Speculation and Volatility in the Crude Oil Futures Market

Yun Pan

A Thesis

in

The John Molson School of Business

Presented in Partial Fulfillment of the Requirements
for the Degree of Master of Science in Administration at
Concordia University
Montreal, Quebec, Canada

June 2009

Your file Votre référence
ISBN: 978-0-494-63096-9
Our file Notre référence
ISBN: 978-0-494-63096-9

Abstract

Speculation and Volatility in the Crude Oil Futures Market

Yun Pan

Speculative activity is often claimed to be the cause of high price and turmoil in the crude oil futures market. We use Working's T and a new speculative index to measure the degree of speculation in the crude oil futures market. We find that the variation in the volatility of the crude oil futures market can be explained by the variation in the speculative index, but the economic significance of this relationship is quite low. Granger causality tests show that changes in the degree of speculation lead to changes in volatility in the crude oil futures market, but reverse relationship does not hold. In addition the causality effect is weak. After the Commodity Futures Modernization Act was enacted, speculative activity played a more important role in explaining futures market volatility, and other fundamental factors' explanatory power decreased significantly.

Acknowledgements

It is difficult to acknowledge all the individuals who have supported me to complete this thesis. However, the debt is real and so are my thanks. I would first and foremost thank Dr. Latha Shanker for her timely guidance, inspirational instruction and incredible patience through the difference phases of this thesis. Special thanks to Dr. Gregory Lypny and Dr. Nilanjan Basu for being part of this committee. I would also like to thank all my friends and classmates for their support and encouragement. Last my thanks would go to my beloved family for their affections and considerations.

Table of Contents

List of Figures

List of Charts and Tables

Chapter 1 Introduction

In recent years, the public has seen a sharp increase in the price of crude oil. The price has increased from $10 per barrel in 1998 to over $145 per barrel in July 2008 (see Figure 1). Because crude oil is the "black blood" of the industry, its historical high price has a huge impact on the economy and everyone's life, thus leading to a hot debate about the cause and consequences of this phenomenon.

Some people argued that market fundamentals are the cause for high crude oil prices. The supply is stable, and the demand is continuously increasing. Not having found big oil fields for more than two decades decreases the probability of future increases in supply. The huge increase in demand from emerging markets is the main force that has driven the price to record highs.

Other people claimed that such a rapid increase cannot be fully explained by market fundamentals. Because of the price discovery mechanism of the crude oil futures market, the industry has been viewing the futures price as a reliable estimate of the future spot price. The crude oil futures price is actually used as a benchmark for the spot price. They argued that rampant speculation in the crude oil futures market resulted in historical high futures prices, which in turn led to high spot prices (e.g., Masters (2008a, 2008b)).

Each party to the debate brought some data to the public to support their arguments. However, these data are more of a descriptive nature. In order to obtain a more reliable picture of the dynamics in the crude oil market, we need to identify the essence of market variables and investigate the relationship between these variables.

In order to quantitatively measure the degree of speculation in crude oil futures markets, we recall a traditional speculative measure - Working's T, and review its characteristics and applications. We also identify some shortcomings of Working's T, and present a new speculative index, which has a straight-forward derivation and relatively clear economic meaning. Based on very limited information released by the Commodity Futures Trading Commission (CFTC), it is difficult to derive a perfect speculative index. However, we believe, using Working's T and our new speculative index together, we can more reliably measure the degree of speculative activity.

Some researchers have addressed the relationship between speculation and market price movement in crude oil futures market (e.g., Sanders, Boris and Manfredo (2004)). Our focus is to examine the relationship between degree of speculation and futures price volatility. We want to know whether the variation in the degree of speculation explains the volatility of the futures market. If such a relationship exists, we also want to identify the cause-effect relationship.

In addition, the Commodity Futures Modernization Act (referred to as CFMA hereafter) is claimed to have created the so-called "Enron Loophole". Big banks make over-the-counter (OTC) transactions with institutional investors, and then use massive long positions in commodity futures contracts to offset their short positions that result from their OTC transactions. These banks are qualified as bone fide hedger under the CFMA, but they actually act as an intermediary for so-called index traders to speculate or invest. We want to investigate whether there is a structural change in the crude oil futures market after the CFMA was enacted. Because the CFTC only releases the Commitment of Traders (COT) report for the crude oil futures market, with index traders still classified as hedgers, we can not separately test the behavior of index traders, or re-classify index traders into the speculative category. This limits the strength of our tests. However, we believe it is still beneficial to examine the effect of the CFMA under the current data constraints.

The rest of the thesis is organized as follows. Chapter 2 provides a brief review of the literature. Chapter 3 describes the data. Chapter 4 presents the methodology. The empirical results and discussion are provided in Chapter 5. The paper concludes in Chapter 6.

Chapter 2 Literature Review

2.1 A Traditional Speculative Measure - Working's T

The development of Working's T

After examining some agricultural futures markets, Working (1960) found that the amount of long speculation, measured by the dollar value of open speculative contracts, differed greatly between commodities, and the difference in the amounts of speculation depend mainly on the amounts of hedging in the market. He also found a close correspondence between changes in the amounts of speculation and hedging. These facts led Working to the conclusion that futures markets are primarily hedging markets.

Working's analysis was based on the regularly published COT reports by the then-current administration. However, because of minimum reporting levels, the positions of non-reporting traders were unknown. The non-reporting traders accounted for a large proportion of the futures markets, thus influencing the ability to make inference. In order to alleviate this problem, Working had developed a technique to allocate the unreported positions to reporting categories, and used a special survey with complete information on positions to test the validity of his allocation method.

Working claimed that, because long hedging and short hedging cannot exactly match in time, they are partially offset. The un-offset part needs to be covered by speculation. The long speculation provided support to short hedging that is not offset by long hedging, and short speculation offered support to long hedging that is not offset by short hedging. Then Working argued that total hedging, rather than net hedging is the better measure of hedging demand. Based on this inference, through regression analysis and mathematical derivation, Working presented a speculative index (which was referred to as Working's T by later researchers) as follows:

$$T = 1 + \frac{SS}{HS + HL} \text{ when HS} \geq \text{HL}$$

$$T = 1 + \frac{SL}{HS + HL} \text{ when HS} < \text{HL}$$

SS refers to Short Speculation, SL refers to Long Speculation, HS refers to Short Hedging, and HL refers to Long Hedging.

The above formula has the form of scaling by total hedging, which was consistent with Working's inference that total hedging is the measure of hedging demand. Working calculated the five-year average speculative index for eleven agricultural futures markets, and found that the overall degree of speculation in agricultural markets is quite low.

Working's T has been widely used in previous research. However, there are some problems with this index. First, the derivation of the index is not straight-forward. Hence, we do not know the exact economic meaning of the index, thus limiting its use. In addition, Working's T is better used as a relative measure. For example, if Working's T for a specific market is 1.3, we need to compare this number with historical numbers, and then conclude whether the degree of speculation has changed. If Working's T increases by 10%, can we infer that the degree of speculation has increased by 10%? It is difficult to say because we do not know the exact meaning of Working's T. Second, the total amount of hedging has been cited by many previous researchers to be a better measure of hedging demand or market activity than net hedging. However, theoretically speaking, the real demand for speculation is somewhere between total hedging and net hedging. The long hedging and short hedging are always partially offset. The long hedging and short hedging after offsetting provide the actual demand for short speculation and long speculation.

Chart 1 Read Hedging Demand in Futures Market

Short hedging:	sssssssss$\boxed{\text{sssssssssssss}}$
Long hedging:	lllllllllll$\boxed{\text{llllll}}$
Short hedging after offsetting:	$\boxed{\text{ssssssssssss}}$
Long hedging after offsetting:	$\boxed{\text{llllll}}$
Actual demand for Long speculation:	$\boxed{\text{llllllllllllllll}}$
Actual demand for Short speculation:	$\boxed{\text{ssss}}$

* The words in gray are matched and offset hedging.

Of course, the CFTC does not provide the details of positions, and we do not know the proportion of matched hedging. This ideal speculative index can not be developed with the currently provided COT data.

According to the above analysis, we expect a new speculative index which has straight-forward derivation and clear economic meaning. We can test how Working's T and this new speculative index behave, and how they explain or correlate to other important market variables.

