2016-01-27

Valuation of Crude Oil Futures, Options and Variance Swaps

Shahmoradi, Akbar

UNIVERSITY OF CALGARY

Valuation of Crude Oil Futures, Options, and Variance Swaps

by

Akbar Shahmoradi

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE

DEGREE OF DOCTOR OF PHILOSOPHY OF SCIENCE

GRADUATE PROGRAM IN MATHEMATICS AND STATISTICS

CALGARY, ALBERTA

JANUARY, 2016

Abstract

In this research we provide a set of practical approaches to value crude oil futures, especially long dated ones given crude oil spot prices. Throughout the research we change the reference point for our data sets from calendar dates to time to expiry and all our models are analyzed based on time to expiry.

We use a set of Levy processes to value crude oil options by calibrating parameters using Fast Fourier Transform algorithm and solving an objective function using Particle-Swap Optimization.

In order to help market participants to use available crude oil storage and refinery data in pricing futures contracts and the spreads between them, we provide a framework that helps crude oil market participants to get fair value of futures and run scenario analysis if a physical factor such as level of inventories at Cushing\Oklahoma or in the US changes.

We also investigated variance risk premia in crude oil prices using information obtained from crude oil option prices. Our results indicate that ***"usually"*** there is a negative risk premium in crude oil prices but that does not necessarily provide trading opportunity for market participants because excess return of shorting the variance swap show huge losses when crude oil market is in turmoil.

Dedication

I dedicate this research to my parents and my wife

Acknowledgements

I would like to sincerely express my gratitude to my dear supervisor Professor Anatoliy Swishchuk from Department of Mathematics and Statistics for his continuous support and supervision throughout the Ph.D. studies.

Besides my supervisor, I would like to sincerely thank the rest of my thesis committee: Professor Antony Ware, Professor Matt Davison, Dr.Alexandru Badescu, and Dr.Pablo Moran, for their insightful comments which helped me to significantly improve the quality of the research.

Table of Contents

List of Figures

List of Tables

Chapter One: Introduction

1.1 Short History of Crude Oil Markets

Because of the key role of oil in the global economy and its integration with financial markets around the world, managing risks associated with price of crude oil is very critical for governments, businesses, producers, refineries and other market participants.

Before 1973, market participants had little concern over managing risk of oil prices as the price of this commodity used to be stable and much more predictable. The Red Line Agreement (1927) and Achnacarry Agreement (1928) was main contributor to the price stability. The Red Line Agreement was intended to prevent members of a cartel consisting of the world's largest oil companies, known as Turkish Petroleum Company (TPC) back then, from exploration, and production of crude oil in the ex-Ottoman territory without partnering with all other members. However, the TPC would also control crude oil prices. Unlike the Red Line Agreement which was focus on upstream of crude oil markets, the Achnacarry Agreement which was made by major international oil companies at Achnacarry Castle in Scotland, was intended to control downstream marketing of oil. These two main agreements successfully brought price stablility to crude markets until 1972.

Since that time, political and economic crises such as the 1973 Arab-Israeli war combined with sharp decline in the US oil production, the 1979 Iranian revolution, the 1990 Gulf war, and the 2008 financial crisis have led to major global oil price movements in either direction. For example, as market had been tightening since decline of US oil production from 10 million barrels a day (MBD) to just over 8 MBD, the 1973 OPEC oil embargo skyrocketed crude oil prices from sub $3/barrel to over $12 in just a few months. Crude oil prices averaged around $20 since then and traded between $10 and $40 through 1990.

As Figure 1.1 shows, crude oil markets show sizable swings and high volatility at times of crises.

