworks, such as the exhaustible resources theory, the supply–demand framework, and the informational approach. In this context, the role of speculators is part of the informational approach, as it is the role of OPEC, the erosion of spare capacity, and the role of inventories.

4 IMPLEMENTATION AND RESULT ANALISIS

The concepts of market efficiency and unbiasedness are difficult to separate empirically.

Market efficiency implies that the futures price will equal the expected future spot price (plus or minus a possible time-varying risk premium), while the futures price will be unbiased predicator of the future spot price only if markets are both efficient and have no risk premium. Thus, the hypothesis that the futures price provides unbiased forecast of the spot price is a joint hypothesis of market efficiency and risk neutrality.

This issue is further complicated by a time dimension – that is, while markets may be efficient and unbiased in the long-run, they may be inefficient and biased in the short-run.

One of the objectives of this section is to empirically test the separate and joint hypotheses of market efficiency and unbiasedness in both the long and the short term.

The conventional process of testing for market efficiency and unbiasedness entails:

1. Co-integration analysis

   This comprises two tests. One is to test whether there are any co-integrating long-run relationships between the spot price and the three pairs of futures prices. And the other is to test separately and jointly the hypotheses of market efficiency and unbiasedness in the long run.

2. Error correction analysis

   This entails the individual and joint test of market efficiency and unbiasedness hypotheses in the short run.

I begin the efficiency test using the traditional regression approach developed by Fama (1984), testing whether the basis contains information about the future spot price and about the risk premium at the expiration of the future contract.

Lastly, in order to extend the task in a more practical direction, I introduce the net commercial and non-commercial positions variable. With this I test the Granger causality relationship between traders' positions and returns and test the impact of the extreme positions on market returns.

4.1. Fama's approach

In section 2 it was mentioned that the most appropriate model to test for efficiency is the model that supports the idea of futures prices being unbiased estimators of future spot prices.

To test whether futures prices have power to forecast future spot prices, I employ Fama's (1984) regression approach. Moreover, since I assume risk-neutral and rational actors, Fama's approach makes it possible to test whether the expected premium in market is nonzero.
More specifically, I implement Fama's (1984) regression approach to test whether the basis in any period contains information about future spot prices or contains information about the risk premium at the expiration of the future contract. Two equations are estimated.

The first is the spot price change regression:

$$S_T - S_t = \propto_1 + \beta_1 \, [F(T)_t - S_t] + e_T^c \qquad (10 *)$$

The second is premium regression:

$$F(T)_t - S_T = \propto_2 + \beta_2 \, [F(T)_t - S_t] + e_T^P \quad (11 *)$$

If β_1 is significantly different from zero then we can deduce that the basis, $[F(T)_t - S_t]$, contains information about the changes in spot price. And, if β_2 is significantly different than zero then the premium, $[F(T)_t - S_T]$ has variations that show up in the basis.

These two regressions, (10*) and (11*), are subject to an adding-up constraint. The sum of the premium, $[F(T)_t - S_T]$, and the change in the spot price, $[S_T - S_t]$, is the basis, $[F(T)_t - S_t]$. Thus, the intercepts in (10*) and (11*) must sum to 0.0; each period's residuals must sum to 0.0; and the slope coefficients must sum to 1.0. Thus, the regressions always assign all variation in the basis to the expected premium, the expected change in the spot price, or some mix of the two (Fama, 1984). However, the allocation can be statistically unreliable. Since estimates of β_1 and β_2 are typically between 0.0 and 1.0, the regressions can fail to identify the source of variation in the basis. The results will depend heavily on the sample period. For instance, in the period 1990(1) – 2008(10) β_1 is significant and β_2 is insignificant for all maturities; while in the period 1992(1) – 2002(12) β_1 is insignificant and β_2 is significant for all maturities. The results are summarized in exhibit 8 that follows and exhibit 32 included in the appendix.

Exhibit 8 Fama's Change Regression - 1990(1) - 2008(10)

Regression: $S_t - S_{t-1} = \alpha + \beta\ [F1_{t-1} - S_{t-1}] + e_t$				Regression: $S_t - S_{t-3} = \alpha + \beta\ [F3_{t-3} - S_{t-3}] + e_t$				Regression: $S_t - S_{t-6} = \alpha + \beta [F6_{t-6} - S_{t-6}] + e_t$			
Diff Log S1	**coeff**	**t-value**	**p-value**	**Diff Log S3**	**coeff**	**t-value**	**p-value**	**Diff Log S6**	**coeff**	**t-value**	**p-value**
Constant	0.008	1.170	0.242	Constant	0.030	2.570	0.011	Constant	0.064	4.170	0.000
Basis% _1	0.705	2.650	0.009	Basis% _3	0.720	3.580	0.001	Basis% _6	0.683	3.960	0.000
R^2	0.0311			R^2	0.055			R^2	0.067		
F(1,219) =	7.032 [0.009]**			F(1,219) =	12.64 [0.000]**			F(1,219) =	15.68 [0.000]**		

NOTE: The sample spans over 1990(1) - 2008(10); thus, 226 observations.
Since the first six observations were used to compute the lagged variables,
the regression sample is running from 1990(6) – 2008(10); thus, the regressions are based on 221 observations.
The numbers in parenthesis represent the probabilities of falsely rejecting the null hypothesis (i.e., p-value)
**and *denote 1% and 5% level of significance, respectively.

The dependent variable is the change in the spot price and independent variable is the basis (prices are expressed as natural logs):
$\Delta S1_t = S_t - S_{t-1}$, $\Delta S3_t = S_t - S_{t-3}$ and $\Delta S6_t = S_t - S_{t-6}$
$Basis\%\_1 = F1_{t-1} - S_{t-1}$, $Basis\%\_3 = F3_{t-3} - S_{t-3}$, and $Basis\%\_6 = F6_{t-6} - S_{t-6}$

The estimated slope in the 1-month change regression is 0.705; taken literally, a 1.0% increase in the basis implies a 0.705% drop in the expected price change.

The change regression indicates reliable forecast power in futures prices for all maturities; that is, the basis $[F(T)_{t-T} - S_{t-T}]$ contains some reliable information regarding the future change in the spot price $[S_t - S_{t-T}]$ for all maturities. The evidence of futures as predicators of future spot prices is supported only for the one-month maturity, since the estimated constant is not significantly different from zero, and the slope coefficient is not significantly different from one. High forecasting power might be due to high storage costs and significant basis seasonals.

Following I examine the expectation hypothesis by restricting individually or jointly the coefficients $\beta_1 = 1$ and $\propto_1 = 0$.

Exhibit 9 Restrictions - Change Regression - 1990(1) - 2008(10)

Ho:	F-test	S1/F1	S3/F3	S6/F6
$\propto_1 = 0$	F(1,219) =	1.38[0.242]	6.60[0.011]*	17.39 [0.000]**
$\beta_1 = 1$	F(1,219) =	1.23 [0.269]	1.92 [0.168]	3.37 [0.068]
$\propto_1 = 0, \beta_1 = 1$	F(2,219) =	1.46 [0.235]	5.16[0.007]**	13.90 [0.000]**

NOTE: The sample spans over 1991(1) - 2008(10); Observations: 214.
The numbers in parenthesis represent the probabilities of falsely rejecting the null hypothesis (i.e., p-value)
**and *denote 1% and 5% level of significance, respectively.

On the basis of exhibit 9 it can be summarized that for one-month maturity both individual and the joint hypotheses are not rejected, that is, there is reliable forecast power in futures prices. While for three- and six-month maturity the $\propto_1 = 0$ hypothesis is rejected, and the

$\beta_1 = 1$ hypothesis moves towards rejection - six-month maturity is significant at 10% significance level.

Regarding the premium regression, the results are summarized in exhibit 10.

Exhibit 10 Fama's Premium Regression - 1990(1) - 2008(10)

Regression: $[F1_{t-1} - S_t] = \alpha + \beta[F1_{t-1} - S_{t-1}] + e_t$				Regression: $[F3_{t-3} - S_t] = \alpha + \beta[F3_{t-3} - S_{t-3}] + e_t$				Regression: $(F6_{t-6} - S_t) = \alpha + \beta[F6_{t-6} - S_{t-6}] + e_t$			
Premium%1	**coeff**	**t-value**	**p-value**	**Premium%3**	**coeff**	**t-value**	**p-value**	**Premium%3**	**coeff**	**t-value**	**p-value**
Constant	-0.008	-1.170	0.242	Constant	-0.03	-2.570	0.011	Constant	-0.064	-4.170	0.000
Basis% _1	0.295	1.110	0.269	Basis% _3	0.280	1.380	0.168	Basis% _6	0.317	1.840	0.068
R^2	0.005			R^2	0.008			R^2	0.015		
F(1,219) =	1.23 [0.269]			F(1,219) =	1.92 [0.168]			F(1,219) =	3.37 [0.068]		

NOTE: The sample spans over 1991(1) - 2008(10); Observations: 214.
The numbers in parenthesis represent the probabilities of falsely rejecting the null hypothesis (i.e., p-value)
**and *denote 1% and 5% level of significance, respectively.

The dependent variable is the change in the spot price and independent variable is the basis (prices are expressed as natural logs):
$Premium\%\_1 = F1_{t-1} - S_t$ $Premium\%\_3 = F3_{t-3} - S_t$ and $Premium\%\_6 = F6_{t-6} - S_t$
$Basis\%\_1 = F1_{t-1} - S_{t-1}$ $Basis\%\_3 = F3_{t-3} - S_{t-3}$ and $Basis\%\_6 = F6_{t-6} - S_{t-6}$

The estimated slope in the 1-month change regression is 0.295; taken literally, a 1.0% increase in the basis implies a 0.295% drop in the expected premium.
The premium regression it indicates no reliable evidence of time-varying expected premiums; that is, the basis $[F(T)_{t-T} - S_{t-T}]$ contains some reliable information regarding the premium $[F(T)_t - S_T]$ for all maturities. The evidence of futures as predictors of future spot prices is supported, since the estimated constant is not significantly different from zero, and the slope coefficient is not significantly different from one.

Now I examine the expectation hypothesis by restricting hypothesis the coefficients $\beta_2 = 1$ and $\propto_2 = 0$.

Exhibit 11 Restrictions - Premium Regression - 1990(1) - 2008(10)

Ho:	F-test	S1/F1	S3/F3	S6/F6
$\propto_2 = 0$	F(1,219) =	1.38 [0.242]	6.6 [0.011]*	17.39 [0.000]**
$\beta_2 = 1$	F(1,219) =	7.03 [0.009]**	12.64 [0.0005]**	15.68 [0.000]**
$\propto_2 = 0, \beta_2 = 1$	F(2,219) =	3.92 [0.021]*	8.13 [0.0005]**	12.74[0.000]**

NOTE: The sample spans over 1991(1) - 2008(10); Observations: 214.
The numbers in parenthesis represent the probabilities of falsely rejecting the null hypothesis (i.e., p-value)
**and *denote 1% and 5% level of significance, respectively.

From the exhibit 11 it can be summarized that for one-month maturity the individual restriction $\propto_2 = 0$ is not rejected, while is rejected for the individual $\beta_2 = 1$ and the joint hypothesis; that is, there are no time-varying expected premiums.

As mentioned, the change and premium regressions assign all basis variation to expected premiums, expected spot-price changes, or some combination of the two, but this can be statistically unreliable. If the sample is changed to, for instance 1992(1) – 2002(12), the opposite conclusion results. The results are presented in exhibits 30 and 32 included in the appendix. While in the period 1990(1) – 2008(10) β_1 is significant and β_2 is insignificant for all maturities, implying that there is evidence for good forecasting power of futures prices and no reliable evidence of time-varying expected premiums; in the period 1992(1) – 2002(12) β_1 is insignificant and β_2 is significant for all maturities, implying no reliable evidence for the forecasting power of futures prices and there is reliable evidence of time-varying expected premiums.

When examining the expectation hypothesis for equation (10*), it can be said that the $\propto_1 = 0$ hypothesis is not rejected for all maturities; the individual $\beta_1 = 1$ and the joint $\propto_1 = 0, \beta_1 = 1$ hypotheses are rejected for all maturities (whether at 10%, 5% or 1% significance level). That suggests no evidence on the forecastability power of the futures prices regarding the future spot prices. For equation (11*) all the hypotheses, individual or joint, are rejected for all maturities. That means that there is strong evidence on the existence of time-varying expected premiums. The results are presented exhibits 31 and 33 in the appendix.

The fact that it is unlikely that the regressions (10*) and (11*) can reliably assign basis variation to expected premiums or expected spot-changes can be confirmed by looking at the variations in the variables. The results are summarized in the exhibit 12. If the basis variation (computed in terms of standard deviation) is low relative to the variation of premiums and spot-price changes, then these regressions produce unreliable results.