Working (1960) opened the door to the quantitative investigation of the degree of speculation of commodity futures markets, primarily agricultural futures markets. The researchers who followed Working (1960) focused on two themes. The first theme was to examine the validity of Working's T, apply this index to assess the adequacy of speculation of a certain futures market, and investigate the relationship between the speculative index and other market variables. The second theme was to develop an estimation technique to examine the composition of unreported positions, in order to estimate a complete distribution of positions to facilitate further research.

The applications and discussions of Working's T

Peck (1979-1980) investigated the relationships between commercial use and total market activity in agricultural and non-agricultural futures markets. She found that total hedging is more correlated to total market participation measured by total open interest than is net hedging. When studying the seasonality of annually produced, continuously consumed commodities, she found that the huge growth in exports and hence in commercial long hedging has offset the storage-related, short hedging needs, and the commercial use is nearly balanced as in grain markets of interest. Though the pattern of futures markets has changed in many aspects, her study was still shown to "lend support to the Working definition of an appropriate measure of hedger demands upon a market. Net hedging is not the most useful view of the demands that commercial users make on a market. Speculation is needed to offset both long hedging and short hedging. Only coincidentally are long and short hedgers sufficiently alike in date and amount to be offsetting, although increased balance increases the probability of such correspondence and differences in seasonal needs between long and short hedgers decreases this probability. The appropriate measure of minimum required speculation must at least begin with total hedging demand." (1979-1980, p.339)

Peck (1981) examined changes in the levels of speculation in the wheat, corn and soybean futures markets from 1964 to 1977. She found significant declines in speculation in these markets. In order to investigate how the decline in speculation influenced the futures price variability, Peck regressed a proxy of price variability on the speculative index. Because the COT Report was released monthly at that time, she used the monthly average of daily trading ranges as a proxy for futures price variability. The control variables included the proxies for scalping and new information. Peck found a negative and statistically significant coefficient for the speculative index. Since decreases in speculation are significantly associated with increases in the average daily trading range, Peck argued that declines in speculation had a clear influence on market performance and speculation became inadequate. As the robustness tests, Peck used some other forms of speculative indices in the above regression, such as short speculation and long speculation, both in actual levels and as percentages of open interest. The results were quite similar.

Sanders, Boris and Manfredo (2004) examined the COT Reports for crude oil, unleaded gasoline, heating oil, and natural gas futures contracts from 1992 through 1999. They examined the collection procedures for the COT data and had the following findings.

First, the data provided no information about the motives of non-reporting traders except that their positions were below the CFTC reporting level. Second, pure hedge positions were a subset of reporting commercial category, and reporting commercial classification may reflect a diverse set of motives. Third, reporting noncommercial positions were probably the most precise category, which captured the positions of a pure subset of speculators.

Sanders, Boris and Manfredo (2004) then used Granger causality tests to check whether there exists a Granger causality relationship between trader positions and market prices. They found a positive correlation between market returns and noncommercial traders' positions, and a negative correlation between market returns and commercial traders' positions. In addition, positive returns led to an increase in noncommercial net long positions in the following week and a decrease in commercial net long positions. However, in general, traders' net positions do not have predictive power for market returns.

Sanders, Irwin and Merrin (2008) investigated the adequacy of speculation in agricultural futures markets. The CFTC provided more information about traders' positions through the Commitment of Index Trader and Bank Participation reports for agricultural futures markets, thus offering new opportunities for deeper research of relevant areas. They found that index fund positions have stabilized as a percent of total open interest after huge increases from early 2004 till mid-2005. No significant changes or shifts are revealed by traditional speculative measures like Working's T. They found, in most markets, an increase in short hedging had equaled or surpassed the increase in long speculation. The speculative measures adjusted for index trader positions are still within the historical ranges. Sanders, Irwin and Merrin inferred that long-only index funds may be a good thing in markets traditionally dominated by short hedging. The administration needs to consider the potential benefits offered by index traders when considering measures to curb speculation.

Estimation Technique for non-reporting positions

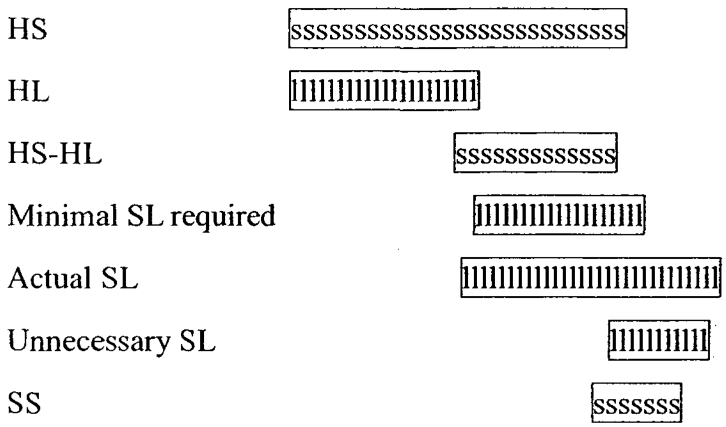
Larson (1961) and Rutledge (1979) suggested procedures to allocate the positions in the non-reporting category to hedging and speculation categories. Sanders, Irwin and Merrin (2008) just allocated the non-reporting traders' positions to the commercial, non-commercial, and index trader categories in the same proportion as observed for reporting traders.

2.2 A New Speculative Index

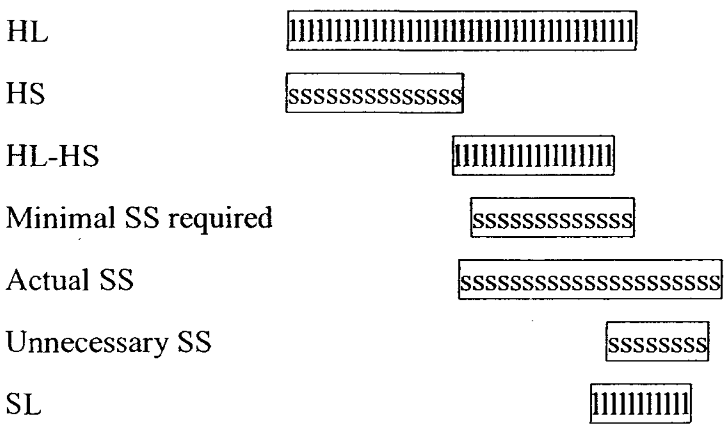
Dr. Latha Shanker proposed a more natural and straight-forward speculative index, which has clear economic meaning. This speculative index also relies on the findings that commodity futures markets are primarily hedging markets, and speculation follows hedging demands.

Chart 2 Derivation of a New Speculative Index

Case 1: When HS ≥ HL:



Case 2: When HL>HS:



* HL refers to Long Hedging, HS refers to Short Hedging, SL refers to Long Speculation, and SS refers to Short Speculation.

Let consider two cases. The first case is when short hedging dominates long hedging. The short hedging that is not offset by long hedging is HS-HL. In an ideal situation, the minimal SL required is equal to HS-HL. Of course, this ideal situation requires some assumptions that short hedging and long hedging have a perfect match in time, only differ in quantity, and long speculation has a perfect match in time and quantity with residual short hedging HS-HL. Of course, these assumptions of perfect matching over time are not realistic in practice. But it makes sense to be an ideal situation and the origination of a new speculative index. Because perfect matching will not occur, the actual long speculation is always greater than the minimal required long speculation, and the excess amount of long speculation equals to SL-(HS-HL) = SL+HL-HS = SS, just the amount of actual short speculation. We further scale this excess amount of long speculation by all the long positions required in the ideal situation, which just equals to short hedging. Then we get a new speculative index SS/HS.

The second case is when long hedging dominates short hedging. The long hedging that is not offset by short hedging is HL-HS. In an ideal condition, the minimal SS required is equal to HL-HS. This ideal situation requires same unrealistic assumptions as mentioned above. Because perfect matching will not occur, the actual short speculation is always greater than minimal required short speculation, the excess amount of short speculation equals to SS-(HL-HS) = SS+HS-HL = SL, just the amount of actual long speculation. We further scale this excess amount of short speculation by all the short positions required in the ideal situation, which just equals to long hedging. Then we get a new speculative index SL/HL.

Now let us compare this speculative index with Working's T in the case that HS $\geq$ HL.