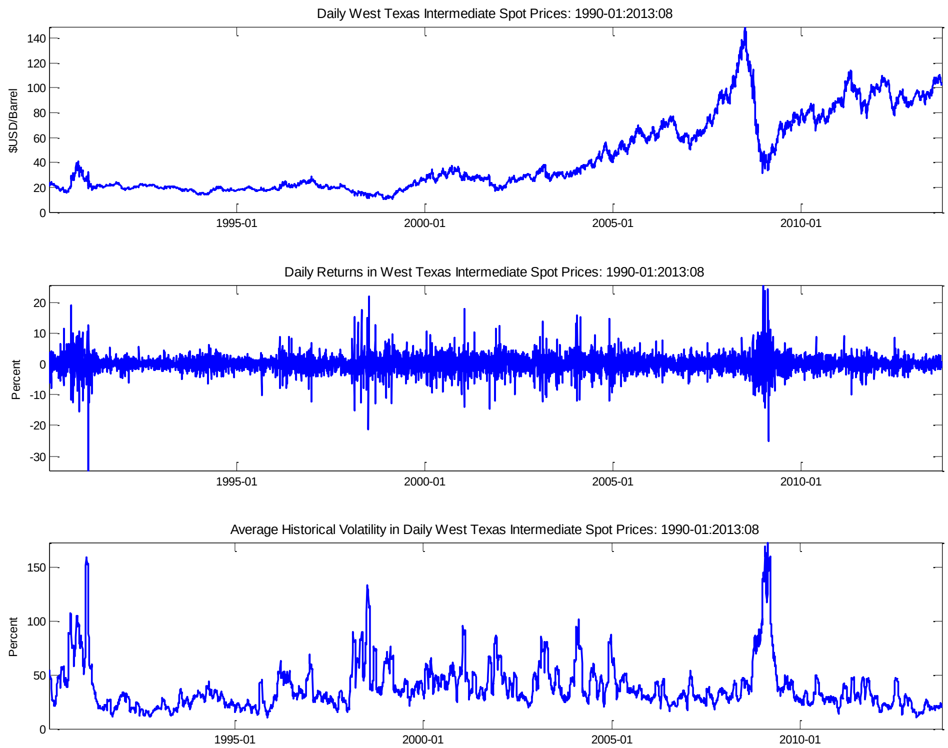


Figure 1.1: Return and Historical Volatility of WTI Crude Oil Prices

For example, during financial crisis crude oil volatility spiked to over 150 percent which was way over historical average of around 40 percent. This makes a relatively accurate approach to measuring risk very crucial to market participants as to when these drastic moves in risk are likely in crude oil market.

From transactional perspective, before 1983 transacting crude oil used to involve bilateral agreements between counterparties. Producers and consumers would use a pre-agreed prices to negotiate long term contracts. By initiating the first crude oil futures contracts on the New York Mercantile Exchange (NYMEX) in 1983, light crude oil futures has eventually become the world's most actively traded commodity. The liquidity of crude futures attracted physical and financial market participants into the crude markets. In general, participants were either trying concerned about managing a risk of being exposed to physical market, or were hoping to make speculative profit by being on the right side of the market.

By late 1990s as number of market participants in crude futures contracts, and consequently the market liquidity increased significantly, crude futures became the main space to determine crude oil prices and the physical traded volume in spot market declined substantially. In fact, producers, refineries and also storage operators take advantage of crude futures markets and manage their exposure well before being exposed to fluctuations in the spot markets. This is the reason modeling dynamics of spot crude oil price cannot be done by simply classifying key factors into supply and demand for *"physical"* crude oil. Instead, the dynamics of crude oil prices should be done in a framework that involves both spot and futures market. In addition, the behavior of players should be taken into account as well, which will be discussed in detail in the Chapter 5.

1.2 Key Trends in Crude Oil Market

Basically the dynamics and structure of global and regional crude oil markets could be discussed from four points of views.

- Supply components of crude oil markets
    - OPEC objective to influence crude market should prices decline significantly by adjusting its production targets if and when OPEC members believe their action could have long term impact on crude oil prices.

    - Non-OPEC objective to maximize shareholder value taking price as given

- Demand components of crude oil
    - Industrialized countries with non-increasing demand
    - Developed countries with significant growth in demand
    - Significant growth in financial side of oil and refined products
    - Hedging and risk-limiting activities of physical buyers and sellers of oil and its refined products
    - Activities of market makers, index-asset investors, banks, hedge funds and pension funds in the market
    - Impacts of managing and optimizing physical assets such as transportation and storage terminals, and refineries on market volatility and term structure of forward curves, and vice versa.

There are many types of crude oil produced around the world that are not necessarily the same in terms of quality. Light-weight, low sulfur grades tend to price higher than heavier, higher-sulfur grades. Location is also a factor as disruptions in a regional production could potentially impact oil prices in that region more than the others. Independent of the quality, supply of oil is very inelastic in the short term in a sense that producers are not able to increase oil production in a few months, if not years.

Crude oil is used in refineries to produce petroleum products such as gasoline, diesel, heating oil, and jet fuel. This is why prices of crude oil and petroleum products are closely tied to each other as both are impacted by factors affecting either market. However, a shock in one market does not necessarily impact the other market in the same direction. For example, unexpected refinery outage puts upward pressure on prices of petroleum product but because it reduces demand for crude oil, puts a downward pressure on crude oil prices. However, an unexpected supply disruption of crude oil production pushes both markets higher. *This will be investigated in Chapter 5*.