Exhibit 12 Standard Deviation for the Spot-Change, Basis and Premium

Sample period: 1990(6) - 2008(10) Observations: 221				Sample period: 1992(6) - 2002(12) Observations: 127			
	Change S1%	**Change S3%**	**Change S6%**		**Change S1%**	**Change S3%**	**Change S6%**
Mean	7.26 %	8.59 %	9.19 %	Mean	3.03 %	3.17 %	3.42 %
Std.Dev.	35.04 %	34.32 %	31.77 %	Std.Dev.	32.28 %	30.60 %	30.09 %
	Basis % 1	**Basis% 3**	**Basis% 6**		**Basis % 1**	**Basis% 3**	**Basis% 6**
Mean	-3.38 %	-4.71 %	-5.14 %	Mean	-5.45 %	-5.64 %	-5.44 %
Std.Dev.	8.72 %	11.04 %	11.93 %	Std.Dev.	8.99 %	11.36 %	12.21 %
	Premia % 1	**Premia % 3**	**Premia % 6**		**Premia % 1**	**Premia % 3**	**Premia % 6**
Mean	-10.47 %	-13.07 %	-14.50 %	Mean	-8.38 %	-8.34 %	-8.09 %
Std.Dev.	34.59 %	33.52 %	30.93 %	Std.Dev.	32.71 %	31.47 %	31.14 %

NOTE: All prices are measured in natural logs, that is the return, basis and premiums are in geometric format.
The the return, basis and premium are annualized- form monthly to annually:

$$LnR(ann.) = LnR(mnth) * T$$

$$St.Dev.R\ (ann.) = St.Dev.R\ (mnth) * \sqrt{T}$$

For one-month maturity $T = 12$; for three-month maturity $T = 4$; and , for six-month maturity $T = 2$.

The standard deviation is higher in first period, which is an expected result since this period includes the Kuwait invasion by Iraqis (August1990), the Gulf War (March 2003), the financial crisis (August 2008).

Basis standard deviation is indeed very small relative to the variances of changes in spot price and premiums, thus the Fama's regressions approach is not reliable and further empirical analysis is required.

The analysis can be extended by employing the co-integration and error correction models if prices prove to be non-stationary and if any linear combination of them is stationary. This is the subject for the next section.

4.2. Unit Root test

If the geometric Brownian Motion is assumed to be a reasonable proxy for the behavior of crude oil prices, then market efficiency per se requires that price changes are uncorrelated, which implies a unit root in the level or logarithm of the price series. That would suggest that unit root test can be thought as *prima face* evidence for market efficiency.

In this context unit root test is a test for random walk. If crude oil spot and futures prices follow a random walk, the result is that the crude oil market is efficient in the weak form, meaning future prices cannot be predicted using historical price data. This implies that an uninformed investor with a diversified portfolio will, on average, obtain a rate of return as good as an expert. If the random walk hypothesis is rejected it follows that it is possible for investors to make profits using technical analysis.

In this paper the unit root test was implemented using the ADF test. The tests were run using one lag, a constant (and a time trend). One reason for including a constant (and a time trend) is to ensure that the test will have the correct rejection frequency under the null hypothesis. Another reason is justified by economic theory and the nature of the commodity.

That oil is a renewable resource, the supply of which is limited relative to demand, is a justification for inclusion of a time trend. After 2000 very few giant discoveries of oil have been made. Even though great discoveries have been made, they are mostly in deep-water. This requires sophisticated technology in terms of exploration and extraction as well as consideration of constant risk of environmental hazards, which have to be included in the spot price. Therefore, based on the recent pattern in oil discoveries and production one would expect to see an upward trend in prices, which would reflect supply constraints. As well, an upward trend in oil prices is consistent with Hotelling's theory, which states that prices for exhaustible resources should increase at an exponential rate over time.

The results of the ADF test are reported in the exhibit 13.

Exhibit 13 Unit Root Test - ADF test

Regression:

$$\Delta y_t = \propto (+\lambda t) + \beta y_{t-1} + \sum_{i=1}^{lag} c_i \, \Delta y_{t-i} + v_t$$

	weekly		monthly	
	without trend	with trend	without trend	with trend
log	t-value	t-value	t-value	t-value
Log S	-0.86	-2.66	-0.68	-2.56
Log F1	-0.72	-2.51	-0.57	-2.50
Log F3	-0.53	-2.34	-0.48	-2.44
Log F6	-0.30	-2.16	-0.36	-2.29
Change S %	-23.04**	-23.04**	-9.647**	-9.635**
Basis %	-7.487**	-7.591**	-4.647**	-4.706**
Premium %	-15.97**	-16.01**	-8.68**	-8.636**

NOTE: **Sample Period:** 1993 (01)- 2008 (10)

Critical values used in ADF test:

weekly	without trend	5%=-2.865, 1%=-3.441
	with trend	5%=-3.418, 1%=-3.974
monthly	without trend	5%=-2.877, 1%=-3.466
	with trend	5%=-3.434, 1%=-4.009

Performing the test on the natural logs[13] of each series we see that the null hypothesis of unit root is not rejected, while the tests on the first differences result in the rejection of the null hypothesis. So, I conclude that the price series are integrated of order one, I(1). The basis and the premium are integrated (or stationary) as well.

As mentioned in section 3 one limitation of the ADF is that it assumes a linear deterministic trend to account for the upward trend in the economic variables. The fact that price series often display changes in the trend when major economic events such as oil crises or financial crunches occur, justifies a unit root test that allows under both null and alternative hypotheses for the presence of a one-time change in the level and/or in the slope of the trend function.
The test consists in adjusting the ADF test by including dummy variables to ensure there are as many deterministic regressors as there are deterministic components.

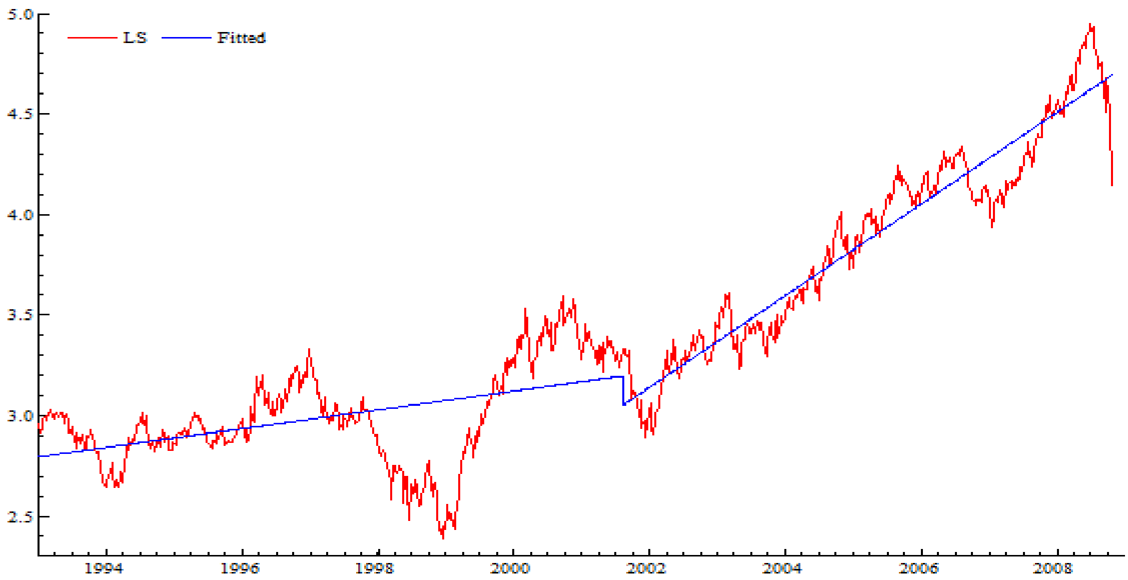
I run a regression that allows for change in the intercept and change in the slope taking place simultaneously (model C). The break point is chosen to be August 2001. This date corresponds with oil prices fall due to weak world demand largely as a result of economic recession in the United States) and OPEC overproduction. Oil prices continued to decline

[13] Test the logarithm of a price because these are assumed to follow geometric processes in continuous time (Alexander, 2008).

sharply following the September 11, 2001 terrorist attacks on the United States, largely on increased fears of a sharper worldwide economic downturn (and therefore sharply lower oil demand). Prior to August 2001, oil prices tripled between January 1999 and September 2000 due to strong world oil demand, OPEC oil production cutbacks, and other factors, including weather and low oil stock levels. With the beginning of 2002 prices embarked on an increasing trend with the rise in prices accelerating from 2004. This date coincides with oil production cuts by OPEC and non-OPEC countries, plus unrest in the Middle East and the possibility of renewed conflict with Iraq.

The increasing trend in oil prices since 2002 is also explained by the introduction of the Commodity Futures Modernization (Deregulation) Act in 2000. As deregulation took hold and ICE grew in popularity, the price of oil began a steady and then rapid rise. In the 12 years from 1988 to 2000, the price of a barrel of oil doubled from \$18 to an average of \$36 per barrel. In just the years from 2000 to 2005, the price doubled again, rising to \$60 per barrel. But the prices in 2007 and 2008 would exceed them all. In just 14 months, from January 2007 to March 2008, the price doubled again, increasing from \$55 to \$110 per barrel. Such a rapid rise in price has only happened twice before in modern history: during the 1973 and 1979 energy crises.

Exhibit 14 The Trend Function with a One -Time Change in the Intercept and Slope



NOTE: The broken straight line is a fitted trend (by OLS) of the form $\Delta y_t = \mu^C + \beta^C t + \theta^C DU_t + \gamma^C DT_t + e_t$, where $DU_t = DT_t = 0 \; if \; t \leq 2001 \; (aug)$ and $DU_t = 1, \; DT_t = t \; if \; t > 2001(aug)$.
Sample spans over 1993 (1)- 2008 (10).

Exhibit 14 shows a plot of the logarithm of the spot price. A feature of this plot is the marked change in the level of the price series in 2001 followed by a higher growth path after August 2001 (denoted to be the break point). The solid line is the estimated trend with two dummy variables added, an intercept dummy (0 prior and at august 2001, 1 after august 2001) and a slope dummy (0 prior and at august 2001, t after august 2001).
The results of the unit root test with a one-time structural break are reported in exhibit 15.

Exhibit 15 The Unit Root Test with a One-Time Structural Break – Model C

Model (C)

$$\Delta y_t = \mu^C + \theta^C DU_t + \beta^C t + \gamma^C DT_t + d^C D(T_B)_t + \alpha^C y_{t-1} + \sum_{i=1}^{k} c_i \, \Delta y_{t-i} + e_t \; (3.c)$$

y=	S		F1		F3		F6	
	coeff	**t-value**	**coeff**	**t-value**	**coeff**	**t-value**	**coeff**	**t-value**
Constant (μ^C)	0.407	1.570	0.367	1.480	0.339	1.450	0.326	1.470
Trend (t)	0.001	0.965	0.001	0.917	0.001	0.880	0.001	0.806
y_{t-1}	-0.028	**-2.980**	-0.025	**2.790**	-0.024	**-2.960**	-0.022	**-2.660**
Δy_{t-1}	0.014	0.387	0.035	0.970	0.066	1.830	0.095	2.660
$D(T_B)_t$(aug01)	0.141	0.064	-0.512	-0.247	-0.625	-0.322	-0.455	-0.251
DU_t (aug01)	-2.521	-1.980	-2.230	-1.810	-2.102	-1.760	-2.076	-1.790
DT_t (aug01)	0.005	2.090	0.002	1.920	0.004	1.880	0.004	1.900
NOTE:	**Sample Period:**		1993 (01) – 2008 (10)					
	Critical Values:		5%=-4.24, 1%=-4.88					

From the exhibit 15 it can be seen that for each of the series the null hypothesis of a unit root cannot be rejected at 5% significance level.

A similar unit root test with one-time structural break was performed in the context of model A with a change in the intercept only. The results are presented in the exhibit 34 included in the appendix. Same conclusion yields for model A as for the model C.

Concluding this section I find that each of the oil price series can be characterized as a random walk process (containing unit roots) and that the structural break is significant and meaningful in terms of events that have impacted on world oil markets.

However, while the test results offer some support for the hypothesis that the price series have unit roots, this evidence by no means is conclusive on evidence for market efficiency and a further test is required to be carried out. Since price series are integrated I(1), a co-integration test is appropriate.

4.3. Co-Integration Test

Once the price series are confirmed to be integrated of identical order I(1), Johansen Maximum Likelihood co-integration test is employed to ascertain the existence of a co-integrating relationship between them.

To carry out the co-integration test the $VAR(q)$ model in logs was formulated and estimated as an unrestricted reduced form (Doornik, Hendry, 2007):

$$X_t = \Phi D_t + BX_{t-1} + \varepsilon_t \qquad (20)$$

where $X_t = (S_{t+n}, F(T)_t)$, the vector of spot and futures prices each being non-stationary.