New Speculative index $= \dfrac{SS}{HS}$

Working's $T = 1 + \dfrac{SS}{HS + HL}$

First, let us compare the origin of the indices. The minimum value for the new speculative is zero, which occurs when SS equals to zero. The minimum value for Working's T is 1, which also occurs when SS equals to zero. According to the market identity HS+SS = HL+SL. SS equals to zero only when HS is perfectly offset by HL and SL, which is unrealistic. In short, the minimum number for the two indices is unreachable in practice.

Second, the derivation of Working's T involved regression analysis and mathematical derivation, which is not clear and straight-forward. The derivation of the new speculative index only uses market clearing identities, which is much more straight-forward than that of Working's T.

Third, the new speculative index has a very clear economic meaning. We scale unnecessary long speculation by all ideal long positions in the market. If the new speculative index equals to 0.3, we know that unnecessary long speculation equals to 30% of short hedging in a market dominated by short hedging. If the new speculative index increases by 10%, we know that holding short hedging constant, the short speculation (or unnecessary long speculation) increases by 10%, or non-matching between short hedging and long hedging increases roughly by 10%.

On the other hand, considering Working's T, HS+HL in the denominator, appears to be a direct reflection of total hedging as a measure of hedging demand rather than net hedging, thus having some theoretical appeal. However, HS+HL has direct economic meaning only when HS and HL are perfectly unmatched. Being perfectly unmatched is even more unrealistic than being perfectly matched or offset. If we want to use actual total hedging demand to scale SS, we should use unmatched parts of HS+HL rather than HS+HL. And using HS+HL may lead to underestimation of the degree of speculation, and difficulties in explaining the economic meaning of the index and changes in the index. As we mentioned before, it is difficult to explain the economic meaning of a value for Working's T of 1.3 and a 10% increase in Working's T. In addition, SS can be considered as unnecessary SL, and is essentially based on the net hedging concept. However, HS+HL is based on total hedging concept. A fraction with SS in the numerator and HS+HL in the denominator has very confusing economic meaning.

Next let us compare the new speculative index with Working's T using numerical examples. We consider three cases.

First case, we consider the origins of the two indices. Suppose short hedging dominates long hedging. Long speculation is exactly equal to the demand by short hedging that is not covered by long hedging. HS=100, HL=50, SL=50, SS=0. In this case, there is no excess speculation.

Working's T = 1+0/(100+50)=1

New speculative index = 0/100=0.

Second case, long speculation exceeds the demand by short hedging that is not matched by long hedging. HS=100, HL=50, SL=100, SS=50.

Working's T=1+50/(100+50)=1.3333

New speculative index=50/100=0.5

Comparing the above results with the first case, Working's T gives the impression that the excess speculation is 33.33%, while the new speculative index shows that the excess speculation is 50%. In the formula of Working's T, HS+HL=150 cannot be considered to be an accurate measure of total hedging demand.

Third case, long speculation exceeds the demand by short hedging that is not matched by long hedging and increases to 200. HS=100, HL=50, SL=200, SS=150.

Working's T = 1+150/(100+50)=2

New speculative index = 150/100=1.5

Comparing the above results with the first case, Working's T gives the impression that the excess speculation is 100%, and the new speculative index shows that the excess speculation is 150%

These three cases illustrate that Working's T may underestimate the degree of speculation in the futures market.

2.3 Degree of Speculation and Volatility

Using daily and weekly data for crude oil, heating oil and gasoline, Pindyck (2004a) examined the role of volatility in short-run commodity price dynamics and the determinants of volatility. He found that the market variables (such as spot price, inventories, production and convenience yield) and macroeconomic variables (such as interest rates) do little to explain the behavior of price volatility for crude oil and heating oil, and that volatility can be viewed as exogenous.

There are many other papers which addressed the volatility in the crude oil spot or futures market. Some papers analyzed the relationship between crude oil price volatility and other market variables such as volume, supply and demand (e.g., Foster (1995), Yang, Hwang and Huang (2002)). Some papers examined the explanatory power of speculative index to some price variability measures in agricultural futures market (e.g., Peck (1981), Leuthold (1983)). However, we have not found papers directly investigating the relationship between degree of speculation and volatility for crude oil futures markets.

Chapter 3 Data

3.1 Commitment of Traders Data for Crude Oil Futures

We downloaded COT data for crude oil futures from the website of the CFTC, and obtained 710 weekly observations from March 21st 1995 to March 10th 2009. The report date (as of date) of COT data is every Tuesday, and the release date is every Friday. CFTC.gov says: "The COT reports provide a breakdown of each Tuesday's open interest for markets in which 20 or more traders hold positions equal to or above the reporting levels established by the CFTC. The weekly reports for Futures-Only Commitments of Traders and for Futures-and-Options-Combined Commitments of Traders are released every Friday at 3:30 p.m. Eastern time." These 710 COT observations are used to calculate speculative indices.

We checked the reliability of the COT data, and found three problems. First, 16 continuous observations (between January 25th 2000 and May 23rd 2000) are missing. The response from CFTC said that the delta factors are needed to calculate the Futures-and-Options-Combined COT reports. Because the delta factors supplied by the NYMEX on its daily tape to the CFTC did not agree with those published by the exchange for this period, the COT reports for that period were not officially released.

Second, among a total of 710 observations, eleven observations have an as of date which is not Tuesday, i.e., eight observations on Monday, two on Thursday, and one on Friday. But they are in the same report week. Third, two Tuesday (as of date) reports are replaced by advance Friday (as of date) reports respectively, thus leading to two observations in one week. We just deleted these two advance Friday observations in order to avoid un-matching problems with other variables.

3.2 Crude Oil Futures Price Data

We collected daily futures price data from the Energy Information Administration (EIA) website. From 6514 daily observations which are from April 04, 1983 to March 24, 2009, we extracted weekly data by the following means. First we looked for Tuesday observations. If Tuesday observation does not exist for a specific week, we use the Monday just before that Tuesday. If Monday does not exist either, we use Wednesday, and then Thursday and Friday. Finally we obtained 1356 weekly observations from April 04, 1983 to March 24, 2009. Among these observations are 1341 Tuesday observations, 6 Monday observations and 9 Wednesday observations. These 1356 observations (which have a longer time period than the COT data) are used to estimate conditional variances of futures prices.

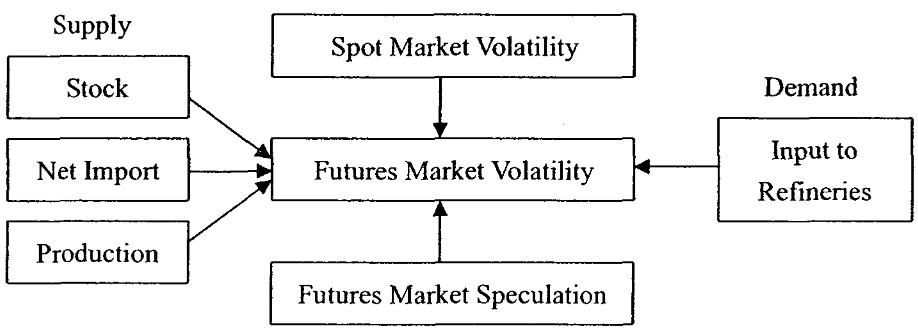
3.3 Crude Oil Spot Price Data

We collected daily spot prices of crude oil from the EIA website. From 5959 daily observations which extend from January 02, 1986 to March 24, 2009, we extracted weekly data by the same method as described above for futures prices. We obtained 1213 weekly observations. Among these observations are 1205 Tuesday observations, 7 Monday observations and 1 Thursday observation.

3.4 Other Control Variables

In order to find the net explanatory power of speculative indices for futures market volatility, we need to use control variables. Besides futures market speculation and spot market volatility, the fundamental factors such as supply and demand are also likely to influence futures market volatility. Input to refineries can be a proxy for demand, and stock (inventory), domestic production, and import and export can be proxies for supply.

Chart 3 Independent and Control Variables of Futures Market Volatility



Considering both data availability and correlation between proxies, we use the following control variables.

- ✓ Proxies for stock: Weekly U.S. Crude Oil Ending Stocks Excluding Strategic Petroleum Reserve (Thousand Barrels) and Weekly U.S. Crude Oil Ending Stocks Strategic Petroleum Reserve (Thousand Barrels).
- ✓ Proxy for import and export: Weekly U.S. Total Crude Oil and Petroleum Products Net Imports (Thousand Barrels per Day).