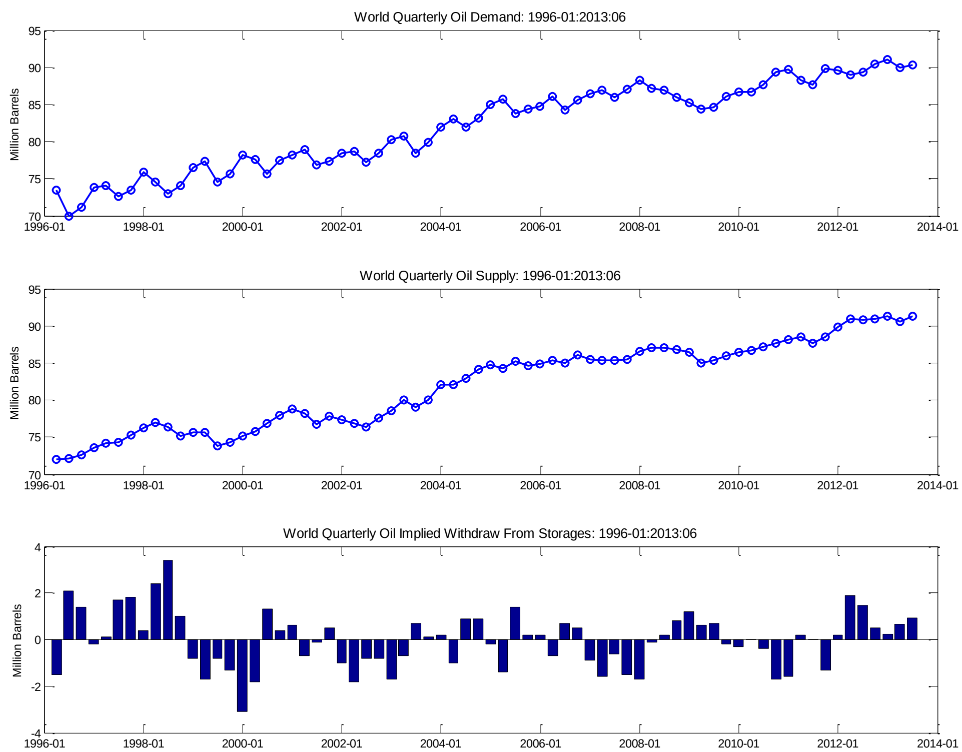


Figure 1.2: World Crude Oil Supply and Demand

Figure 1.2 shows oil supply and demand since mid-1990s. Currently word oil demand is around 90 million barrels a day (MBD), a sharp increase from 70 MBD in the late 90s. The third chart shows implied withdraw from storages, at times when demand outpaces supply. However, the rise in crude oil demand which resulted in increase in oil prices in the last decade has been almost entirely due to growth in developing countries such as China, Brazil, and India.

For example, as Figure 1.3 shows, demand for oil by members of Organization of Economic Cooperation and Development (OECD) countries was over 52% back in 2008 but been declining since then, while non-OECD consumption increased by over 10 MBD in the same period. Given non-OECD countries have almost no other source of energy to replace crude with, or perhaps increase fuel efficiency in the near term, it is expected that crude market to experience much more fluctuations and volatility because of demand stickiness. This is why any shock to the system could most likely result in significant price adjustment in order to balance physical crude oil markets.