When using this methodology it is necessary to establish the appropriate order of the VAR to whiten the error term (Doornik, Hendry, 2007). For the choice of the lag order q, the Akaike information criterion (AIC) is applied. The idea behind the AIC criterion is to select the model which has a minimal loss of information (that is, the smallest AIC). The process is initiated with a generous inclusion of lags, proceeding with gradually excluding of the insignificant ones. Lastly, one compares the AIC criterion between the models that have no serially correlated residuals. Most researchers would begin with a lag length of approximate $T^{1/3}$, where T is the number of observation (Enders, 2004). The number of lags can be extended if a substantial amount of seasonality is suspected.

Provided that the sample for this paper contains 190 observations, the test is initiated with a lag length of 6. Excluding the insignificant lag, the model is reduced to 1 lag for all 6 combinations of equations. Then, the AIC criterion is used for comparing the models containing 6 lags and 1 lag. The results of the AIC are reported in exhibit 16. The lag length resulted to be 1 for each regression.

Furthermore, diagnostic checks - autocorrelation, heteroscedasticity, normality and model specification - on the residuals of equation (3) are performed as a prerequisite to co-integration tests. The autocorrelation test reveals that the null hypothesis of no autocorrelation cannot be rejected for all equations at a 1% significance level. The Ramsey RESET test indicates that the equations are properly specified. The Jarque-Bera test confirms that the residuals are normally distributed. The White test detects no heteroscedasticity in the residuals. Exhibit 16 reports the diagnostic statistic, including the AIC criterion.

Exhibit 16 The Diagnostic Statistics for the pair VAR equations

	Lag Length = 6			Lag Length = 1		
S/F1		**S**	**F1**		**S**	**F1**
AR 1-7 test:	F(7,169)	1.244 [0.282]	1.255 [0.276]	F(7,179)	0.41506 [0.8921]	0.39473 [0.9046]
Normality test:	Chi^2(2)	11.480 [0.003]**	14.747 [0.0006]**	Chi^2(2	14.978 [0.0006]**	20.113 [0.000]**
Vector AR 1-7 test:	F(28,322)	0.99049 [0.4824]		F(28,342)	1.1156 [0.3165]	
Vector Normality test:	Chi^2(4)	63.278 [0.0000]**		Chi^2(4)	61.043 [0.0000]**	
AIC	***-7.1821***			***-7.2328<***		
from a system with 6 lags --> system with 1 lag: Chi^2(20) = 30.365 [0.0642]						
S/F3		**S**	**F3**		**S**	**F3**
AR 1-7 test:	F(7,169)	1.3502 [0.2297]	1.3972 [0.2095]	F(7,179)	0.43661 [0.8782]	0.60414 [0.7520]
Normality test:	Chi^2(2)	11.601 [0.0030]**	12.671 [0.0018]**	Chi^2(2	14.825 [0.0006]**	17.053 [0.0002]**
Vector AR 1-7 test:	F(28,322)	1.1133 [0.3199]		F(28,342)	0.89855 [0.6176]	
Vector Normality test:	Chi^2(4)	70.069 [0.0000]**		Chi^2(4)	69.956 [0.0000]**	
AIC	***-6.3432***			***-6.4250<***		
from a system with 6 lags --> system with 1 lag: Chi^2(20) = 24.450 [0.2233]						
S/F6		**S**	**F6**		**S**	**F6**
AR 1-7 test:	F(7,169)	1.4328 [0.1951]	1.4839 [0.1760]	F(7,179)	0.40803 [0.8965]	1.0664 [0.3870]
Normality test:	Chi^2(2)	11.848 [0.0027]**	10.174 [0.0062]**	Chi^2(2	14.870 [0.0006]**	14.860 [0.0006]**
Vector AR 1-7 test:	F(28,322)	1.1557 [0.2721]		F(28,342)	0.93864 [0.5582]	
Vector Normality test:	Chi^2(4)	47.195 [0.0000]**		Chi^2(4)	52.549 [0.0000]**	
AIC	***-6.0504***			***-6.1457<***		
from a system with 6 lags --> system with 1 lag: Chi^2(20) = 21.878 [0.3472]						
F1/F3		**F1**	**F3**		**F1**	**F3**
AR 1-7 test:	F(7,169)	1.5100 [0.1669]	1.5290 [0.1605]	F(7,179)	0.48217 [0.8467]	0.63091 [0.7299]
Normality test:	Chi^2(2)	14.071 [0.0009]**	12.075 [0.0024]**	Chi^2(2	20.058 [0.0000]**	17.017 [0.0002]**
Vector AR 1-7 test:	F(28,322)	0.84067 [0.7013]		F(28,342)	1.0696 [0.3735]	
Vector Normality test:	Chi^2(4)	59.214 [0.0000]**		Chi^2(4)	48.476 [0.0000]**	
AIC	***-7.7607***			***-7.8356<***		
from a system with 6 lags --> system with 1 lag: Chi^2(20) = 25.770 [0.1735]						
F1/F6		**F1**	**F6**		**F1**	**F6**
AR 1-7 test:	F(7,169)	1.6305 [0.1299]	1.3714 [0.2204]	F(7,179)	0.49204 [0.8396]	1.0734 [0.3823]
Normality test:	Chi^2(2)	13.267 [0.0013]**	9.0716 [0.0107]*	Chi^2(2	20.152 [0.0000]**	14.833 [0.0006]**
Vector AR 1-7 test:	F(28,322)	0.68785 [0.8836]		F(28,342)	1.2159 [0.2120]	
Vector Normality test:	Chi^2(4)	31.653 [0.0000]**		Chi^2(4)	36.010 [0.0000]**	
AIC	-6.8452			-6.9193<		
from a system with 6 lags --> system with 1 lag: Chi^2(20) = 25.930 [0.1681]						
F3/F6		**F3**	**F6**		**F3**	**F6**
AR 1-7 test:	F(7,169)	1.6358 [0.1284]	2.0929 [0.0468]*	F(7,179)	0.72050 [0.6547]	1.1126 [0.3571]
Normality test:	Chi^2(2)	9.7074 [0.0078]**	8.3908 [0.0151]*	Chi^2(2	16.966 [0.0002]**	14.806 [0.0006]**
Vector AR 1-7 test:	F(28,322)	1.0212 [0.4390]		F(28,342)	1.2564 [0.1776]	
Vector Normality test:	Chi^2(4)	33.294 [0.0000]**		Chi^2(4)	39.471 [0.0000]**	
AIC	***-8.3445***			***-8.4195<***		
from a system with 6 lags --> system with 1 lag: Chi^2(20) = 25.742 [0.1745]						

NOTE: The sample spans over (1993(1)- 2008(10)

AR test is a test for serial correlation in the residuals, under the null hypothesis of serial independence. Normality test is a Jarque-Bera Statistic, being distributed as χ^2 with 2 degrees of freedom for individual variable and 4 degrees of freedom for the vector, under the null hypothesis of normal distribution.

The diagnostic check suggests that Johansen's results should be reliable.
The fact that residuals are serially uncorrelated is a condition needed to infer about market efficiency and unbiasedness hypothesis in the next section.

The normality test provides evidence of significant non-normality problem. This appears to be mainly due to extreme events like the Iraqi war in 2003 or the surge in the price in 2008. The problem of non-normality in the data could be overcome by introducing dummy variables relating to these extreme observations. This is not done in this section since it delivers same results concerning the co-integration test.

Another issue to be considered is the inclusion of deterministic terms. Firstly, an unrestricted $constant$ was included to allow for non-zero drift in any unit-root processes found by the co-integration analysis. Secondly, since the trend term was found statistically significant, a restricted $trend$ variable was included. These will ensure that the co-integrating-rank tests are not dependent on 'nuisance parameters' related to the deterministic terms (Doornik, Hendry, 2007). The constant is included as an unrestricted variable and the trend as restricted variable under the hypothesis that the variables X and $B'X$ are linear (a quadratic trend seems unlikely in economics).

When determining the co-integrating rank, two likelihood ratio tests, λ_{trace} and λ_{max}, are employed to identify the co-integration between the four series. Both tests have as null hypothesis that there are at most r co-integration vectors. The λ_{trace} alternative hypothesis is that there exist more than r co-integrating vectors. The λ_{max} alternative hypothesis is that there are exactly $r + 1$ co-integration vectors.

The test can be carried out in a multivariate and bivariate format and the results are summarized in the exhibit 17 and 18.

Exhibit 17 Multivariate Co-Integration Test between Spot and Futures Prices

$$X_t = \Phi D_t + BX_{t-1} + \varepsilon_t$$

	Lag	Ho:	Ha:	Trace Statistic	Ho:	Ha:	Max Statistic
S/F(T)	1	r=0	r>0	220.26 [0.000]**	r=0	r=1	149.14 [0.000]**
		r≤1	r>1	71.13 [0.000]**	r=1	r=2	38.45 [0.000]**
		r≤2	r>2	32.68 [0.005]**	r=2	r=3	26.84 [0.002]**
		r≤3	r>3	5.84 [0.491]	r=3	r=4	5.84 [0.492]
Vector AR 1-7 test:	***F(112,606)= 1.2536 [0.0520]***						

NOTE: The sample spans over (1993(1)- 2008(10)
The numbers in parenthesis represent the probabilities of falsely rejecting the null hypothesis (i.e., p-value)
**and *denote 1% and 5% level of significance, respectively.

The maximum eigenvalue test, as well as the trace test, suggests that there are three co-integration vectors in the system, that is, three stationary linear combinations of $S, F1, F3, F6$; and only one stochastic trend that the spot and the three futures prices share. The conclusion is therefore that the spot and the futures price with different time to maturity are co-integrated and hence there is a long-run relationship between the prices.

This result is further confirmed by the bivariate test indicating that all prices are bilaterally co-integrated.

Exhibit 18 Bivariate Co-Integration Test between Spot and Futures Prices

	Lag	Ho:	Ha:	Trace Statistic	Ho:	Ha:	Max Statistic
S/F1	1	r=0	r>0	41.88 [0.000]**	r=0	r=1	34.96 [0.000]**
		r≤1	r>1	6.92 [0.363]	r=1	r=2	6.92 [0.364]
Vector AR 1-7 test:	***F(28,342) = 1.1156 [0.3165]***						
S/F3	1	r=0	r>0	29.95 [0.013]*	r=0	r=1	23.68 [0.009]**
		r≤1	r>1	6.28 [0.437]	r=1	r=2	6.28 [0.437]
Vector AR 1-7 test:	***F(28,348) = 1.272 [0.166]***						
S/F6	1	r=0	r>0	29.31 [0.016]*	r=0	r=1	23.48 [0.010]**
		r≤1	r>1	5.82 [0.493]	r=1	r=2	5.82 [0.494]
Vector AR 1-7 test:	***F(28,342) = 0.94 [0.56]***						
F1/F3	1	r=0	r>0	30.57 [0.011]*	r=0	r=1	24.38 [0.007]**
		r≤1	r>1	6.18 [0.448]	r=1	r=2	6.18 [0.449]
Vector AR 1-7 test:	***F(28,344) = 1.048 [0.402]***						
F1/F6	1	r=0	r>0	32.09 [0.006]**	r=0	r=1	26.23 [0.003]**
		r≤1	r>1	5.85 [0.489]	r=1	r=2	5.85 [0.491]
Vector AR 1-7 test:	***F(28,342) = 1.2159 [0.2120]***						
F3/F6	1	r=0	r>0	33.87 [0.003]**	r=0	r=1	28.05 [0.001]**
		r≤1	r>1	5.82 [0.494]	r=1	r=2	5.82 [0.495]
Vector AR 1-7 test:	***F(28,342) = 1.256 [0.178]***						

NOTE: * and ** denotes 1% and 5% level of significance, respectively.
The numbers in parenthesis represent the probabilities of falsely rejecting the null hypothesis (i.e., p-value)

Since the hypothesis that $r = 0$ is rejected in the bivariate Johansen's test, it is concluded that there is at most one co-integrating vector for every pair of variables, that is, the spot price and the futures price that are I(1), have a linear combinations being I(0).

The existence of co-integration between the crude spot prices and the three months futures prices, using the Johansen tests, confirms the first necessary condition for long-run market efficiency. This result justifies the use of a vector error correction model for gauging on the short-run dynamics later. Additionally, the evidence provided on the non-autocorrelation in residuals is another necessary condition for long-run market efficiency that justifies the parameter restriction test in the section that follows.

While co-integration is necessary condition for market efficiency, is not a sufficient one. In addition to that, the values of $\propto$ and β parameters should be considered by carrying out a restricted co-integration test.

As mentioned the co-integration regression is conventionally specified as:

$$S_{t+n} = \propto + \beta\, F_t + \varepsilon_T$$

This regression will be inappropriate when non-stationary data are used, but since these variables have been proved to be co-integrated in addition to be non-stationary, I(1), the co-integration methodology is the most appropriate for gauging the parameters of the variables and on the long-run equilibrium relationship between them.