- ✓ Proxy for production: Weekly U.S. Crude Oil Field Production (Thousand Barrels per Day).
- ✓ Proxy for demand: Weekly U.S. Crude Oil Inputs into Refineries (Thousand Barrels per Day).

We downloaded the data for these control variables from the EIA website, and obtained 1383 weekly observations from August 20, 1982 to March 20, 2009. But since the control variables are only used with speculative indices, we only need 710 weekly observations for the same time period as COT data. The default as of date is Friday. Because we could not find daily data for these variables, Friday data are what we can obtain. The summary statistics for all the raw data are shown in Table 1.

Table 1 includes three sections: Crude Oil Futures Market Positions, Crude Oil Futures Price and Spot Price, and Other Control Variables. The data series were carefully created and checked by the respective administrations due to the extreme importance of crude oil market and data's widely use in various researches. We also examined all the identity relationships between market positions, such as the relationship that all reportable long positions plus all non-reportable long positions should be equal to the total open interest, and did not find noticeable problems.

Some other variables such as cuts to output by the OPEC are also recognized as influential factors in crude oil markets by previous research. We do not include these variables for two reasons. First is the data unavailability. We could not find the weekly data. Some of these variables have monthly series, others have only annual data. Second, these variables are more likely to influence the long-term equilibrium rather than short-term price dynamics in the crude oil futures market.

Chapter 4 Methodology

4.1 Calculation of Speculative Indices

We followed the definition of Working' T and the new speculative index as follows:

when HS ≥ HL: New Speculative index $= \frac{SS}{HS}$

$$\text{Working's } T = 1 + \frac{SS}{HS + HL}$$

when HS < HL: New Speculative index $= \frac{SL}{HL}$

$$\text{Working's } T = 1 + \frac{SL}{HS + HL}$$

In addition, among 710 COT observations, 572 observations have HS greater than or equal to HL, 138 observations have HS less than HL.

As of the treatment of non-reporting positions, we follow Sanders, Irwin and Merrin (2008), and allocate the non-reporting traders' positions to the commercial, non-commercial, and index trader categories in the same proportion as observed for reporting traders.

In order to estimate the potential influence of this procedure on the reliability of the derived HS, HL, SS and SL, we present the proportion of non-reportable positions (long and short) to the total open interest in Figure 2. From 1995 to 1997, the average proportion (for non-reportable long or non-reportable short) was about 20%. From 1998 to 1999, however, the proportion decreased dramatically to less than 10%. After that, the proportion decreased gradually to 2%~3% in 2008 and 2009. We do not know the actual composition of non-reportable traders, but considering that for the major part of our data the proportion is less than 10%, the potential bias from our allocation procedure is likely to be limited. We also expect the CFTC to make a complete market survey at a future date, so we can actually test the accuracy of this procedure.

4.2 GARCH Model to Estimate Conditional Variance of Futures Price

In order to obtain a stationary time series, we calculate the logarithm of the futures price, and then difference it. We use the GARCH (1,1) model proposed by Pindyck (2004b) as follows:

Mean equation: $Log\_price\_change = mean + \varepsilon_t$

Variance equation: $h_t = c + a * \varepsilon_{t-1}{}^2 + b * h_{t-1}$

h_t is the estimated conditional variance series, ε_t is the error term. The mean, a, b and c are the parameters.

Besides conditional variance estimates, we also use the squared futures return, as a measure of futures market volatility, in robustness tests. The summary statistics for speculative indices and volatility measures are shown in Table 2. Figure 3 and Figure 4 present the two speculative indices and conditional variance estimates along with the time axis.

4.3 Linear Regression Models

In order to identify the explanatory power of the two speculative indices to futures market volatility, we construct the following four models. Model 1 and Model 2 use conditional variance estimates as the volatility measures for the futures market, and regress the volatility on the new speculative index (Model 1) or Working's T (Model 2), and on other control variables. Model 3 and Model 4 use the squared futures return as the volatility measure for the futures market, and regress it on the new speculative index (Model 3) or Working' T (Model 4) and on other control variables.

Model 1:

$$h_t = \beta_0 + \beta_1 * t_{new} + \beta_2 * S\_ret\_squared + \beta_3 * Stock\_no\_SPR + \beta_4 * Stock\_for\_SPR$$
$$+\beta_5 * \Pr oduction + \beta_6 * Net\_\operatorname{Im} port + \beta_7 * Input + \varepsilon_t$$

Model 2:

$$h_t = \beta_0 + \beta_1 * t_{old} + \beta_2 * S\_ret\_squared + \beta_3 * Stock\_no\_SPR + \beta_4 * Stock\_for\_SPR$$
$$+\beta_5 * \Pr oduction + \beta_6 * Net\_\operatorname{Im} port + \beta_7 * Input + \varepsilon_t$$

Model 3:

$$F\_ret\_squared = \beta_0 + \beta_1 * t_{new} + \beta_2 * S\_ret\_squared + \beta_3 * Stock\_no\_SPR$$
$$+\beta_4 * Stock\_for\_SPR + \beta_5 * \Pr oduction + \beta_6 * Net\_\operatorname{Im} port + \beta_7 * Input + \varepsilon_t$$

Model 4:

$$F\_ret\_squared = \beta_0 + \beta_1 * t_{old} + \beta_2 * S\_ret\_squared + \beta_3 * Stock\_no\_SPR$$
$$+\beta_4 * Stock\_for\_SPR + \beta_5 * \Pr oduction + \beta_6 * Net\_\operatorname{Im} port + \beta_7 * Input + \varepsilon_t$$

h_t is the estimated conditional variance for the crude oil futures market, and $F\_ret\_squared$ is the squared return for the futures market. t_{new} refers to the new speculative index, and t_{old} refers to Working's T. $S\_ret\_squared$ is the squared return for the crude oil spot market. *Stock_no_SPR* refers to the U.S. Crude Oil Ending Stocks

Excluding Strategic Petroleum Reserve, and *Stock_for_SPR* refers to U.S. Crude Oil Ending Stocks Strategic Petroleum Reserve. *Production* refers to U.S. Crude Oil Field Production. *Net_Import* refers to U.S. Total Crude Oil and Petroleum Products Net Imports. *Input* refers to U.S. Crude Oil Inputs into Refineries.

The CFMA was passed by Congress and signed into law by President Bill Clinton in December 2000, just before Christmas recess. This act is often claimed to have created the Enron Loophole, which opened the door for big banks to help index traders to take huge speculative positions. Banks subsequently hedged their positions with index traders in the futures markets. The positions of these banks were treated as hedging positions in the COT Reports. We want to examine whether the passage of the CFMA changed the pattern of the volatility-speculation relationship in the futures market. So we use December 25th, 2000 as the cut-off time point, split the data, and run the above four regression models for each part of the data.

4.4 Granger Causality Test

In order to test whether the change in futures market volatility is caused by the change in speculative indices, we run the Granger causality test. The basic model is as follows:

Test of whether x Granger causes y is an F-test of whether $b_1 = b_2 = \ldots = b_L = 0$ in the regression:

$$y_t = a_0 + \sum_{i=1}^{L} a_i y_{t-i} + \sum_{i=1}^{L} b_i x_{t-i}$$

The failure to reject the null hypothesis that $b_1 = b_2 = ... = b_L = 0$ means lack of Granger causality from x to y, i.e., x does not Granger cause y.

We actually run two models:

Model 5: $Futures\_Volatility_t = a_0 + \sum_{i=1}^{L} a_i Futures\_Volatility_{t-i} + \sum_{i=1}^{L} b_i Speculative\_Index_{t-i}$

Model 5 is used to test whether the change in speculative index Granger causes the change in volatility of the crude oil futures market.

Model 6: $Speculative\_Index_t = a_0 + \sum_{i=1}^{L} a_i Speculative\_Index_{t-i} + \sum_{i=1}^{L} b_i Futures\_Volatility_{t-i}$

Model 6 is used to check whether the change in volatility of the crude oil futures market Granger causes the change in the speculative index.

4.5 Treatment of Multicollinearity

When we conduct linear regression models as discussed in section 4.3, we find high variance inflation factors, which is a sign of multicollinearity. The correlation matrix for all dependent and independent variables supports this judgment. Because the treatment of multicollinearity is a very important part of model selection, we include a brief discussion of multicollinearity in this methodology section, rather in the empirical results section.