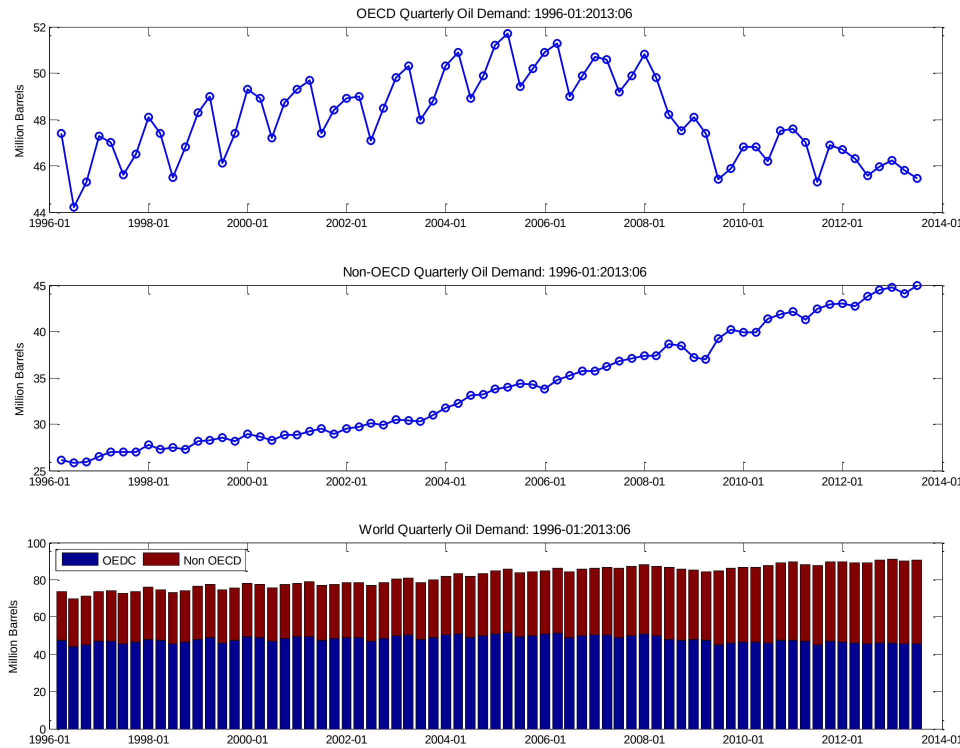


Figure 1.3: OECD vs Non-OECD Oil Demand

Organization of the Petroleum Exporting Countries (OPEC) is an important factor in affecting crude oil prices. OPEC member countries used to produce over 60% of the world's crude oil but recently their share has dropped to 40% (Figure 1.4). OPEC tends to actively influence global oil prices by setting production targets for each of its member countries. For example, OPEC took drastic actions twice in the last decade and reduced its production target by 4 MBD when crude prices dropped sharply in late 2002, and also during financial crisis in early 2009, which caused immediate rally in crude oil prices following the announcements (EIA). Even though share of OPEC members in the world oil production is declining but given these countries are mostly located in regions with many conflicts and also significant surprising social and political events, historically many of shocks to global oil prices have been caused by supply disruptions directly due to those events. The Arab Oil Embargo in 1973-74, the Iranian revolution, Iran-Iraq war in early 1980s, Persian Gulf War in early 1990s, and also political events in Nigeria, and Venezuela, Libya and Iraq are only few of those examples (EIA).

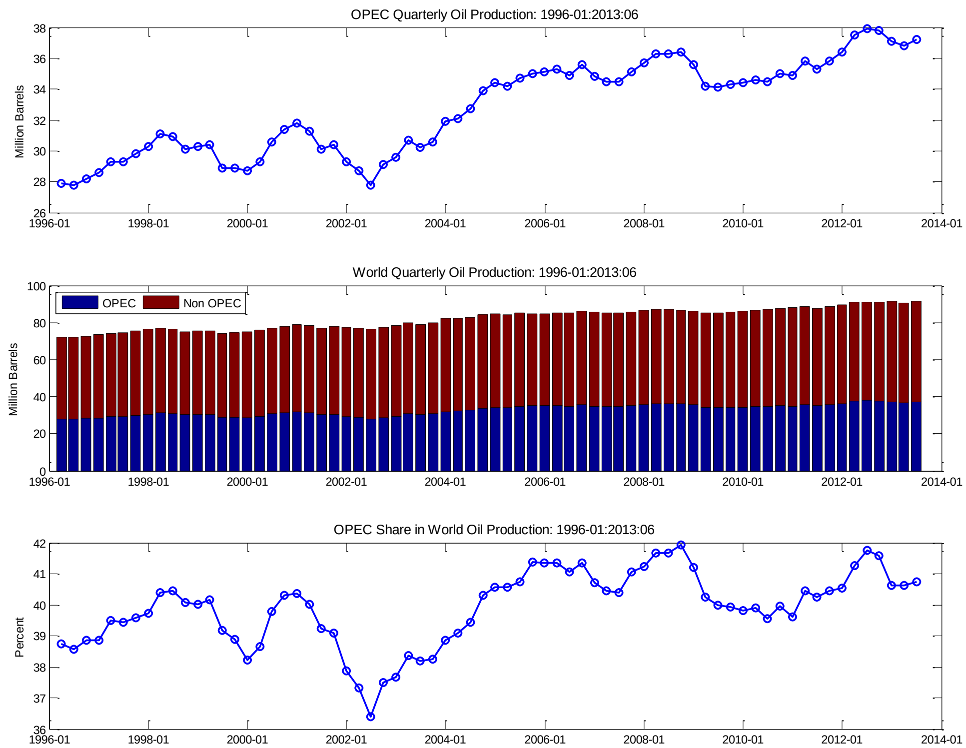


Figure 1.4: OPEC and non-OPEC Oil Production

Existence of oil futures markets has attracted many investors and other participants to the financial side of crude oil markets. This is very beneficial for both producers and consumers of oil as well, as they could hedge their production or input price risks by buying or selling energy derivatives. For example, an oil refinery may want to buy crude oil futures and sell petroleum products in order to lock in some spread between oil and refined products and hedge the risk of spread tightening. The refiner's counter-party on its crude leg could be a producer who wants to hedge its price risks, while an airline could be the one who wants to hedge its exposure to increasing fuel prices.