The long-run relationship between the future spot price and the futures price consistent with the unbiasedness hypothesis requires that $S_{t+n} - F_t = 0$; that is, deviations between them should have a mean of zero and price series should be serially uncorrelated.

Thus, based on the spot-futures parity, $S_t = \propto + \beta\, F_{t-n} + e_t$, and conditioned on the no-autocorrelation hypothesis, futures price unbiasedness will be confirmed when the joint restriction $\propto = 0$ and $\beta = 1$ holds true [14]. This joint test assumes that agents are risk-neutral and that they rationally use all available information. Violation of either hypothesis causes the rejection of the joint hypothesis. Nevertheless, the rejection of the joint hypothesis will not necessarily mean that agents earn abnormal profits.

[14] This can be seen if rewrite the spot-futures parity, in the following form: $S_t - F_{t-n} = \propto + (\beta - 1)F_{t-n} + \varepsilon_T$
The new equation is similar to the spot-futures parity, if $\beta = 1$.

The test of the restrictions on the parameters in the co-integrating vectors can be carried out using the Johansen's method.

As X_t is a vector of variables $(S_{t+n}, F_t, 1)$ we define B' as a vector of coefficients $(1, -\beta, -\propto)$. Thus, $B'X_t = Z$ is considered as defining the underlying economic relationships and assumes that the agents react to the disequilibrium error, Z, through adjustment coefficient $\propto$ to restore equilibrium, $(Z = 0)$.

The long-run test for unbiasedness is carried out by imposing $B' = (1, -1, 0)$, which normalizes $\mathrm{S_t}$ to unity and gives the joint restrictions that $\propto = 0$ and $\beta = 1$. The joint null hypothesis is that there is no premium and that the market is efficient, thus, the futures price is an unbiased predicator of future spot price.

The JM approach does not put any constraint on the normalization of the co-integration vectors. In this case, the co-integration relations are more meaningful when interpreted in terms of spot price (as dependent variable).

The individual test $\propto = 0$ implies that there are no premiums in the market, while the individual test $\beta = 1$ entails a unitary elasticity between the future spot price and futures price. Based on this test we can observe whether the prices are proportional and predictable from each other, implying market efficiency.

These individual and joint restriction, are assessed through a likelihood ratio test. Asymptotically the test statistic is distributed as a $\chi^2(m)$ with m number of degrees of freedom that equals the number of restrictions imposed on the coefficients.

If the joint null hypothesis for unbiasedness is rejected three separate conclusions can be drawn (Mckenzie and Holt, 2002):

(1) the market may indeed be inefficient;

(2) a constant risk premium may exist, which makes market forecasts biased but possibly efficient; or

(3) a time-varying risk premium may be inherent in the market, thus preventing futures prices in isolation from providing unbiased forecasts of future spot prices.

The last two conclusions are relevant in the case of the crude oil futures market. This is because agents in futures markets may be risk averse and because oil is a storable commodity. This storability feature may require a premium for agents to cover their storage costs. As a result, a risk premium may be required for agents to use the futures contract to hedge their output.

In the following exhibit the results of the restriction imposed on the regression coefficients are presented. Note that the restrictions embedded in the null hypothesis are binding if the calculated value of the test statistic exceeds the critical value χ^2. If not significant the restriction is not binding.

Exhibit 19 Unbiasedness Test

	Lag Length	Parameter estimates for $(1, -\beta, -\propto)$	Ho: $\propto = 0$	Ho: $\beta = 1$	Ho: $\propto = 0$ and $\beta = 1$
S/F1	1	(1, -1.01, -0.090)	0.093 [0.761]	0.229 [0.632]	0.533 [0.766]
S/F3	1	(1, -0.979, -0.007)	0.003 [0.955]	0.197 [0.657]	0.229 [0.892]
S/F6	1	(1, -0.923, 0.022)	0.112 [0.738]	1.177 [0.278]	2.038 [0.361]
F1/F3	1	(1,-0.975, -0.057)	0.080[0.777]	0.868 [0.352]	0.902 [0.637]
F1/F6	1	(1, -0.924, 0.021)	0.060 [0.810]	2.247 [0.134]	3.230 [0.200]
F3/F6	1	(1, -0.952, 0.016)	0.010 [0.920]	3.544 [0.060]	4.510 [0.110]

NOTE: The sample spans over (1993(1)- 2008(10)
The numbers in parenthesis represent the probabilities of falsely rejecting the null hypothesis (i.e., p-value)
**and *denote 1% and 5% level of significance, respectively.

The individual test, distributed as a $\chi^2(1)$ with one degree of freedom, indicates that individual hypotheses of market efficiency, $\beta = 1$, and no risk premium, $\propto = 0$, are not rejected.

The joint test, distributed as a $\chi^2(2)$ with two degrees of freedom, indicates that the null hypothesis $\propto = 0$ and $\beta = 1$ cannot be rejected at the 1% significance level or higher. Therefore, the results indicate that for the 1-, 3, and 6-month contract the futures price is an unbiased predicator of the spot market in the long-term.

The acceptance of the above restrictions imposed to $\propto$ and β (both jointly and individually) and the serial independence of residuals is a second necessary condition for market efficiency. The above two conditions (unbiasedness and serial independence) are met, therefore it can be concluded that markets are efficient and futures prices provide unbiased estimates of future spot prices in the long run.

However, co-integration test does not reveal the short run market efficiencies, whereby past information can improve future market forecasts of future spot prices. The short run efficiency is studied in the following section.

4.4. Error Correction Model

In this section the short run efficiency of the futures market is to be tested, since in the short run it is possible that there will be considerable departures from the long run equilibrium relationship.

The short run efficiency can be tested by using an error correction model (ECM) in the following form:

$$\Delta Y_t = \Phi_y D_t + \theta Z_{t-1} + \varphi \Delta X_{t-1} + \sum_{i=1}^{m} \varphi_i \Delta X_{t-i} + \sum_{i=1}^{n} \gamma_i \Delta Y_{t-i} + \varepsilon_{yt} \quad (23)$$

$$\Delta X_t = \Phi_x D_t + \theta_f Z_{t-1} + \gamma \Delta Y_{t-1} + \sum_{i=1}^{m} \gamma_i \Delta Y_{t-i} + \sum_{i=1}^{n} \varphi_i \Delta X_{t-i} + \varepsilon_{xt} \quad (24)$$

It was mentioned that the JM approach does not put any constraint on the normalization of the co-integration vectors, but in this case the co-integration relations are more meaningful when interpreted in terms of spot price (as dependent variable). This reduces to EG's ECM:

$$S_t - S_{t-n} = \alpha + \theta(S_t - \beta\, F_{t-n})_{-1} + \varphi(F_{t-1} - F_{t-n-1}) + \sum_{i=1}^{m} \varphi_i (F_t - F_{t-n})_{-i}$$
$$+ \sum_{i=1}^{m} \gamma_i (S_t - S_{t-n})_{-i} + \varepsilon_{st} \quad (25)$$

The vector error correction model (VECM) provides a framework for valid inference in the presence of I (1) variable.

The magnitude of the error correction term coefficient, θ , indicates the seed of adjustment of any disequilibrium toward the long run equilibrium state of market efficiency and unbiasedness.

Co-integration in this form implies that $\theta < 0$, because the spot price responds to movements from the long-term equilibrium position in the long-run equilibrium equation.
Unbiasedness implies that $\varphi = 1$, because any new information affecting movements in the future spot rate will be incorporated immediately in the current futures price (because new information, which also affects the futures price, affects the future change in the spot price).

The coefficients on the lagged values, $\varphi_i = 0$ and $\gamma_i = 0$, because all past information has already been incorporated into the current futures price.

Thus, the restrictions imposed for testing market efficiency are $\varphi_i = \gamma_i = 0$, $\theta = -1$, $\varphi = 1$ and $\alpha = 0$ (not allowing for the presence of a risk premium according to the unbiasedness hypothesis). In this context any short run market inefficiencies cannot be due to long run market bias, and the two concepts of unbiasedness and market efficiency may be regarded as synonymous.

If these restrictions hold, then the above equation reduces to $S_t = F_{t-n} + e_t$. And it can be concluded that the crude oil futures market is efficient and futures prices provide unbiased estimates of future spot prices both in the long-run and the short-run.

If these restrictions do not hold true, the markets would be deemed inefficient, that is past futures and spot prices would contain relevant information not completely incorporated into current future prices, which could be used to predict the future spot price.

In order to assess the model adequacy, at each lag the VAR residuals are checked for satisfying the white noise assumption. Tests for serial correlation, normality, heteroskedasticity and autoregressive conditional heteroskedasticity, are performed. The model was estimated with zero to twelve lags, with significant coefficients retained.

The error correction term was computed as the lagged residual from $(S_t - \propto - \beta F_{t-n})$ cointegrating regression.

Exhibit 20 reports the best ECM specifications for nested tests in more general form of equation 23. The diagnostic statistics are presented in exhibit 20 as well.

Exhibit 20 (a) Estimated Error Correction Model

$$S_t - S_{t-n} = \alpha + \theta(S_t - \beta\, F_{t-n})_{-1} + \varphi(F_t - F_{t-n})_{-1} + \sum_{i=1}^{m} \varphi_i (F_t - F_{t-n})_{-i} + \sum_{i=1}^{m} \gamma_i (S_t - S_{t-n})_{-i} + \varepsilon_{st}$$

Long Run OLS

Log S	coeff	t-value	p-value
Constant	0.117	2.030	0.044
Log F 1_1	0.949	37.500	0.000
Trend	0.000	1.840	0.068

Log S	coeff	t-value	p-value
Constant	0.352	3.730	0.000
Log F 3_3	0.835	20.200	0.000
Trend	0.002	4.200	0.000

Log S	coeff	t-value	p-value
Constant	0.567	4.570	0.000
Log F 6_6	0.720	13.500	0.000
Trend	0.003	6.430	0.000

ECM

	coeff	t-value	p-value
Constant	0.003	0.423	0.673
ECT 1 _1	-0.636	-2.740	0.007
Diff Log S 1_7	0.811	2.110	0.037
Diff Log S 1_8	1.033	2.590	0.010
Diff Log S 1_9	1.231	3.180	0.002
Diff Log S 1_10	0.174	2.360	0.020
Diff Log F 1_1	0.599	2.400	0.017
Diff Log F 1_7	-0.892	-2.130	0.034
Diff Log F 1_8	-1.034	-2.410	0.017
Diff Log F 1_9	-1.350	-3.230	0.002
R^2	0.133		
F(9,179) =	3.04 [0.002]**		
DW	2.03		

	coeff	t-value	p-value
Constant	-0.002	-0.302	0.763
ECT 3 _1	-0.365	-3.370	0.001
Diff Log S 3_1	0.899	5.320	0.000
Diff Log S 3_3	-0.565	-3.820	0.000
Diff Log S 3_4	0.648	4.320	0.000
Diff Log S 3_7	0.496	3.110	0.002
Diff Log S 3_8	0.255	3.480	0.001
Diff Log F 3_1	0.473	2.470	0.014
Diff Log F 3_2	0.416	2.200	0.029
Diff Log F 3_6	-0.783	-3.940	0.000
Diff Log F 3_7	-0.556	-2.680	0.008
Diff Log F 3_8	-0.841	-4.210	0.000
Diff Log F 3_9	-0.930	-4.520	0.000
Diff Log F 3_12	0.174	1.880	0.062
R^2	0.704		
F(13,175) =	32.03 [0.000]**		
DW	2.06		

	coeff	t-value	p-value
Constant	0.000	0.0548	0.956
ECT 6_1	-0.153	-2.090	0.038
Diff Log S 6_1	0.800	7.460	0.000
Diff Log S 6_6	-0.169	-2.050	0.042
Diff Log S 6_7	0.215	2.760	0.006
Diff Log S 6_12	-0.119	-2.900	0.045
Diff Log F 6_1	0.451	2.700	0.008
Diff Log F 6_3	0.418	2.560	0.011
Diff Log F 6_4	0.300	1.800	0.073
Diff Log F 6_5	0.542	3.280	0.001
Diff Log F 6_6	-0.854	-4.180	0.000
R^2	0.814		
F(11,177) =	70.64 [0.000]**		
DW	1.99		

Test	Model 1	Model 3	Model 6
AR 1-7 test:	F(7,172) = 0.790[0.597]	F(7,168) = 0.694 [0.677]	F(7,170) = 0.75 [0.630]
ARCH 1-7 test:	F(7,165) = 0.635[0.727]	F(7,161) = 1.161 [0.328]	F(7,163) = 2.249 [0.033]*
Normality test:	Chi^2(2) = 3.523[0.001]**	Chi^2(2) = 10.51[0.005]**	Chi^2(2) = 7.821 [0.020]*
Hetero test:	F(18,160) = 1.157[0.304]	F(26,148) = 0.88[0.636]	F(22,154) = 0.977 [0.497]
RESET test:	F(1,178) = 0.645[0.423]	F(1,174) =0.003[0.954]	F(1,176) = 0.002 [0.964]