Multicollinearity occurs when independent variables in the regression model are highly correlated with each other. Multicollinearity results in spuriously insignificant predictors (which should be significant, but are insignificant due to high standard errors). On the other hand, multicollinearity indicates that some independent variables share the same information.

Though multicollinearity poses a threat to the reliability of coefficient estimation of our regression, it provides an opportunity to try different set of independent variables, drop redundant independent variables, thus resulting in parsimonious, intuitive and easy-to-explain models.

As mentioned above, multicollinearity can be detected by variance inflation factors and a correlation matrix. Generally speaking, a good candidate predictor to keep is the independent variable that is highly correlated to the dependent variable and less correlated with other predictors. In our context, we want to test the significance and explanatory power of speculative indices. So the control variables that are less correlated with the speculative index are preferred as well.

According to the correlation matrix in Table 3, *PRODUCTION* is highly correlated to almost all predictors, so it is the first candidate to drop. *STOCK_FOR_SPR* and *NET_IMPORT* are also likely to be dropped.

Since the correlation matrix only shows pairwise correlation, there may also exist correlation among more than two predictors. The variance inflation factors (VIF) can detect pairwise and "multiple" correlation. The rule of thumb is that VIFs greater than 5 are worthy of notice. Some people argue that VIFs greater than 10 are noticeable. Other people claim that for weaker models, VIFs greater than 2.5 need treatment. For each regression (combination of model and data set), we try dropping independent variables with high VIFs to see whether we get a more meaningful model. The final models are shown with regression results in Table 4.

Chapter 5 Empirical Results and Discussion

5.1 Regression of Volatility Measures on Speculative Indices

According to Model 1, we regress conditional variance estimates of futures prices on the new speculative index and other control variables. The coefficient of the speculative index is statistically significant, and the t-value is 6.41. Other control variables, except for inputs into refineries, all have significant coefficients. The F-value of 23.48 supports the overall significance of the set of independent variables. The explanatory power of the whole regression is 13.80%. In another word, 13.80% of variation in conditional variance estimates can be explained by the variation of the new speculative index.

According to Model 2, we regress conditional variance estimates of futures prices on Working's T and other control variables. The coefficient of Working's T is also significant, but the significance level is slightly lower than that for the new speculative index. Other control variables, except for input into refineries, are significant, and the significance levels are similar to Model 1. According to the adjusted R squared, 13.65% of variation in the conditional variance estimates can be explained by the variation of Working's T.

After finding the statistical significance of the new speculative index and Working's T in explaining the futures market volatility, we are interested in their economic significance. In Model 1, the coefficient of the new speculative index is 0.00469. So when the index increases from the first quartile (0.216643) to the third quartile (0.506610), the conditional variance increases by (0.506610-0.216643)*0.00469=0.00136. If we ignore the curvature of the square root function, the conditional standard deviation increases by about 0.0369.

Compared with such a big hypothetical change in the index, the increase in volatility seems relatively small.

For Working's T, we have similar results. When Working's T increases from the first quartile (1.103228) to the third quartile (1.260627), the conditional variance increases by (1.260627-1.103228)* 0.00907=0.001428. The approximate increase in standard deviation is 0.03778.

Model 3 and Model 4 are the robustness tests for Model 1 and Model 2. We regress squared futures returns on speculative indices and control variables. Squared spot returns become the most significant explanatory variables in these two equations. The new speculative index and Working's T become marginally significant. Because the squared spot return explains most of the variation in the dependent variable, the explanatory power and significance level of the new speculative index and Working's T are quite low. However, the significance level of the new speculative index is still a little higher than Working's T. Due to the very high explanatory power of the squared spot return, the overall R squareds of model 3 and model 4 are 50.97% and 50.94% respectively.

5.2 Regression of Volatility Measures on Speculative Indices on Split Data

According to Model 1 and Model 2, we regress estimated conditional variance on speculative indices and control variables on split data, and the results are shown in Table 5. After comparing the coefficients and corresponding significance levels, we can make some inferences. After CFMA was enacted, the new speculative index changes from an insignificant factor to a significant factor. In another word, the speculation in crude oil futures market began to have explanatory power. Squared spot returns became more significant, which may imply that the futures market was more integrated with the spot market. Ending Stock excluding SPR becomes less significant, but Ending Stock for SPR becomes significant. Input to refineries becomes less significant, and its coefficient changes in sign. The overall explanatory power decreases from about 20% to 15%.

According to Model 3 and Model 4, we regress squared futures return on speculative indices and control variables on split data. We found in both parts of the data, the speculative index (new speculative index or Working's T) is not statistically significant. It seems strange that we found significance of the speculative index on all data but not on the spilt data. One of the possible explanations comes from the statistical property of OLS regression. On either part of the data, the observations are clustered, like clustered points on a scatter plot. It is difficult to find the fitted linear model for each part. But if we combine the split data (or use the full data), there may be two clustered regions on the scatter plot, and easy to draw the fitted line. This explanation may infer that the observations did change in some respects. Of course, this is just one of many possible explanations, and subject to validation. Before the validation, it is difficult to make further inferences. Though the speculative index is not significant before and after CFMA, its significance level becomes higher after CFMA was enacted.

Squared spot returns become less significant after CFMA was enacted, which looks contrary to the results based on estimated conditional variance. However, considering that the overall explanatory power of models decreases from 82% to 46% and squared spot returns is the only significant predictor on the post-CFMA data, the relative significance of squared spot return in the model does not decrease. In addition, after CFMA was enacted, the Ending Stock without SPR becomes insignificant.

Comparing all results on the split data, we can identify some systematic patterns. After CFMA was enacted, the overall explanatory power of the models decreases, but the absolute and relative significance of the speculative index increase. We can infer that in the post-CFMA period, speculation plays a more important role in explaining futures market volatility, and other fundamental factors' explanatory power decrease dramatically. We have to emphasize that we get this finding when index traders' positions are categorized as hedges. If we can get separate data for index traders and classify them as speculators, we may find that speculation plays an even more significant role in explaining futures market volatility dynamics.

5.3 Granger Causality Test Results

According to Table 6, it is difficult to identify the Granger Causality relationship for one week lag. Nevertheless, for two and three week lags, we can find that the change in the speculative index leads to change in the conditional variance, and the causality does not exist for the opposite direction. In addition, the significance level of the causality relationship is low.

Chapter 6 Conclusion

In this thesis, we use the new speculative index developed by Dr. Latha Shanker and Working's T to measure the degree of speculation in the crude oil futures market. We estimate the conditional variance of crude oil futures prices, and also use squared futures return to measure the volatility of the futures market. We find the variation of futures price volatility can be explained by the variation of speculation as measured by speculative indices, but the economic significance of this relationship is quite low.

We also find that, after the Commodity Futures Modernization Act was enacted, speculation has become a more important factor in explaining the futures market volatility, and that the explanatory power of other fundamental variables decreases dramatically.

Through Granger causality tests, we find the change in speculation does lead to change in volatility in futures market for the second and third week lag. However, the cause-effect relationship is weak. In addition, such a causality relationship does not exist in the opposite direction.

References:

Antoniou, A., and A.J. Foster, 1992, The Effect of Futures Trading on Spot Price Volatility: Evidence for Brent Crude Oil Using GARCH, Journal of Business Finance & Accounting, 19, 473-484.

Board, J., G. Sandmann, and C. Sutcliffe, 2001, The Effect of Futures Market Volume on Spot Market Volatility, Journal of Business Finance & Accounting, 28, 799-819.

Bryant, H.L., D.A. Bessler, and M.S. Haigh, 2006, Causality in Futures Markets, Journal of Futures Markets, 26, 1039-1057.

Buchananan, W.K., P. Hodges, and J. Theis, 2001, Energy Economics, 23, 279-293.

Chicago Board of Trade, 1989, Commodity Trading Manual.

Christian, J.M., 2006, Commodities Rising, John Wiley & Sons, Hoboken, NJ.

Committee on Homeland Security and Government Affairs (CHSGA), U.S. Senate, 2008, "Lieberman, Collins Say Commodities Market Speculation Contributes to High Cost of Food, Oil", May 20.