Other market participants such as market makers, banks, hedge funds, pension funds and other investors also get involved in the financial side of crude markets without necessarily having any intention of buying or selling physical quantities of oil. However, they play a key role in price discovery mechanisms in this market. The Figure 1.5 shows a measure known as open interest published by the **U.S. Commodity Futures Trading Commission (**CFTC) which indicates the number of contracts in the market that have not been settled or closed.

The open interests has been going up significantly since mid-90s and it was around 2.664 million on May 12, 2015 which was close to all-time high of 3.22 million made back in second half of 2008. Usually any large price movement is associated with a significant change in open interests. It is notable that investors who take position in index funds or exchange-traded funds (ETFs) investors are ultimately contributing to higher open interests in crude oil markets as well, and could cause significant market volatility if they decide to exit their positions.

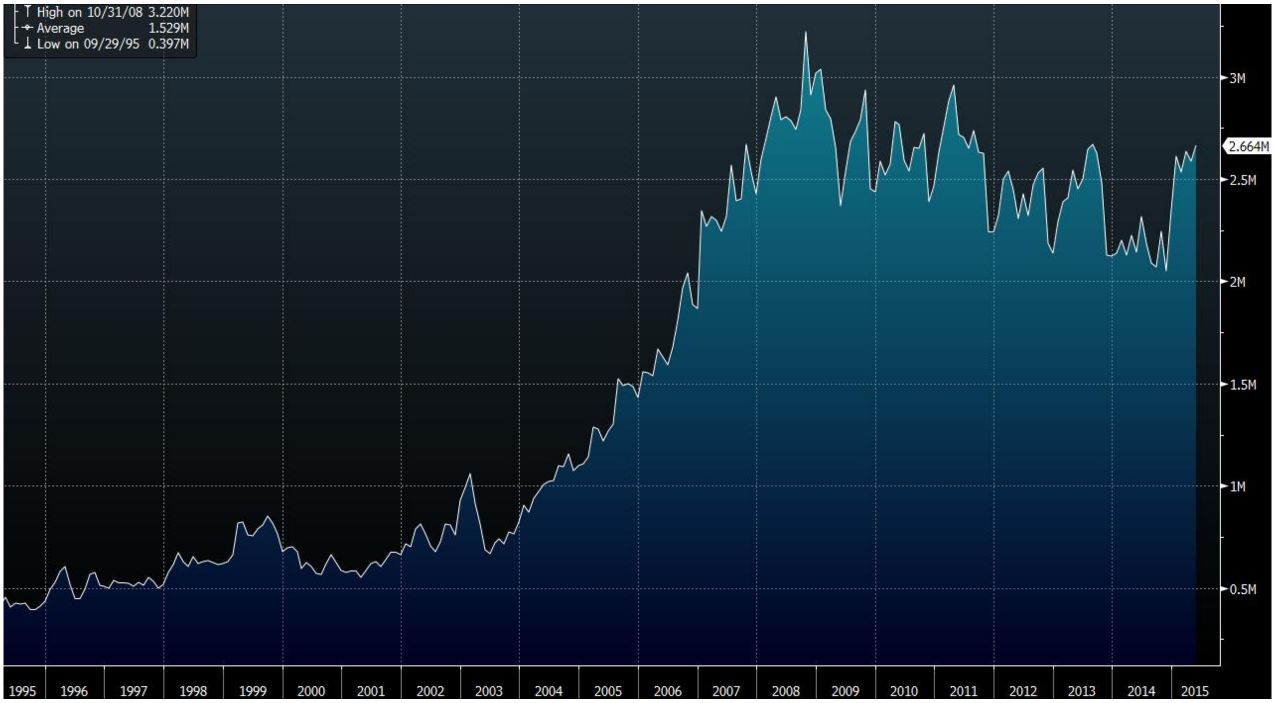


Figure 1.5: Open Interest in Crude Oil Futures Market (Source: Bloomberg)

The crude oil storage capacities and the level of oil inventories are two key factors affecting crude oil risks at times of crises. In addition, crude oil products such as gasoline, and heating oil, are two other factors affecting oil price dynamics and consequently its risk. Furthermore, refinery utilization and refinery margin are also factors market has an eye on while balancing crude oil markets. All these so called "physical factors" are main drivers behind most fluctuations in crude oil prices and play key role in forming the shape of the forward curve, spot price volatilities and also the risk associated with the level of prices.