NOTE: The sample spans over (1993(1)- 2008(10))

**and *denote 1% and 5% level of significance, respectively.

The numbers in parenthesis represent the probabilities of falsely rejecting the null hypothesis (i.e., p-value)

AR test is a test for serial correlation in the residuals, under the null hypothesis of serial independence.
ARCH test is a Engle test with the null of conditional homoscedasticity
NORMALITY test is a Jarque-Bera Statistic, being distributed as χ^2 with 2 degrees of freedom for individual variable, under the null hypothesis of normal distribution.
Heteroscedasticity is a test based on White test (using residuals squares on the original regressor and all their squares), under the null of unconditional homoscedasticity.
RESET test is a Regression Specification Test with the null hypothesis of correct specification of the model

Exhibit 20 (b) Estimated Error Correction Model

Long Run OLS

Log F 1	coeff	t-value	p-value
Constant	0.245	3.270	0.001
Log F 3_2	0.886	26.800	0.000
Trend	0.001	3.560	0.001

Log F 1_1	coeff	t-value	p-value
Constant	0.457	4.240	0.000
LFT 6_5	0.771	16.600	0.000
Trend	0.003	6.080	0.000

Log F 3_3	coeff	t-value	p-value
Constant	0.309	3.860	0.000
Log F 6_3	0.847	24.100	0.000
Trend	0.002	5.060	0.000

ECM

	coeff	t-value	p-value
Constant	-0.001	-0.079	0.937
ECT 1/3_1	-0.562	-3.430	0.001
Diff Log F1 _3	0.930	8.520	0.000
Diff Log F1 _4	-0.804	-7.140	0.000
Diff Log F1 _5	0.830	6.990	0.000
Diff Log F1 _7	0.700	5.840	0.000
Diff Log F1 _8	1.203	3.510	0.001
Diff LogF1 _10	1.241	3.440	0.001
Diff Log F3_1	1.591	8.620	0.000
Diff Log F3_2	-1.131	-10.90	0.000
Diff Log F3_6	-0.813	-6.150	0.000
Diff Log F3_8	-2.017	-5.120	0.000
Diff Log F3_9	0.620	4.850	0.000
Diff Log F3_10	-1.652	-3.860	0.000
Diff Log F3_11	0.302	3.570	0.001

R^2	0.593
F(14,174) =	18.08 [0.000]**
DW	1.91

AR 1-7 test:	F(7,167)	=	0.594 [0.76]
ARCH 1-7 test:	F(7,160)	=	0.752[0.628]
Normality test:	Chi^2(2)	=	7.675 [0.022]*
Hetero test:	F(28,145)	=	0.678 [0.886]
RESET test:	F(1,173)	=	5.693 [0.01]

	coeff	t-value	p-value
Constant	0.003	0.385	0.700
ECT 1/6_1	-0.233	-2.570	0.011
Diff Log F1_1	0.482	2.270	0.025
Diff Log F1_2	0.483	2.430	0.016
Diff Log F1_5	-0.594	-7.420	0.000
Diff Log F1_6	0.571	6.650	0.000
Diff Log F1_8	0.537	3.270	0.000
Diff Log F6_1	0.861	3.130	0.002
Diff Log F6_2	-0.786	-2.950	0.004
Diff Log F6_8	-0.655	-3.050	0.003
Diff Log F6_10	-0.244	-2.140	0.034
Diff Log F6_11	0.235	2.340	0.020

R^2	0.759043
F(11,177) =	50.69 [0.000]**
DW	1.92

AR 1-7 test:	F(7,170)	=	0.84 [0.56]
ARCH 1-7 test:	F(7,163)	=	1.02[0.42]
Normality test:	Chi^2(2)	=	1.73[0.42]
Hetero test:	F(22,154)	=	1.15[0.29]
RESET test:	F(1,176)	=	0.95 [0.33]

	coeff	t-value	p-value
Constant	0.001	0.145	0.885
ECT 1/6_3	-0.372	-2.960	0.004
Diff Log F3_1	0.729	2.270	0.024
Diff Log F3_4	0.645	7.330	0.000
Diff Log F3_6	0.696	1.720	0.087
Diff Log F3_7	-1.093	-2.250	0.026
Diff Log F3_8	2.022	5.170	0.000
Diff Log F3_9	-0.265	-2.880	0.005
Diff Log 3_10	0.400	5.510	0.000
Diff Log F6_1	0.682	1.830	0.068
Diff Log F6_2	-0.252	-2.570	0.011
Diff Log F6_3	-0.674	-6.400	0.000
Diff Log F6_6	-1.365	-2.920	0.004
Diff Log F6_7	1.829	3.180	0.002
Diff Log F6_8	-2.407	-5.120	0.000

R^2	0.701922
F(14,174) =	29.27 [0.000]**
DW	1.9

AR 1-7 test:	F(7,167)	=	1.6 [0.139]
ARCH 1-7 test:	F(7,160)	=	1.33 [0.241]
Normality test:	Chi^2(2)	=	8.22 [0.02]*
Hetero test:	F(28,145)	=	1.333 [0.14]
RESET test:	F(1,173)	=	1.937 [0.166]

NOTE: The sample spans over (1993(1)- 2008(10)
**and *denote 1% and 5% level of significance, respectively.
The numbers in parenthesis represent the probabilities of falsely rejecting the null hypothesis (i.e., p-value)

AR test is a test for serial correlation in the residuals, under the null hypothesis of serial independence.
ARCH test is a Engle test with the null of conditional homoscedasticity
NORMALITY test is a Jarque-Bera Statistic, being distributed as χ^2 with 2 degrees of freedom for individual variable, under the null hypothesis of normal distribution.
Heteroscedasticity is a test based on White test (using residuals squares on the original regressor and all their squares), under the null of unconditional homoscedasticity.
RESET test is a Regression Specification Test with the null hypothesis of correct specification of the model

The results indicate no serial independence of residuals, no ARCH effect, no heteroscedasticity and the models seem to be well specified, while there is a problem of non-normality. This appears to be mainly due to extreme events that lead to dramatic fall or increases in oil prices.

The coefficients on the error correction terms are significant in all regressions. This is consistent with the result that the futures and spot markets are co-integrated.

Plots of error correction terms suggest that substantial departures from long-run market efficiency took place during the high volatility periods. This may be regarded as supporting evidence that futures markets provide inadequate forecasts of future spot prices during periods of unexpectedly high volatile rises in the general price level.

An F-test was used to test restrictions $\varphi_i = \gamma_i = 0, \theta = -1, \varphi = 1$ and $\alpha = 0$.
On the basis of F-test, one can see whether the crude oil futures market is efficient in the short run, as it is long run it is efficient. If this is not true arbitrage opportunities are possible.

Exhibit 21 (a) Restriction Results for the Estimated Error Correction Model

S/F1	S/F3	S/F6
(1) JOINT RESTRICTION: $\varphi_i = \gamma_i = 0, \theta = -1, \varphi = 1, \alpha = 0$		
F(1,179) = 6.7499 [0.0102]*	F(1,175) = 13.113 [0.0004]**	F(1,177) = 9.7225 [0.0021]**
(2) RESTRICTION ON THE LAGGED VARIABLES: $\varphi_i = \gamma_i = 0$		
F(7,179) = 2.8131 [0.0084]**	F(11,175) = 15.656 [0.0000]**	F(9,177) = 41.718 [0.0000]**
(3) RESTRICTION ON THE CONSTANT: $\alpha = 0$		
F(1,179) = 0.17894 [0.6728]	F(1,175) = 0.091099 [0.7631]	F(1,177) =0.0030039 [0.9564]
NOTE:	The sample spans over (1993(1)- 2008(10) **and *denote 1% and 5% level of significance, respectively. The numbers in parenthesis represent the probabilities of falsely rejecting the null hypothesis (i.e., p-value)	

Exhibit 21 (b) Restriction Results for the Estimated Error Correction Model

F1/F3	F1/F6	F3/F6
(1) JOINT RESTRICTION: $\varphi_i = \gamma_i = 0, \theta = -1, \varphi = 1, \alpha = 0$		
F(1,174) = 40.055 [0.000]**	F(1,177) = 13.431 [0.0003]**	F(1,174) = 6.516 [0.012]*
(2) RESTRICTION ON THE LAGGED VARIABLES: $\varphi_i = \gamma_i = 0$		
F(12,174) = 12.360 [0.000]**	F(9,177) = 15.102 [0.000]**	F(12,174) = 12.270 [0.000]**
(3) RESTRICTION ON THE CONSTANT: $\alpha = 0$		
F(1,174) =0.01[0.94]	F(1,177) = 0.148 [0.700]	F(1,174) = 0.021 [0.885]
NOTE:	The sample spans over (1993(1)- 2008(10) **and *denote 1% and 5% level of significance, respectively. The numbers in parenthesis represent the probabilities of falsely rejecting the null hypothesis (i.e., p-value)	

Exhibit 21 reports the F-tests on parameters restrictions of the final ECM relating to the second and third necessary condition for market efficiency.

The first case test (1) refers to the unbiasedness hypothesis, imposing $\varphi_i = \gamma_i = 0$, $\theta = -1$, $\varphi = 1$ and $\alpha = 0$. The short-term inefficiency is significant, with a reported p-value of 1.02% one month futures, 0.4% for three month futures and 0.2% for six month futures. Thus, the result rejects the hypothesis of unbiasedness.

The test (2) $\varphi_i = \gamma_i = 0$, is rejected in all cases, which means that the past information is not incorporated immediately and completely in the current futures prices. There is evidence that lagged future and spot prices influence current spot price. Thus, it may be that agents are unable to exploit arbitrage opportunities fully while they are learning about changes in market fundamentals.

The restriction $\alpha = 0$ (3) is not rejected in either of the cases, supporting the non-existence of a risk premium conditional on the form that has been assumed for it. Such a result means that a risk premium could exist, but in any case will not be of a linear form. A nonlinear or time varying risk premium is very possible, which is advocated by the existence of ARCH in the initial data.

Thus, further testing revealed the existence of important short run deviations from unbiasedness. This result strongly suggests that there are short run deviations from the long run efficiency conditions. However, this rejection of the joint hypothesis of market efficiency and unbiasedness in futures prices does not allow the identification of the reason for the rejection. Given that the unbiasedness of futures prices is the most commonly accepted model to test efficiency and the risk premium is assumed to be linear and constant over time, here this rejection could be due to a positive time varying risk premium.

Finally, our findings have two important implications for market participants in NYMEX crude oil futures market. First, it suggests that there are opportunities for consistent speculative profits to be made. Second, in relation to the price discovery role of the copper futures market, it appears that the market does not fulfill this function and hence the information incorporated in futures prices is not considered as important in order to forecast future spot prices.

4.5. Granger Causality Test - Price Leadership Test

Futures contracts were originally developed as new financial instruments for price discovery and risk transfer. The essence of the price discovery function depends on whether new information is reflected first in the futures markets or cash markets. Both markets contribute to the discovery of a unique and common unobservable price, which is the efficient price. Consequently, the analysis of price discovery and information flow across cash and futures markets has received much attention from academicians, regulators and practitioners.

Within the same error correction model one can test the lead-lag relationship. The test for price leadership is carried out by using the Granger causality test. The dependent variable is the change in prices - spot price or futures price of different maturities; and the independent variable is the basis, $S_{t-n} - F(n)_{t-n}$. The following regressions are relevant[15]:

$$S_t - S_{t-n} = \alpha_s + \theta_s[S_{t-n} - F(n)_{t-n}] \quad (26)$$

$$F(n)_t - F(n)_{t-n} = \alpha_f + \theta_f[S_{t-n} - F(n)_{t-n}] \quad (27)$$

The coefficient of the basis, θ_s, contains information about exogeneity. When $\theta_s = 0$ the spot price is weakly exogenous for the futures price and therefore leads the futures price, while if $\theta_f = 0$ there will be no long-run causation towards this variable in the system, hence this variable will be exogenous to the system, and therefore the futures price leads the spot price.

Given that there are four variables in the system there can be 6 pairs of equations (or 12 equations totally). The results for the first three pairs are summarized in the exhibit 22.