Commodity Futures Trading Commission (CFTC), 2006a, Comprehensive Review of the Commitments of Traders Reporting Program, Federal Register, 71, FR 35627, June 21.

Commodity Futures Trading Commission (CFTC), 2006b, Commodity Futures Trading Commission Actions in Response to the "Comprehensive Review of the Commitments of Traders Reporting Program", December 5.

Commodity Futures Trading Commission (CFTC), September, 2008, Staff Report on Commodity Swap Dealers & Index Traders with Commission Recommendations.

Deaves, R., and I. Krinsky, 1992, The Behavior of Oil Futures Returns around OPEC Conferences, Journal of Futures Markets, 12, 563-574.

Epstein, G., 2008, Commodities: Who's Behind the Boom?, Barron's, March 31.

Enders, W., 2004, Applied Econometric Time Series (2nd edition), John Wiley & Sons, Hoboken, NJ.

Ferderer, J.P., 1996, Oil Price Volatility and the Macroeconomy, Journal of Macroeconomics, 18, 1-26.

Foster, A.J., 1995, Volume-Volatility Relationships for Crude Oil Futures Markets, Journal of Futures Markets, 15, 929-951.

Goodwin, B.K., and R. Schnepf, 2000, Determinants of Endogenous Price Risk in Corn and Wheat Futures Markets, Journal of Futures Markets, 20, 753-744.

Gorton, G., and K.G. Rouwenhorst, 2006, Facts and Fantasies about Commodity Futures, Financial Analysts Journal, 62, 47-68.

Holihan, M.B., 2008, The Complete Guide to Investing in Commodity Trading and Futures, Atlantic Publishing Group, Ocala, Florida.

Krichene N., 2002, World crude oil and natural gas: a demand and supply model, Energy Economics, 24, 557-576.

Leuthold, R.M., 1983, Commercial Use and Speculative Measures of the Livestock Commodity Futures Markets, Journal of Futures Markets, 3, 113-135.

Masters, M.W., 2008a, Testimony before the Committee on Homeland Security and Government Affairs, U.S. Senate, May 20.

Masters, M.W., 2008b, Testimony before the Committee on Homeland Security and Government Affairs, U.S. Senate, June 24.

Nixon, J.W., 2008, RE: Commodity Futures Trading Commission (CFTC) - Action on Convergence in Agricultural Commodities Futures, April 22.

Peck, A.E., 1979-1980, Reflections of Hedging on Futures Markets, Food Research Institute Studies, 17, 327-349.

Peck, A.E., 1981, The Adequacy of Speculation on the Wheat, Corn, and Soybean Futures Markets, Research in Domestic and International Agribusiness Management, 2, 17-29.

Peck, A.E., 1982, Estimation of Hedging and Speculative Positions in Futures Markets Revisited, Food Research Institute Studies, 18, 181-195.

Pindyck, R.S., 2004a, Volatility in Natural Gas and Oil Markets, Journal of Energy and Development, 30, 1-19.

Pindyck, R.S., 2004b, Volatility and Commodity Price Dynamics, Journal of Futures Markets, 24, 1029-1047.

Rogers, Jim, 2007, Hot Commodities: How Anyone Can Invest Profitable in the World's Best Market, Random House.

Rutledge, D.J.S., 1977-78, Estimation of Hedging and Speculative Positions in Futures Markets: An Alternative Approach, Food Research Institute Studies, 16, 205-211.

Sanders, D.R., K. Boris, and M. Manfredo, 2004, Hedgers, Funds, and Small Speculators in the Energy Futures Markets: An Analysis of the CFTC's Commitments of Traders Reports, Energy Economics, 26, 425-445.

Sanders, D.R., S.H. Irwin, and R.P. Merrin, 2007, Smart Money? The Forecasting Ability of CFTC Large Traders, Proceedings of the NCCC-134 Conference on Applied Commodity Price Analysis, Forecasting, and Market Risk Management.

Sanders, D.R., S.H. Irwin, and R.P. Merrin, 2008, The Adequacy of Speculation in Agricultural Futures Markets: Too Much of a Good Thing?, Marketing and Outlook Research Report 2008-02, Department of Agricultural and Consumer Economics, University of Illinois at Urbana-Champaign.

Streeter, D.H., and W.G. Tomek. 1992, Variability in Soybean Futures Prices: An Integrated Framework, Journal of Futures Markets, 12, 705-728.

Till, H., and J. Eagleeye, 2007, Intelligent Commodity Investing: New Strategies and Practical Insights for Informed Decision Making, Risk Books, Incisive Financial Publishing.

Working, H., 1960, Speculation on Hedging Markets, Food Research Institute Studies, 1, 185-220.

Working, H., 1962, New Concepts Concerning Futures Markets and Prices, American Economic Review, 62, 432-459.

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

Appendices:

Table 1 Summary Statistics for Raw Data

Variable	Description	# of Usable	# of Missing	Mean	Median	Std Dev	Skewness	Excess Kurtosis	Minimum	Maximum
Crude Oil Futures Market Positions										
TOI	Total Open Interest	710	681[1]	1134489	779250	770578.3	1.293252	0.362182	391567	3329695
NCL (SL)	Non-Commercial Long	710	681	94346.91	64585.5	77921.88	0.883143	-0.432	9552	302206
NCS (SS)	Non-Commercial Short	710	681	52248.53	35056	44900.43	0.94791	-0.43727	1716	169941
NCSP	Non-commercial Spread	710	681	299979	141782	346303.5	1.6478	1.686943	21106	1471270
CL (HL)	Commercial Long	710	681	665739.1	512875	361965.1	1.136301	0.009064	230673	1607317
CS (HS)	Commercial Short	710	681	706489.7	553689	389043.7	1.103348	-0.03587	254878	1684802
RL	Total Reportable Long	710	681	1060065	713620.5	769010.3	1.277148	0.346011	313771	3271788
RS	Total Reportable Short	710	681	1058717	721740	765072.6	1.281013	0.35904	309792	3268723
NRL	Total Non-Reportable Long	710	681	74423.84	74370.5	18950.33	0.087264	-0.59904	32548	129900
NRS	Total Non-Reportable Short	710	681	75771.59	77953	19353.32	0.116745	-0.83001	38181	124958
Crude Oil Futures Price and Spot Price										
FUTURES_PRICE	Crude Oil Futures Price	1356	35	31.00611	22.74	21.27826	2.375516	6.204273	11.07	140.97
SPOT_PRICE	Crude Oil Spot Price	1209	182	31.21042	21.51	22.46458	2.229689	5.208464	10.88	141.06
Other Control Variables										
STOCK_NO_SPR	Ending Stock Excluding SPR	1382	9	325378.5	328243	23326.01	-0.25856	-0.30763	263666	391907
STOCK_FOR_SPR	Ending Stock for SPR	1382	9	570262.2	571951	86239.5	-0.98593	1.836564	270455	709343
NET_IMPORT	Net Import	946	445	9747.94	9816	2037.04	-0.13706	-0.90478	4678	14370
INPUT	Inputs to Refineries	1382	9	14005.02	14126.5	1222.21	-0.34484	-0.7237	10639	16464
PRODUCTION	Production	1365	26	6726.11	6493	1252.02	0.310462	-0.98472	3813	8969

[1] Because we consider the longest time period among all variables (from August 20, 1982 to March 24, 2009) to be the full time period. The variables with shorter time periods appear to have many missing observations, but those observations are not really "missing".