1.3 Organization of the Thesis

In Chapter 2 we review the literature regarding quantitative analysis of crude oil prices. In general literature related to crude oil prices is partly focused on analyses of crude oil prices from statistical perspectives and their behavior in a time series framework over time. Some of the literature deals with developing risk metrics for crude oil prices, and finally there are some works on pricing instruments trading in the crude oil markets.

In Chapter 3 we focus on coming up with fair value of crude oil futures contracts in any given day after having spot prices. We use a set of one and two factor models to model crude oil prices. We carefully construct our data sets by changing reference of data from calendar date to time to expiry of each future contract. In this way we do not jump from one contract to the next one by ignoring how and when contracts role. We use daily spot and forward data and calibrate parameters using an optimization approach known as Particle-Swamp Optimization (PSO), and argue why my approach is superior to others in the literature. We provide a practical method which could help producers, market makers and other market participants to take spot prices traded in the market as given, and price long dated crude forward contracts on a real time basis.

In Chapter 4 we focus on pricing crude oil options using Merton's Jump Diffusion Model (MJDM), Normal Inverse Gaussian Model (NIGM), and Variance Gamma Model (VGM), which belong to family of Levy processes. We present characteristic functions of these processes, and then we use current market option prices to calibrate parameters of these models. We used Fast Fourier Transform (FRFT) algorithm to calibrate parameters of NIGM and VGM, and used PSO to calibrate MJDM. Our results are satisfactory of options not very far from ATM strikes.

We allocate Chapter 5 to building a bridge between risk-neutrality and structure of the crude oil markets. This is an improvement to what we have done in Chapter 3. We try to provide a framework for producers and other market participants to use information of crude oil inventories at Cushing and the US and feed them into futures pricing methods. We first present a set of structural models to show how the dynamics of crude market works. We use the framework to calibrate crude oil prices, and then we use the framework of our two-factor model presented in Chapter 3 to build a relationship between physical variables of crude oil markets and valuation methods. We provide out-of-sample results and compare them with actual data as well.

In Chapter 6 we investigate variance risk premia in crude oil prices using information obtained from crude oil option prices. We provide detailed steps on constructing the required data sets, and designing a realistic experiment. Our results indicate that there is a negative risk premium in crude oil prices but that does not necessarily provide trading opportunity for market participants because excess return of shorting the variance swap show huge losses when crude oil market is in turmoil.

We summarized our main findings, and contributions in Chapter 7.

1.4 Summary of Main Results

We used spot and futures data sets and imposed risk neutrality on our two factor models to calibrate our parameters. We believe this is the first time this is done for crude oil markets. Our models help to solve a problem that market participants have to deal with in a given day; what is the fair value of a future contract when spot market settles? Our models in Chapter 3 provide a practical solution for this question.

We used physical crude oil information and provided a practical solution to calculate fair value of crude oil future contracts by utilizing inventories and capacity data in the process. Our out of sample results are very encouraging.

We believe this is most likely the first application of NIG, JDM, and VG models in crude oil markets by calibrating their parameters using option markets. In crude oil market each option contract on a future has a different underlying compared to other option contract. This adds significant complexity to calibration and implementation process. We compared our results to actual data as well.

We also investigated Variance Risk Premia crude oil market, and unlike other researches in the literature our findings indicates that the negative excess return in Variance Swap does not necessarily provide trading opportunity because the strategy losses significantly at the times of crisis. We designed a very realistic and real-world type experiment by building forward-type curves on crude oil options for this reason.

Chapter Two: Review of Literature

In this chapter we review previous studies on modeling crude oil spot, futures, options and variance swap that are used to either measure risk or to value crude oil derivatives. Given our research uses models that are developed for the other asset classes, we review some of those works as well even though there are not applied to crude oil prices. The review of literature will be put into three categories in order to discuss their relevance to our work in this thesis. **The first category** is time series analysis of crude oil prices. These studies are not directly used in our research but they highlight one of the areas of crude oil market that researchers study them most. **The second category** is one and multi-factor models that are developed to study behavior of asset prices but our focus will be on commodities especially crude oil. **The third category** is allocated to variance swaps in crude oil markets.