[15] The lagged terms of the dependent and independent variables can be introduced if they are significant – for the case at hand these lagged terms are not significant, thus not included

Exhibit 22 Exogeniety Test - Causality between Spot and Futures Prices

Pair 1: $S_t - S_{t-1} = \alpha_s + \theta_s[S_{t-1} - F(1)_{t-1}]$

Diff Log S	coeff	t-value	p-value
Constant	0.007	0.997	0.320
Basis%_1	0.561	2.130	0.034
R^2	0.021		
F(1,212) =	4.557 [0.03]*		

Pair 2: $S_t - S_{t-3} = \alpha_s + \theta_s[S_{t-3} - F(3)_{t-3}]$

Diff LogS_3	coeff	t-value	p-value
Constant	0.022	1.970	0.050
Basis%_3	0.520	2.650	0.009
R^2	0.032		
F(1,212) =	6.999 [0.009]**		

Pair 3: $S_t - S_{t-6} = \alpha_s + \theta_s[S_{t-6} - F(6)_{t-6}]$

Diff Log S_6	coeff	t-value	p-value
Constant	0.054	3.540	0.001
Basis%_6	0.551	3.220	0.002
R^2	0.047		
F(1,212) =	10.38 [0.001]**		

Pair 1: $F(1)_t - F(1)_{t-1} = \alpha_f + \theta_f[S_{t-1} - F(n)_{t-1}]$

Diff Log F1	coeff	t-value	p-value
Constant	0.006	0.902	0.368
Basis%_1	0.216	0.857	0.392
R^2	0.003		
F(1,212) =	0.735 [0.392]		

Pair 2: $F(3)_t - F(3)_{t-3} = \alpha_f + \theta_f[S_{t-3} - F(3)_{t-3}]$

Diff Log F3	coeff	t-value	p-value
Constant	0.005	0.871	0.385
Basis%_3	-0.036	-0.356	0.722
R^2	0.001		
F(1,212) =	0.127 [0.722]		

Pair 3: $F(6)_t - F(6)_{t-6} = \alpha_f + \theta_f[S_{t-6} - F(6)_{t-6}]$

Diff Log F6	coeff	t-value	p-value
Constant	0.004	0.858	0.392
Basis%_6	-0.056	-1.000	0.318
R^2	0.005		
F(1,212) =	1.003 [0.318]		

NOTE: Sample period 1991(1) - 2008(10)
Basis%_1=(LogF1-LogS)_1; Basis%_3 = (LogF3-LogS)_3; Basis%_6 = (LogF6-LogS)_6

From exhibit 22 the coefficients of the error term, θ_s in three equations (where change of spot prices is the dependent variable) are significant at conventional levels and thus the null hypothesis of $\theta_s = 0$ is rejected for all 3 equations. That is, there is no evidence for spot price to be weakly exogenous to the system, or for spot price to lead the futures price. On the other hand, the other three equations, containing the futures prices with different maturities, have coefficients that are not significant enough to be rejected. Thus, the null hypothesis of $\theta_f = 0$ is not rejected and it can be concluded that there will be no long-run causation towards the futures price and that futures price is exogenous to the system.

These results imply that there is causality from futures to spot, i.e. the futures market leads the spot market. That also implies that futures prices tend to discover new information more rapidly than spot prices.

In the following, I test lead-lag relationship between futures prices of different maturities. The results are summarized in exhibit 23.

Exhibit 23 Exogeniety Test - Causality between Futures Prices of Different Maturities

Pair 4				Pair 5				Pair 6			
$F(1)_t - F(1)_{t-2} = \alpha_1 + \theta_1[F(3)_{t-2} - F(1)_{t-2}]$				$F(1)_t - F(1)_{t-5} = \alpha_1 + \theta_1[F(6)_{t-5} - F(1)_{t-5}]$				$F(3)_t - F(3)_{t-3} = \alpha_1 + \theta_1[F(6)_{t-3} - F(3)_{t-3}]$			
Diff Log F1	**coeff**	**t-value**	**p-value**	**Diff Log F1**	**coeff**	**t-value**	**p-value**	**Diff Log F3**	**coeff**	**t-value**	**p-value**
Constant	0.015	1.640	0.102	Constant	0.043	3.160	0.002	Constant	0.024	2.270	0.024
Basis%_2	0.493	1.810	0.072	Basis%_5	0.528	2.630	0.009	Basis%_3	0.492	1.680	0.095
R^2	0.015			R^2	0.032			R^2	0.013		
F(1,212) =	3.267 [0.072]			F(1,212) =	6.903 [0.009]**			F(1,212) =	2.806 [0.095]		
$F(3)_t - F(3)_{t-2} = \alpha_2 + \theta_2[F(3)_{t-2} - F(1)_{t-2}]$				$F(6)_t - F(6)_{t-5} = \alpha_2 + \theta_2[F(6)_{t-5} - F(1)_{t-5}]$				$F(6)_t - F(6)_{t-3} = \alpha_2 + \theta_2[F(6)_{t-3} - F(3)_{t-3}]$			
Diff Log F3	**coeff**	**t-value**	**p-value**	**Diff Log F6**	**coeff**	**t-value**	**p-value**	**Diff Log F6**	**coeff**	**t-value**	**p-value**
Constant	0.013	1.580	0.117	Constant	0.036	3.100	0.002	Constant	0.021	2.190	0.030
Basis%_2	0.213	0.854	0.394	Basis%_5	0.077	0.458	0.648	Basis%_3	0.180	0.680	0.497
R^2	0.003			R^2	0.001			R^2	0.002		
F(1,212) =	0.7293 [0.394]			F(1,212) =	0.2093 [0.648]			F(1,212) =	0.4629 [0.497]		

NOTE: Sample period 1991(1) - 2008(10)
Basis%_2=(LogF3-LogF1)_2; Basis%_5= (LogF6-LogF1)_5; Basis%_3=(LogF6-LogF3)_3.

From exhibit 23 the coefficients of the error term, θ_s in three equations (where futures of shorter maturity are the dependent variable) are significant at 10% significant level and thus the null hypothesis of $\theta_1 = 0$ is rejected at 10% significance level for all three equations. That is, there is no evidence for futures prices of shorter maturities to lead the futures prices of longer maturities. On the other hand, the other three equations have coefficients that are not significant. Thus, the null hypothesis of $\theta_f = 0$ is not rejected and it can be concluded that there is no long-run causation towards the futures price of longer maturities and therefore the six month futures price leads the one- and three month futures prices.

As an overall conclusion it can be said that weak exogeneity could not be rejected for the six months futures contract in any of the relationships containing the six month futures prices (in the left-hand side of the equation). Moreover, weak exogeneity could not be rejected for the three month contract in the formulation containing three month and spot, and three month and one month. Finally, the weak exogeneity cannot be rejected for the one month contract in the formulation containing one month and spot.

Overall, it appears that futures prices lead the spot prices, and that futures prices on contract of longer time to maturity lead futures contracts with shorter time to maturity, therefore we can conclude that it is always the longest contract that binds the price series together in the long-run.

The fact that the price discovery first takes position in the futures market and then it is transmitted to underlying spot market could be explained by the fact that futures markets are different form spot markets in terms of cost of transaction, capital required and other issues. These issues suggest that futures markets would be forerunners of the spot markets as far as the information discounting is concerned.

4.6. Granger Causality Test - Do returns lead traders' positions?

In this section, we establish the methodology and present results that are used to determine whether crude oil futures traders' positions relate to crude oil futures prices. Granger causality test is used to determine whether returns lead traders' positions, and, vice versa, whether traders' positions lead returns. Finally, tests are conducted to determine if extreme trader positions impact energy futures market prices.

Before specifically examining the lead–lag relationships between traders' positions (PNL) and market returns (R), it would be worthwhile to examine the contemporaneous relationships between PNL_t and R_t. To measure this, simple correlation coefficients are calculated and presented in exhibit 24.

Exhibit 24 Contemporaneous Correlation Coefficients between Futures Returns and Percent Net Long Positions

Trader Category (PNL)	R (F1)	R (F3)	R (F6)
Reporting Noncommercial	0.162	0.165	0.161
Reporting Commercial	-0.196	-0.198	-0.190
Non-reporting	0.122	0.111	0.100
NOTE:	Sample period: 1993 (1) - 2008 (10). Simple correlation coefficients are calculated over 826 weekly observations. Using a two-tailed t-test, any correlation greater than 0.1 in absolute value is statistically significant at the 5% level.		

All of the correlation coefficients in exhibit 24 are statistically different from zero at the 5% level. The results indicate a positive contemporaneous correlation between the PNL for reporting non-commercials and returns and a negative relationship between reporting commercials and returns. That is, reporting non-commercials are net buyers in rising markets, while commercial hedgers are net sellers. It is not surprising that correlations are opposite of sign since the market as a whole must hold a neutral net position.

In the following, the lead–lag relationship between net positions and returns is considered. The test is initiated by considering whether traders' positions relate to past price changes. That is, do traders adjust their positions based on market movement? Traders who buy following price increases or sell following price declines are considered to be trend followers or positive feedback traders. Conversely, traders who buy following price declines are considered to be value hedgers or negative feedback traders. In either case, it is interesting to understand how positions respond (if at all) to past market returns.

Hamilton (1984) suggests a bivariate Granger test for examining the lead–lag relationship between two series. The null hypothesis that futures returns do not lead trader positions is tested by running the OLS regression in eq. 28 and the null hypothesis is that $\lambda_j = 0$ for all j:

$$PNL_t = \alpha + \sum_{i=1}^{m} \gamma_i PNL_{t-i} + \sum_{j=1}^{n} \lambda_j R_{t-j} + \omega_t \quad (28)$$

The lag structure (m,n) in regression 28 is selected by estimating the models for all values of $i = 1,2,\dots,12$ and $j = 1,2,\dots,12$, and dropping lags that were insignificant at 10% level. This procedure removes all evidence of residual serial correlation. The model is tested for heteroskedasticity with White's test. If the model is heteroskedastic, then White's heteroskedastic consistent covariance estimator is used. The p-values from the F-test tests are presented in exhibit 25.

Exhibit 25 Granger Causality Test - Returns lead the Percent Net Long Positions

	F1	F3	F6
Rep Noncomm	(m=1,6,7); (n=1, 2)	(m=1,6,7); (n=1, 2)	(m=1,6,7); (n=1, 2)
	F(2,820) = 33.03 [0.00]** (JHCSE)	F(2,820) = 31.30 [0.00]** (JHCSE)	F(2,820) = 25.58 [0.00]** (JHCSE)
Rep Comm	(m=1); (n=1,2,7)	(m=1); (n=1,2,7)	(m=1); (n=1,2,7)
	F(3,825) = 17.30[0.00]**	F(3,821) = 16.95 [0.00]**	F(3,821) = 13.767 [0.00]**
Nonrep	(m=1,2,4); (n=1, 2, 5, 7)	(m=1,2,4); (n=1, 2, 5, 7)	(m=1,2,4); (n=1, 2, 5, 7)
	F(4,818) = 13.82 [0.00]**	F(4,818) = 13.16 [0.00]**	F(4,818) = 10.431 [0.00]**
NOTE:	The lag structure (m, n) specified for each OLS regression: $PNL_t = \alpha + \sum_{i=1}^{m} \gamma_i PNL_{t-i} + \sum_{j=1}^{n} \lambda_j R_{t-j} + \omega_t$ Only the significant lagged dependent and independent variable s are retained. The p-values from the *F-test* of the null hypothesis, $\lambda_j = 0$ for $\forall_j$. Rejection of the null hypothesis implies that returns lead the traders' positions.		

The null hypothesis is rejected at conventional levels in all cases. That is, the p-values suggest that the lagged returns are important in determining net positions held by non-reporting traders. Specifically, positive futures returns result in reporting non-commercials – they increase their net long position the following week as prices increase. This could be indicative of a class of trend followers who buy in rising markets.

Concerning the reporting commercials, results suggest that they increase their net long positions when prices fall. This could be indicative of a class of value hedgers who sell in rising markets. Otherwise, this could be an expression of the data constraint that long positions must equal short positions, where negative feedback commercial traders take the opposite position of positive feedback noncommercial traders.

Next, I proceed with testing whether traders' positions can predict subsequent market returns. If they do, traders may develop profitable trading strategies and impact the market returns. This is tested by running the Granger Causality test represented by the regression:

$$R_t = \mu + \sum_{i=1}^{m} \psi_i R_{t-i} + \sum_{j=1}^{n} \varphi_j PNL_{t-j} + \omega_t \ (29)$$

In equation (29) the series PNL_t is said to lead futures returns R_t if they are useful in predicting R_t. The null hypothesis that PNL_t does not lead R_t , $H_o: \varphi_j = 0$ for $\forall_j$ is tested with a F-test. The p-values for testing the null that PNL_t does not lead $R_t$are presented in exhibit 26.

Exhibit 26 Granger Causality Test - Percent Net Long Positions lead Returns

	F1	F3	F6
Rep Noncomm	(m=3,7); (n=1,2) F(2,821) = 2.56 [0.08]	(m=3,7); (n=1,2) F(2,821) = 2.0918 [0.1241]	(m=3,7); (n=1,2) F(2,821) = 1.8737 [0.1542]
Rep Comm	(m=3,7); (n=1,2) F(2,825) = 3.38 [0.03]*	(m=3,7); (n=1,2) F(2,821) = 2.82 [0.06]	(m=3,7); (n=1,2) F(2,821) = 2.73 [0.07]
Nonrep	(m=1,7); (n=2) F(1,822) = 1.6315 [0.2019]	(m=1,7); (n=2) F(1,822) = 1.3361 [0.2481]	(m=1,7); (n=2) F(1,822) = 1.2798 [0.2583]
NOTE:	The lag structure (m, n) specified for each OLS regression: $R_t = \mu + \sum_{i=1}^{m} \psi_i R_{t-i} + \sum_{j=1}^{n} \varphi_j PNL_{t-j} + \omega_t$ Only the significant lagged dependent and independent variable s are retained The p-values from the *F-test* of the null hypothesis, $\varphi_j = 0$ for $\forall_j$. Rejection of the null implies that traders' positions lead the returns.		