Table 2 Summary Statistics for Speculative Indices and Volatility Measures

Variable	Description	# of Usable	# of Missing	Mean	Median	Std Dev	Skewness	Excess Kurtosis	Minimum	Maximum
Speculative Indices										
T_OLD	Working's T	710	681	1.19626	1.163946	0.11866	1.028669	0.118895	1.046321	1.532088
T_NEW	New Speculative Index	710	681	0.379061	0.309012	0.229816	1.033389	0.140568	0.085764	1.038327
Volatility Measures										
H_T	Conditional Variance Estimates of Futures Price	1355	0	0.002997	0.00203	0.003182	2.821017	9.278239	0.000233	0.023009
F_RET_SQUARED	Squared Futures Return	1355	0	0.002679	0.000776	0.006645	7.534156	76.43354	0	0.098323
S_RET_SQUARED	Squared Spot Return	1204	151	0.003101	0.001031	0.008186	9.443783	124.976	0	0.136944

Table 3 Correlation Matrix of the Variables Used in the Linear Regression Models[1]

	H_T	F_RET_ SQUARED	T_NEW	T_OLD	S_RET_ SQUARED	PRODUCTION	STOCK_ NO_SPR	STOCK_ FOR_SPR	NET_ IMPORT	INPUT
H_T	1.00000	0.31339	0.14695	0.14265	0.29759	-0.07665	0.01852	0.13796	0.12585	0.06380
F_RET_SQUARED	0.31339	1.00000	0.10512	0.10163	0.82505	-0.04833	0.00208	0.07471	0.10299	0.01801
T_NEW	0.14695	0.10512	1.00000	0.99948	0.14428	-0.84538	0.13678	0.85539	0.68951	0.11744
T_OLD	0.14265	0.10163	0.99948	1.00000	0.14278	-0.84532	0.13173	0.85958	0.69122	0.12058
S_RET_SQUARED	0.29759	0.82505	0.14428	0.14278	1.00000	-0.01548	0.03347	0.03310	0.10316	-0.04459
PRODUCTION	-0.07665	-0.04833	-0.84538	-0.84532	-0.01548	1.00000	0.53499	-0.82369	-0.89147	-0.83829
STOCK_NO_SPR	0.01852	0.00208	0.13678	0.13173	0.03347	0.53499	1.00000	-0.37777	-0.33072	-0.51911
STOCK_FOR_SPR	0.13796	0.07471	0.85539	0.85958	0.03310	-0.82369	-0.37777	1.00000	0.68494	0.71168
NET_IMPORT	0.12585	0.10299	0.68951	0.69122	0.10316	-0.89147	-0.33072	0.68494	1.00000	0.72913
INPUT	0.06380	0.01801	0.11744	0.12058	-0.04459	-0.83829	-0.51911	0.71168	0.72913	1.00000

[1] H_T refers to the conditional variance estimates of futures price, F_RET_SQUARED refers to the squared futures return, T_NEW refers to the new speculative index, and T_OLD refers to Working's T. S_RET_SQUARED refers to the squared spot return, PRODUCTION refers to the U.S. crude oil field production, STOCK_NO_SPR refers to U.S. crude oil ending stocks excluding Strategic Petroleum Reserve, and STOCK_FOR_SPR refers to U.S. crude oil ending stocks Strategic Petroleum Reserve. NET_IMPORT refers to U.S. total crude oil and petroleum products net imports, and INPUT refers to U.S. crude oil inputs into refineries.

Table 4 Regression of Volatility Measures of Crude Oil Futures Market On Working's T or New Speculative Index

Variables	Model 1	Model 2	Model 3	Model 4
Dependent	Conditional variance estimates of futures price	Conditional variance estimates of futures price	Squared Futures Return	Squared Futures Return
Intercept	0.00635 (2.39)**[1]	-0.00273 (-1.16)	0.00667 (1.44)	0.00208 (0.51)
New speculative index	0.00469 (6.41)***		0.00239 (1.87)*	
Working's T		0.00907 (6.31)***		0.00434 (1.73)*
Squared Spot Return	0.07838 (7.08)***	0.07861 (7.09)***	0.50684 (26.28)***	0.50729 (26.30)***
Stocks Excluding SPR[2]	1.226671E-8 (2.81)***	1.259326E-8 (2.88)***	2.592326E-9 (0.34)	2.755013E-9 (0.36)
Stocks for SPR	-1.71162E-8 (-5.88)***	-1.71787E-8 (-5.82)***	-1.08197E-8 (-2.13)**	-1.03599E-8 (-2.01)**
Inputs into Refineries	1.033459E-7 (0.79)	9.918433E-8 (0.76)	-2.74318E-8 (-0.12)	-2.96333E-8 (-0.13)
F value	23.48	23.19	146.98	146.78
Adj R-Sq	0.1380	0.1365	0.5097	0.5094

[1] The number in parentheses is the t value. *** refers to the 99% confidence level, ** refers to the 95% confidence level, and * refers to the 90% confidence level.

[2] SPR refers to Strategic Petroleum Reserve. It may seem strange that we include Stock_no_SPR and Stock_for_SPR in one model. But the correlation between these two variables is only -0.3777, they are not highly correlated, and thus are not different proxies for the same economic variable.

Table 5 Regression of Volatility Measures of Crude Oil Futures Market on Working's T or New Speculative Index on Split Data[1]

Variables	Before CFMA	After CFMA	Before CFMA	After CFA
Dependent	Conditional Variance	Conditional Variance	Conditional Variance	Conditional Variance
Intercept	-0.01657 (-1.73)*	0.01518 (4.25)***	-0.01759 (-1.17)	0.00940 (2.46)**
New Index	0.00465 (1.30)	0.00301 (2.95)***		
Working's T			0.00477 (0.66)	0.00583 (2.92)***
S_ret_squared	0.04953 (2.22)**	0.07770 (5.83)***	0.04999 (2.24)**	0.07775 (5.83)***
Stock_no_SPR	1.74574E-8 (2.94)***	1.300737E-8 (2.05)**	1.764101E-8 (2.96)***	1.340583E-8 (2.12)**
Stock_for_SPR	-2.54704E-9 (-0.18)	-1.35898E-8 (-3.71)***	-8.53602E-9 (-0.62)	-1.37087E-8 (-3.69)***
Input	9.86575E-7 (5.49)***	-5.87584E-7 (-2.77)***	9.884127E-7 (5.49)***	-5.94026E-7 (-2.81)***
F value	15.59	15.94	15.27	15.90
Adj R-Sq	0.2072	0.1504	0.2036	0.1500

Variables	Before CFMA	After CFMA	Before CFMA	After CFMA
Dependent	F_ret_squared	F_ret_squared	F_ret_squared	F_ret_squared
Intercept	-0.00134 (-0.11)	0.01310 (2.09)	-0.00302 (-0.16)	0.00828 (1.23)
New Index	0.00192 (0.44)	0.00267 (1.49)		
Working's T			0.00267 (0.30)	0.00475 (1.36)
S_ret_squared	0.98557 (36.06)***	0.43212 (18.49)***	0.98575 (36.06)***	0.43252 (18.49)***
Stock_no_SPR	1.56132E-8 (2.14)**	-3.65923E-9 (-0.33)	1.563786E-8 (2.14)**	-3.13013E-9 (-0.28)
Stock_for_SPR	-4.7731E-9 (-0.28)	-1.02402E-8 (-1.60)	-6.33945E-9 (-0.37)	-9.72939E-9 (-1.49)
Input	-7.40224E-8 (-0.34)	-3.41964E-7 (-0.92)	-7.38941E-8 (-0.34)	-3.62852E-7 (-0.98)
F value	265.18	75.55	265.06	75.41
Adj R-Sq	0.8256	0.4690	0.8255	0.4685

[1] The number in parentheses is the t value. *** refers to the 99% confidence level, ** refers to the 95% confidence level, and * refers to the 90% confidence level.

Table 6 Granger Causality Test: New Speculative Index, Working's T and Conditional Variance[1]

Lags	New index -> Conditional variance	Working' t -> Conditional variance	Conditional variance -> New index	Conditional variance -> Working' t
1	Yes*	Yes*	Yes*	Yes*
2	Yes**	Yes*	No	No
3	Yes*	Yes*	No	No
4	Yes*	No	Yes*	Yes*

[1] Test of x->y is an F-test of the exclusion restrictions that $b_1 = b_2 = ... = b_L = 0$ in the regression:

$$y_t = a_0 + \sum_{i=1}^{L} a_i y_{t-i} + \sum_{i=1}^{L} b_i x_{t-i}$$

[2] "No" means fails to reject the null hypothesis that $b_1 = b_2 = ... = b_L = 0$ at 90% confidence level, i.e., lack of Granger causality. "Yes" means rejection of null hypothesis at 90% level (*), 95% level (**), or 99% level (***).