2.1 Time Series Analysis of Crude Oil Prices

Kuper (2002) takes Generalized ARCH type volatility model to measure risk in spot Brent crude oil price from January 1982 through April 2002. He uses the "volatility" as a measure of risk and tests variety of model specifications to find a better fit. Kuper does not extend his model to come up with typical risk measures such as Value-at-Risk (VaR) for crude oil prices. Also, his study ignores futures contracts and is only focused on spot market.

Giot and Laurent (2003) propose a skewed Autoregressive Conditional Heteroskedasticity (ARCH) approach to measuring market risk in crude oil market using daily spot Brent and WTI prices for the period of 20/5/1987 to 18/3/2002. They use a typical Autoregressive Conditional Heteroskedasticity, ARCH, model defined as

$$r_t = R_t \Gamma + e_t \tag{2.1}$$

$$h_t^2 = a_0 + a_1 e_{t-1}^2 + ... + a_q e_{t-q}^2 \tag{2.2}$$

where $R_t = (1, r_{t-1}, ..., r_{t-q})$, and $\Gamma = (\rho_0, ..., \rho_p)'$. They use estimated models and forecast out-of-sample Value-at-Risk (VaR) metric for longer than 1-day but argue that given crude prices change course over long term due to supply-demand condition their model cannot be a good candidate for long term risk measures.

Cabedo *et al* (2003) uses an Autoregressive Moving Average (ARMA) historical simulation approach to model VaR in crude oil market using daily spot Brent oil prices. They use Historical Simulation Approach, Monte Carlo Simulation, and Variance-Covariance approaches as three common methodologies to calculate VaR with the assumption of standard normal distribution for oil price returns. They conclude that the historical simulation with ARMA forecasts (HSAF) generates most flexible and efficient risk metrics.

Fan et al (2008) estimate VaR of crude oil prices using variety of GARCH-type models, and also investigate existence of risk spill-over affect between Brent and WTI prices. Their GARCH model is similar to the set of equation were specified in (2.1) and (2.2). However, the (2.2) is replaced by the following equation:

$$h_t = a_0 + \sum_{i=1}^{p} a_i e_{t-i}^2 + \varphi e_{t-1}^2 d_{t-1} + \sum_{j=1}^{q} \beta_j h_{t-i} \tag{2.3}$$

The difference of their specification of h_t in (2.3) compared to typical GARCH models is the introduction of *d*, which returns 1 if $e_{t-1} < 0$ and 0 otherwise. This helps to add some asymmetric response to the model. They use daily spot data from May 1987 to August 2006. Their results show that their approach is superior to VaR calculated by other alternative normal-distribution based models, and also it seems there is significant spill over affect between Brent and WTI which should be expected of course.

Marimoutou *et al* (2006) use Extreme Value Theory to forecast VaR for both long and short crude oil positions. They investigate the relative predictive performance of some VaR and Extreme Value Theory (EVT) models, and compare them to GARCH, also historical simulation modeling approaches. They use Generalized Pareto Distribution (GPD) to estimate VaR. They fix a sufficiently high threshold for crude oil returns, *u*, and then for *q>F(u)* estimate VaR as follows:

$$VaR_t = \hat{u} + \frac{\hat{\beta}}{\hat{\xi}}\left[(\frac{n}{N_u}(1-q))^{-\hat{\xi}} - 1 \right] \tag{2.4}$$

where $F(x) = (1 - F(u))G_{\xi,\beta(u)}(y) + F(u)$, and

$$\hat{F}(x) = 1 - \frac{N_u}{n}\left[1 + \hat{\xi}(\frac{x - \hat{u}}{\hat{\beta}}) \right]^{-\frac{1}{\hat{\xi}}} \tag{2.5}$$

$$\hat{F}(n) = \frac{n - N_u}{n} \tag{2.6}$$

They use daily spot Brent oil prices from May 21, 1987 through January 24, 2006. By comparing the EVT models to other alternatives, they conclude that that their approach is superior to other alternatives.