From exhibit 26 there is little evidence to reject the null that reporting non-commercial positions do not Granger cause the returns at the 5% level; or, more explicitly, non-commercial positions do not lead returns. This implies that reporting noncommercial or "fund" positions do not contain any predictive information about returns. One possible interpretation of the results is that funds do not increase long (short) positions prior to rising (falling) futures prices. That is, funds do not exhibit systematic forecasting ability over 1-week intervals.

For commercial traders, the p-value is significant at 5% level, though it loses of its' significance the greater maturity on the futures contract gets.

Further, the results for non-reporting traders yield same conclusion as for reporting non-commercial traders- non-reporting positions do not contain any predictive information about returns.

In summary, the Granger causality tests suggest the following. First, there is no evidence that traders' (net long) positions contain general predictive information about market returns. Second, there is consistent evidence that positive futures returns cause the net long positions held by noncommercial traders to increase. Conversely, commercial traders show a tendency to be net sellers of futures positions the week following an increase in prices. The results for non-reporting traders are mixed.

Finally, the evidence that speculators are potentially trend followers is consistent with similar work using sentiment indices in energy futures markets (Sanders, 2000).

Impact of extreme trader positions

It is argued that only extreme level of traders' positions may impact the market. To test this assertion, I follow Wang's framework and define an extreme position as the upper and lower 20^{th} percentile of the prior 3-year range. So, the extreme level dummies are defined as follows:

LO=1 if PNL is in the lower 20^{th} percentile of its range from the prior 3 years, and
LO=0 otherwise.
HI=1 if PNL is in the upper 20^{th} percentile of its 3-year range, and
HI=0 otherwise.

Exhibit 27 Lower and Upper 20^{th} percentile for Traders' Positions

The percent net long (PNL):	Rep. Noncomm.	Rep. Comm.	Nonrep.
Lower 20th percentile	-0.193 %	-3.893 %	-11.207 %
Higher 20th percentile	7.397 %	0.957 %	-1.509 %
NOTE:	Sample period: November 2005 – October 2008 The percent net long (PNL) positions are calculated as the long position minus the short position divided by their sum.		

The following OLS regression is then used to test the impact of extreme positions on market returns:

$$R_t = \propto_o + \propto_1 LO_{t-1} + \propto_2 HI_{t-1} + \varepsilon_t \qquad (30)$$

The null hypothesis that extreme positions do not impact market returns $\propto_1 = \propto_2 = 0$ is tested with a F- test. Equation 30 is a version of the market timing test proposed by Cumby and Modest (1987). Within their framework the market timing test is a difference in means test, where $\propto_o + \propto_1$ is the mean return conditioned on extremely low net long positions, and $\propto_o + \propto_2$ is the expected return following extremely high net long positions. If the mean return conditioned on extremely short positions $\propto_o + \propto_1$or extremely long positions $\propto_o + \propto_2$ is

different from the unconditional mean $\propto_o$, then extreme PNL positions are useful in forecasting market returns.
The estimates of equation 30 for individual markets are presented in exhibit 28.

Exhibit 28 Extreme Level Regressions 2005(11) –2008(10)

Regression:

$$R_t = \propto_o + \propto_1 LO_{t-1} + \propto_2 HI_{t-1} + \varepsilon_t$$

		F1	F3	F6
Reporting Noncommercial	Constant	0.0003	0.0009	0.0015
	LO=1	0.0047	0.0039	0.0028
	HI=1	-0.0009	-0.0028	-0.0037
	p-value	0.8679	0.8424	0.8344
Reporting Commercial	Constant	0.0032	0.0035	0.0039
	LO=1	-0.0047	-0.0059	-0.0063
	HI=1	-0.0060	-0.0061	-0.0066
	p-value	0.7826	0.7219	0.6605
Nonreporting	Constant	-0.0027	-0.0022	-0.0015
	LO=1	0.0020	0.0030	0.0038
	HI=1	0.0162	0.0129	0.0098
	p-value	0.2396	0.3771	0.5385

NOTE:

Sample period: Nov 2005 – Oct 2008, with 156 weekly observation

Null hypothesis: $\propto_1 = \propto_2 = 0$
p-value from F-test that all slope coefficients equal zero

$$F - test = \frac{(R^2_{UR} - R^2_R)/r}{(1 - R^2_{UR})/(t-k-1)}$$

where subscripts UR and R distinguish between the R^2 from the unrestricted and restricted regression models. The number of restrictions being tested is r, and k is the number of explanatory variables in the unrestricted regression.

In no case does the F-test reject the null hypothesis that extreme position levels do not predict returns (or, there is no timing ability) at the 5% level. Though the coefficients are not statistically significant, an illustration of how to read the results follows: for instance, when non-commercial traders show an extremely small PNL position, the resulting week's return is a statistically positive 0.5% (0.0003+0.0047). Likewise, when non-commercial traders have a relatively large PNL position, subsequent weekly return is a statistically negative 0.06% (0.0003-0.0009). Or, when non-reporting traders have a relatively large PNL position, subsequent weekly return is a statistically positive 1.35% (-0.0027+0.0162).
The main conclusion form exhibit 28 is that there is no evidence that extreme position levels are consistent predictors of price movement, or there is no evidence of a systematic pattern of price continuation.

CONCLUSIONS

Was the recent rise in oil prices generated by changes in market fundamentals or speculation? Most probably the recent price spike was generated by both. This should not be too surprising, despite efforts by politicians or the popular press to spotlight one or the other. Increasing demand from the emerging countries along with stagnant production levels changed the supply-demand balance in a way that required higher prices to clear the market. This change in market fundamentals was recognized by speculators, who reckoned that ongoing changes in fundamentals would raise prices further, and took positions accordingly. This was reinforced by the fact that oil as a commodity has increasingly been viewed as a way of diversifying portfolio risk and hedging against future inflation and exchange rate fluctuations. The increase of market participants drove price beyond levels justified by the existing supply/demand balance. This increase slowed demand directly by raising energy prices and indirectly by slowing rates of economic activity. Consequently, this allowed prices to decline towards levels that are consistent with the existing supply-demand balance.

The fundamental purpose of this paper was to determine empirically the causal relationship between spot and futures prices and between futures prices and traders' futures positions; this way being able to draw tentatively a conclusion on whether there any possibilities for noncommercial traders to influence the price in the oil market. These results have implications for various players in the oil market such as international organizations, oil companies and governments when making investments decisions and policy recommendations.

I investigated (1) the long-run equilibrium relationship between the futures price and the spot price; (2) the long- and short-run efficiency of futures market as an unbiased predictor of spot prices; (3) lead-lag relationship between spot and futures prices; (4) the causality relationships between net traders' futures positions and market returns and (5) the impact of extreme traders' positions on market returns.

Concerning the test of market efficiency hypothesis it was found that while a market may appear to be long-term efficient, the same market may indeed be inefficient in the short-term, thus allowing the possibility of speculation/arbitrage opportunities. The long-run efficiency of the crude oil futures market was tested using the Johansen Maximum Likelihood procedure and short-run efficiency is examined by constructing and investigating an error correction model proposed by Engle and Granger. The study of market efficiency in oil futures market is important to both the regulators and the producers/hedgers. From the regulator point of view, an efficient market means a better alternative to market interventions such as imposing price stabilization policies. For marketers, it provides a reliable forecast of spot prices in the future to allow them effectively manage their risks in the production or investment process.

When examining the lead-lag relationships between spot price and futures price of different maturities, it was found that the futures market leads the spot market and futures prices tend to discover new information more rapidly than spot prices, furthermore, that futures prices on contract of longer maturity lead futures contracts with shorter maturity. This test was carried out using the Granger causality framework. The results show the importance of taking into account the long-run relationship between the futures and the spot prices in forecasting future spot prices.

Regarding the causality relationships between net traders' futures positions and market returns, a positive correlation was found between returns and positions held by noncommercial traders, and a negative correlation between commercial positions and market returns. Furthermore, positive returns result in an increase in noncommercial net positions in the following week, whereas the net long positions held by commercial hedgers decline following price increases. Generally, it can be concluded that traders' net positions do not lead market returns in general, implying that trader's positions do not contain any predictive information about returns. For this test I employed the Granger causality framework.

When testing the impact of extreme traders' positions on market returns no evidence was found that extreme position levels are consistent predictors of price movement, implying that there is no evidence of a systematic pattern of price continuation. This test was carried out using the market timing framework proposed by Cumby and Modest (1987).

Overall, the results obtained suggest that any long-term trends in oil prices have been dictated by market fundamentals rather than by investors' sentiment.

However, this conclusion might be misleading due to the limitations of the test used. First, the tests do not account for a possible short-run time varying risk premium. Second, the tests above assume that oil prices are characterized by constant variance. Third, current CFTC classification of traders in the futures markets does not allow reliable distinction between purely financial and hedging-related activities. Moreover, even if trading in futures markets were to be reported by type of activity rather than type of entity, it would not provide a complete picture of the activities of financial investors. The coverage of existing data does not extend to off-exchange/over the-counter (OTC) activities.

Overall, this study has opened up more interesting research questions regarding the efficiency of oil futures markets as well as the formation of oil price and its driving factors. First, a further study could explore the role of a time varying risk premium into the model of the efficiency and unbiasedness of futures prices. Second, as oil prices exhibit extensive volatility over the sample period, attention has to be given to price volatility in examining futures market efficiency. Lastly, an analysis of the influence of the fundamental factors and monetary policy are warranted

References

Asche, F., Guttormsen A.G., (2002), Lead lag relationships between futures and spot prices.

Alexander, C., (2008) Practical Financial Econometric-Market Risk Analysis II. John Wiley & Sons, ltd

Beninga, S., 2008, Financial Modeling, (3rd Edition) The MIT Press

Bodie. , Kane, Marcus, Investmnets (7th Edition), Mc GRAW – Hill International Edition

Brooks, C., (2008), Introductory Econometrics for Finance, Second Edition, Cambridge University Press, UK

Cabinet Office UK, Global Energy team, The rise and fall o foil prices; analysis of fundamental and financial drivers, December, 2008

Charemza, W.W, Deadman, D.F. (1997) New Directions in Econometric Practice: General to Specific Modelling, Cointegration, and Vector Autoregression, Cheltenham : Elgar

Cuthbertson, K. and Nitzsche, D. (2004). Quantitative Financial Economics-Stocks, Bonds & Foreign Exchange, West Sussex: John Wiley & Sons Ltd.

Commodity Futures Trading Commission
http://www.cftc.gov

Dickey, D.A. and Fuller, W.A., (1979). "Distribution of the Estimators for Autoregressive Time Series with a Unit Root," Journal of the American Statistical Association, Vol. 74, pp 427-431.

Doornik, J. A., and Hendry, D. F. (2007). Empirical Econometric Modelling- PcGive 12, OxMetrics 5: Volume I, 5th ed., London: Timberlake Consultants Press.

Dimitris Kenourgios & Aristeidis Samitas, 2005. "Testing efficiency of the copper futures market: new evidence from London Metal Exchange," Finance 0512010, EconWPA.

Dimitris Kenourgios, 2005. "Testing efficiency and the unbiasedness hypothesis of the emerging Greek futures market," Finance 0512015, EconWPA

Engle, R., Granger, C. (1987). Co-integration and error-correction: representation, estimations and testing, Econometrica 55, 251-276.

Enders, W., (1995), Applied Econometrics Time Series, First Edition, Wiley Series in Probability and Mathematics Statistics, John Wiley & Sons, Canada

Energy Information Administration, 2008. Glossary.
http://www.eia.doe.gov/glossary/index.html.

Energy Information Administration, 2004. Monthly Energy Chronology.
http://www.eia.doe.gov/cabs/chrn2004.html.

Fama, E., 1970. Efficient capital markets: a review of theory and empirical work. Journal of Finance 25, 383-417.

Fama, E., French, K., 1987. Commodity futures prices: some evidence on forecast power, premium and the theory of storage. Journal of Business 60, 55-73.

Fama, E., 1991. Efficient capital markets: II. Journal of Finance 46, 1575-1617

Farrell, G. N., Kahn, B. and Visser, F. J. (2001) Price determination in international oil markets: developments and prospects, Quarterly Bulletin, March 2001, South African Reserve Bank

Fattouh, B. (2007) The Drivers of oil prices: the usefulness and limitations of non-structural model, the demand-supply framework and informal approach, Oxford Institute for Energy Studies

Fujihara, R.A., Mougoue, M. (1997), Linear dependence, nonlinear dependence and petroleum futures market efficiency. The Journal of Futures Markets, 17, 75-99

Johansen, S. (1988), Statistical analysis of cointegrating vectors, J. Econ. Dyn. Control 12, 231-254.