Figure 1 Crude Oil Spot Price from January 1986 to March 2009

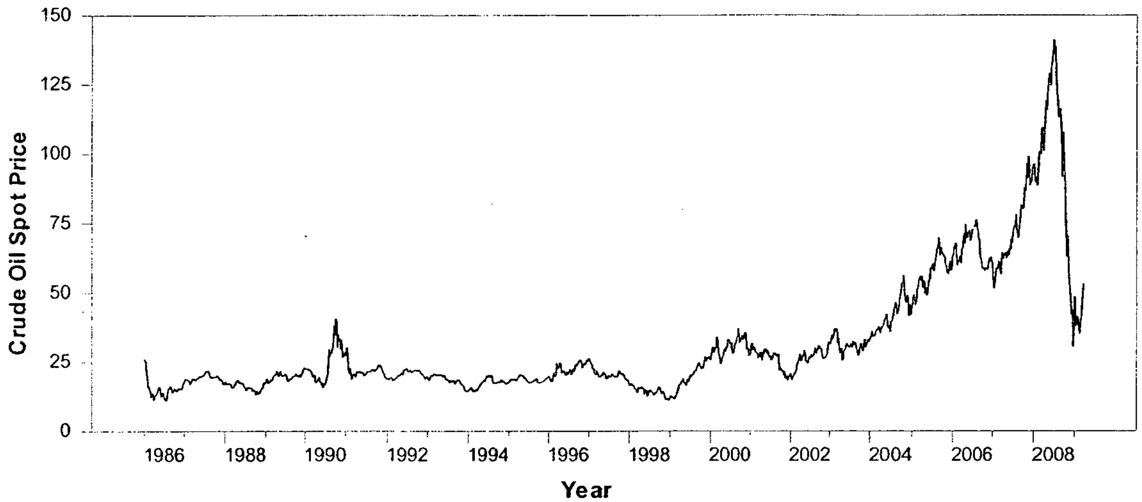
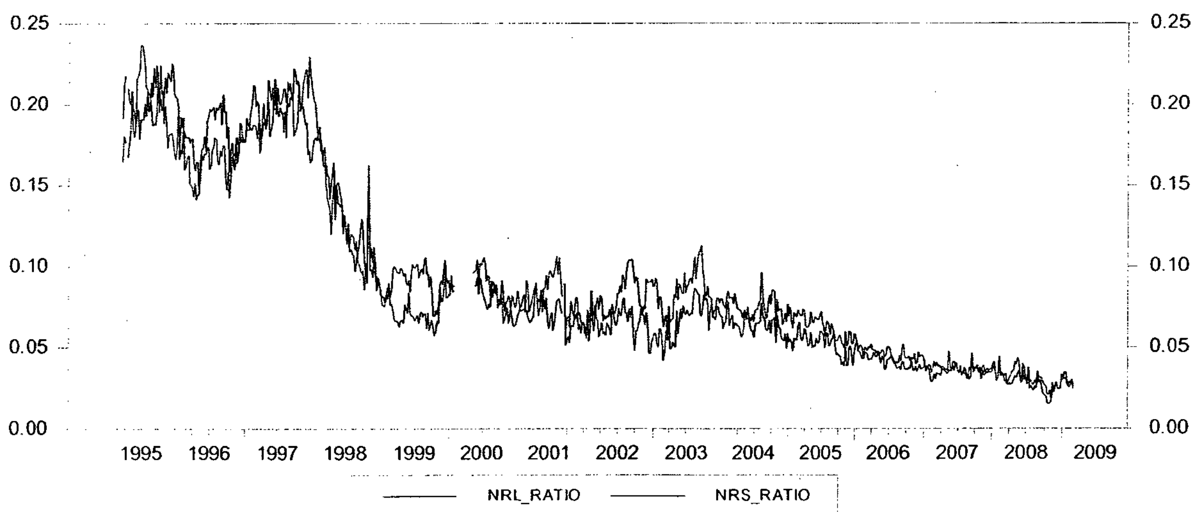
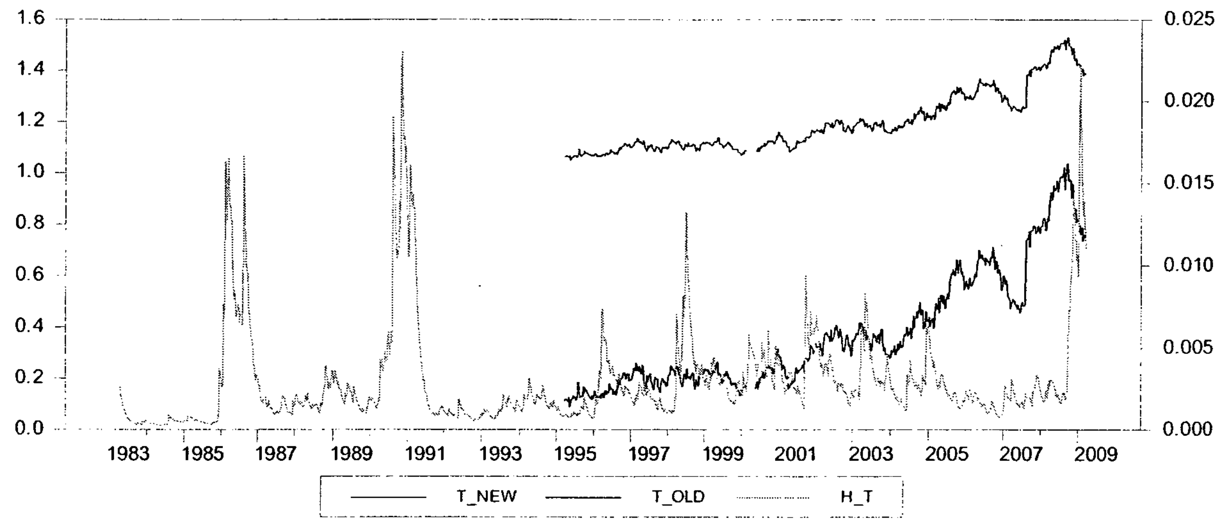


Figure 2 Proportion of Non-reportable Positions in Total Open Interest



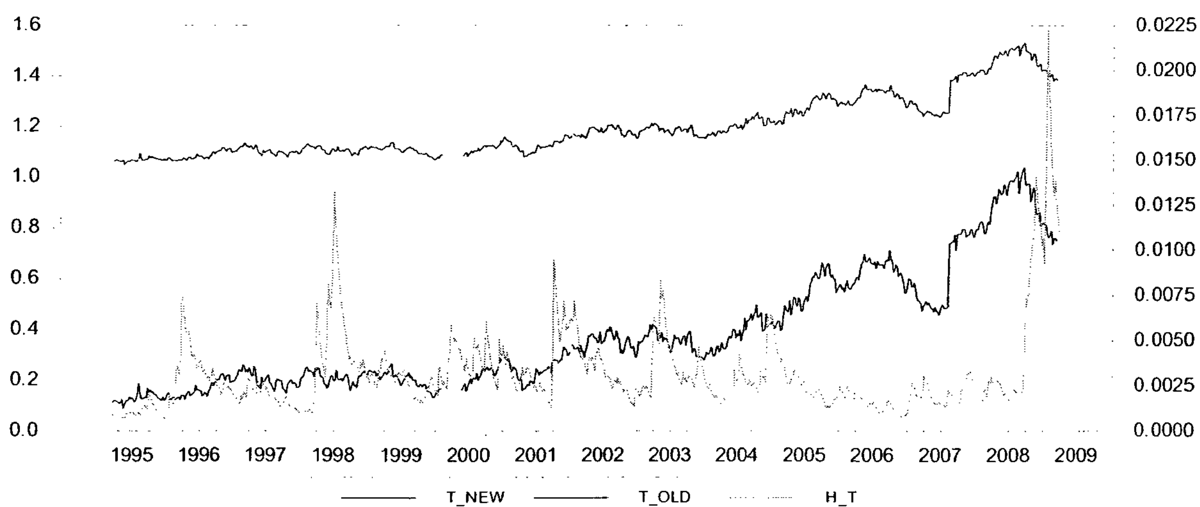
* NRL_Ratio refers to all non-reportable long positions divided by total open interest, and NRS_Ratio refers to all non-reportable short positions divided by total open interest.

Figure 3 Speculative Indices and Conditional Volatility



* T_new refers to the new speculative index, T_old refers to Working's T, and H_T refers to the conditional variance estimate from the GARCH model.

Figure 4 Speculative Indices and Conditional Volatility on Common Time Period



* T_new refers to the new speculative index, T_old refers to Working's T, and H_T refers to the conditional variance estimate from the GARCH model.