Johansen, S. (1991) Estimation and hypothesis testing of cointegrating vectors in Gaussian vector autoregressive models, Econometrica 59, 1551-1580.

Johansen, S. (1995), Likelihood-based inference in co-integrated vector autoregressive models, Oxford University Press, Oxford.

Haigh, M., S. (1998), Co-integration analysis of unbiased expectations in the BIFFEX freight futures market, 1998 Annual meeting, August 2-5, Salt Lake City, UT 20873, American Agricultural Economics Association.

Haigh, M., S., Hranaiova, J., Overdahl, (2005) Price volatility, liquidity provision and the role of managed money traders in energy futures markets, U.S. Commodity Futures Trading Commission & bPublic Company Accounting Oversight Board

Harris, R., Sollis, R. (2003) Applied Time Series Modelling and Forecasting, Chichester : Wiley

Hendry, D.F. and Juselius K. (1999). Explaining cointegration analysis: Part I, Energy Journal, Vol. 21.

Hendry, D.F. and Juselius K. (2000). Explaining cointegration analysis: Part II, Energy Journal, Vol. 21.

Hill, R. C., Griffiths, W. E., Lim, G. C. (2008) Principles of econometrics, Hoboken, N.J. : John Wiley

Hull, J. C., (2005), Fundamentals of futures and options markets, 5th Edition, International Edition, Pearson Education, International.

Kaufmann, R. K, Ullman, B. , (2008), Oil prices, speculation, and fundamentals: Interpreting causal relations among spot and futures prices, Energy Economics, Elsevier.

Kaldor, N. (1939), Speculation and economic stability, The Review of Economic Studies, Vol. 7, No. 1 (Oct., 1939), pp. 1-27.

Kellard, N., Newbold, P., Rayner, T., Ennew, C., (1999), The relative efficiency of commodity futures, The Journal of Futures Markets, Vol 19, No 4, 413-432.

Koop, G. (2008) Introduction to Econometrics, Chichester , West Sussex : Wiley

Krichene Noureddine, (2006) World crude oil markets: monetary policy and the recent oil shock, IMF Working Paper, International Monetary Fund

P. Perron, (1989), The great crash, the oil price shock, and the unit root hypothesis, Econometrica 57, 1361-1401.

Perron, P. (1994) Trend, unit root and structural change in macroeconomic time series, cointegration for applied economist, Macmillian Press, Basinggstoke, UK

Sander, D. R., Boris, K., Manfredo, M. (2004) Hedgers, funds, and small speculators in the energy futures markets: an analysis of the CFTC's Commitments of Traders reports, Elsevier Science Inc.

Serletis, A., (2007) Quantitative and Empirical Analysis of Energy Markets , World Scientific

Stevans, K. L., Sessions, D. N. (2008) Speculation, futures prices and the U.S. real price of crude oil, SSRN

Weiner, R. J. (2002) Sheep in wolves' clothing? Speculators and price volatility in petroleum futures, Elsevier Science Inc.

Wooldridge, J.M., (2006). Introductory Econometrics – A Modern Approach, 3rd ed., Mason: Thomson South-Western.

http://www.sciencedirect.com/science/article/B6W5X-45MW05T-D/2/3fe7531fbbfdb448be452c5d1e7b6107

http://www.bushagenda.net/article.php?id=546

http://www.cftc.gov/newsroom/speechestestimony/opajeffery-12.html

Appendix

Exhibit 29 Level Representation of Spot and Futures Prices, Basis and Premium

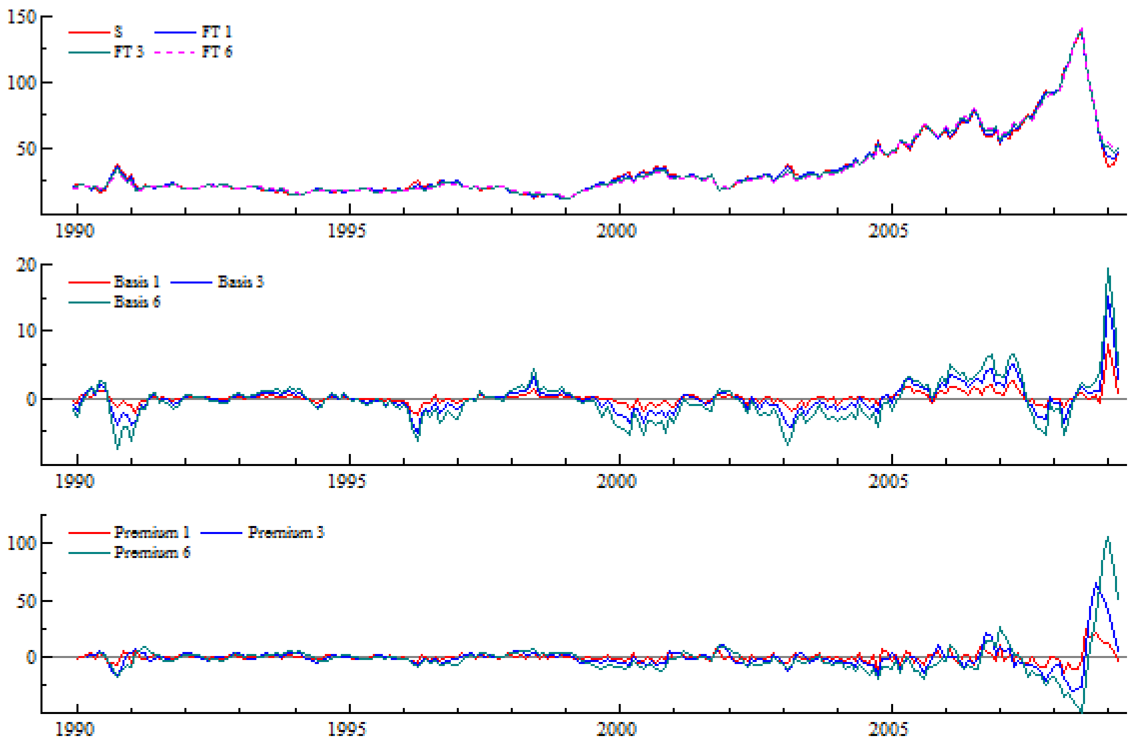


Exhibit 30 Fama's Change Regression- 1992(1) - 2002(12)

Regression:
$S_t - S_{t-1} = \alpha + \beta\ [F1_{t-1} - S_{t-1}] + e_t$

Diff Log S_1	coeff	t-value	p-value
Constant	0.004	0.473	0.637
Basis% _1	0.327	1.020	0.307
R^2	0.008		
F(1,125) =	1.05 [0.307]		

Regression:
$S_t - S_{t-3} = \alpha + \beta\ [F3_{t-3} - S_{t-3}] + e_t$

Diff Log S_3	coeff	t-value	p-value
Constant	0.012	0.838	0.404
Basis% _3	0.291	1.220	0.227
R^2	0.012		
F(1,125) =	1.477 [0.227]		

Regression:
$S_t - S_{t-6} = \alpha + \beta[F6_{t-6} - S_{t-6}] + e_t$

Diff Log S_6	coeff	t-value	p-value
Constant	0.024	1.210	0.229
Basis% _6	0.282	1.280	0.202
R^2	0.013		
F(1,125) =	1.643 [0.202]		

NOTE: The sample spans over 1992(1) - 2002(12); Observations: 127.
Since the first six observations were used to compute the lagged variables,
the regression sample is running from 1990(6) – 2008(10); thus, the regressions are based on 221 observations.
The numbers in parenthesis represent the probabilities of falsely rejecting the null hypothesis (i.e., p-value)
**and *denote 1% and 5% level of significance, respectively.

The dependent variable is the change in the spot price and independent variable is the basis (prices are expressed as natural logs):
$\Delta S1_t = S_t - S_{t-1}$, $\Delta S3_t = S_t - S_{t-3}$ and $\Delta S6_t = S_t - S_{t-6}$
$Basis\%\_1 = F1_{t-1} - S_{t-1}$, $Basis\%\_3 = F3_{t-3} - S_{t-3}$, and $Basis\%\_6 = F6_{t-6} - S_{t-6}$

Exhibit 31 Restrictions - Change Regression- 1992(1) - 2002(12)

Ho:	F-test	S1/F1	S3/F3	S6/F6
$\propto_1 = 0$	F(1, 125)	0.22 [0.637]	0.70 [0.404]	1.46[0.23]
$\beta_1 = 1$	F(1, 125)	4.43 [0.034]*	8.75 [0.004]**	10.62[0.001]**
$\propto_1 = 0, \beta_1 = 1$	F(1, 125)	2.57 [0.081]	5.55 [0.005]**	7.60[0.001]**

NOTE: The sample spans over 1992(1) - 2002(12); Observations: 127.
The numbers in parenthesis represent the probabilities of falsely rejecting the null hypothesis (i.e., p-value)
**and *denote 1% and 5% level of significance, respectively.

Exhibit 32 Fama's Premium Regression - 1992(1) - 2002(12)

Regression:
$[F1_{t-1} - S_t] = \alpha + \beta[F1_{t-1} - S_{t-1}] + e_t$

Premium%1	coeff	t-value	p-value
Constant	-0.004	-0.473	0.637
Basis% _1	0.673	2.100	0.037
R^2	0.034		
F(1,125) =	4.431 [0.037]*		

Regression:
$[F3_{t-3} - S_t] = \alpha + \beta[F3_{t-3} - S_{t-3}] + e_t$

Premium%3	coeff	t-value	p-value
Constant	-0.012	-0.838	0.404
Basis% _3	0.709	2.960	0.004
R^2	0.07		
F(1,125) =	8.748 [0.004]**		

Regression:
$(F6_{t-6} - S_t) = \alpha + \beta[F6_{t-6} - S_{t-6}] + e_t$

Premium%3	coeff	t-value	p-value
Constant	-0.024	-1.210	0.229
Basis% _6	0.718	3.260	0.001
R^2	0.08		
F(1,125) =	10.62 [0.001]**		

NOTE: The sample spans over 1992(1) - 2002(12); Observations: 127.
The numbers in parenthesis represent the probabilities of falsely rejecting the null hypothesis (i.e., p-value)
**and *denote 1% and 5% level of significance, respectively.

The dependent variable is the change in the spot price and independent variable is the basis (prices are expressed as natural logs):
$Premium\%\_1 = F1_{t-1} - S_t$ $Premium\%\_3 = F3_{t-3} - S_t$ and $Premium\%\_6 = F6_{t-6} - S_t$
$Basis\%\_1 = F1_{t-1} - S_{t-1}$ $Basis\%\_3 = F3_{t-3} - S_{t-3}$ and $Basis\%\_6 = F6_{t-6} - S_{t-6}$

Exhibit 33 Restrictions - Premium Regression-1992(1) - 2002(12)

Ho:	F-test	S1/F1	S3/F3	S6/F6
$\propto_2 = 0$	F(1, 125)	0.22[0.637]	0.70[0.404]	1.46[0.23]
$\beta_2 = 1$	F(1, 125)	1.05[0.31]	1.48[0.23]	1.64[0.202]
$\propto_2 = 0, \beta_2 = 1$	F(1, 125)	0.57[0.566]	0.91[0.406]	1.23[0.296]

NOTE: The sample spans over 1992(1) - 2002(12); Observations: 127.
The numbers in parenthesis represent the probabilities of falsely rejecting the null hypothesis (i.e., p-value)
**and *denote 1% and 5% level of significance, respectively.

Exhibit 34 Unit Root Test with Structural Break – Model A

Model (A)

$$\Delta y_t = \mu^A + \theta^A DU_t + \beta^A t + d^A D(T_B)_t + \alpha^A y_{t-1} + \sum_{i=1}^{k} c_i \Delta y_{t-i} + e_t$$

y=	S		F1		F3		F6	
	coeff	t-value	coeff	t-value	coeff	t-value	coeff	t-value
Constant (μ^A)	0.104	0.553	0.187	0.553	0.107	0.652	0.106	0.685
Trend (t)	0.001	1.150	0.001	1.070	0.001	0.945	0.000	0.830
y_{t-1}	-0.013	**-2.440**	-0.012	**-2.420**	-0.011	**-2.400**	-0.010	**-2.380**
Δy_{t-1}	0.005	0.140	0.027	0.751	0.058	1.620	0.087	2.460
$D(T_B)_t$ (mars03)	-0.340	-0.268	-0.033	-0.028	-0.208	-0.186	-0.246	-0.234
DU_t (mars03)	0.288	0.969	0.302	1.070	0.329	1.250	0.345	1.390

NOTE: **Sample Period:** 1993-01-05 - 2008-10-28
Critical Values: 5%=-3.80, 1%=-4.42