Gorton et al (2007) take a macro approach to investigate relationship between commodity futures risk premiums and level of physical inventories. They use the following relationship between spot and forward prices and its relationship with storages:

$$F_{t,T} - S_t = S_t r_t + w_t - c_t \tag{2.7}$$

where *F* is the forward price, *S* the spot price, *r* interest rates, *w* marginal costs of storage, and *c* convenience yield. In order to estimate risk premium, $\pi_{t,T}$, they specify the following relationship based on the Theory of Normal Backwardation:

$$F_{t,T} - S_t = E_t(S_T) - S_t - \pi_{t,T} \tag{2.8}$$

The Normal Backwardation theory indicates that the expected future spot price at expiry of the contract, T, trades at a discount implying that $\pi_{t,T} > 0$. This is not necessarily the case in crude oil markets as we will discuss it in Chapter 3-Chapter 5. They test these two relationships using data for 31 commodities futures on energy, crops, precious metals, and materials from 1969 to 2006. Their main finding is to highlight the negative and non-linear relationship between month-on-month spreads between futures contracts, and level of inventories.

Gronwald (2009) employs a GARCH model to investigate jumps in crude oil prices. His model is specified as follows:

$$y_t = \alpha + \sum_{i=1}^{l} \phi_i y_{t-i} + \sqrt{h_t}\, z_t + \sum_{k=1}^{n_t} X_{t,k} + e_t \tag{2.9}$$

$$h_t = \omega + \sum_{i=1}^{q} \alpha_i e_{t-i}^2 + \sum_{i=1}^{p} \beta_i h_{t-i} \tag{2.10}$$

where $z_t \sim NID(0,1)$ and $X_{t,k}$ is jump size that follows normal distribution, and the number of jumps follow a Poisson distribution:

$$P(n_t = j \mid \varphi_{t-i}) = \frac{\lambda_t^j}{j!} e^{-\lambda_t} \tag{2.11}$$

where $\lambda_t > 0$, and $\varphi_{t-1} = \{y_{t-1},..., y_1\}$ is the history of data. They use daily returns on WTI spot crude oil prices from 30/03/1983 to 24/11/2008 to calibrate the model. Their results show the mean of the jumps is negative indicating that slow gradual increase in crude oil prices are followed by sudden drops.

Yang (2010) investigates the risk of commodity futures return by month-over-month spread and maturity, by introducing a dynamic partial equilibrium two-factor model which takes into account both producers and speculators behavior. His model replicates the average returns and volatilities of most of his selected portfolio of commodities. To do so, he starts from the future excess return of a one-period long position:

$$R^e_{i,t,T} = \frac{F_{i,t+1,T-1}}{F_{i,t,T}} - 1 \qquad (2.12)$$

and a measure called monthly basis as follows:

$$B_{i,t} = \frac{\log(F_{i,t,T_2}) - \log(F_{i,t,T1})}{T_2 - T_1} \qquad (2.13)$$

and define his two-factor model as:

$$R^e_{j,t} = \alpha_j + \beta_{j,Mkt} Mkt_t + \beta_{j,LMH_{CF}} LMH_{CF,t} + \varepsilon_{j,t} \qquad (2.14)$$

Mkt and LMH are returns of two portfolios of futures contracts he introduces. Mkt is the equally weighted futures excess returns across all commodities and maturities in the sample but LMH only focuses on shorter maturities futures contracts. His calibration is based on monthly data of 34 commodities from January 1970 to December 2008, and in some occasions he also borrows some of the parameters from other studies. His main findings is that the average excess return in shorter term month-on-month spreads is about 10% higher than longer term month-on-month spreads.

2.2 One and Multi-Factor Models of Crude Oil Prices

Schwartz et al (1990) introduce a two-factor model in order to price crude oil futures contracts. Two factors affecting crude oil prices are spot price of crude, and also the instantaneous convenience yield of oil, *δ*, and time to maturity of a future contract:

$$dS / S = \mu dt + \sigma_1 dz_1 \tag{2.15}$$

$$d\delta = \kappa(\alpha - \delta)dt + \sigma_2 dz_2 \tag{2.16}$$

where $\kappa > 0$ measures degree of mean-reversion to the long run mean log price , α , z_1 and z_2 are two Wiener processes and $dz_1.dz_2 = \rho dt$, where ρ is the correlation between the two Brownian motions. It is notable that Schwartz uses closest maturity crude oil futures contract trading on the NYMEX as a proxy for crude oil spot prices. In order to calculate the instantaneous convenience yield they use the following relationship between futures, F, and spot

$$F(S,T) = Se^{(r-\delta)(T-t)} \qquad (2.17)$$

where T is annualized time to expiry of a futures contract. They use the two consecutive forward contracts to determine the annualized monthly forward convenience yields

$$\delta_{T-1,T} = r_{T-1,T} - 12\ln\frac{F(S,T)}{F(S,T-1)} \qquad (2.18)$$

They use average weekly data for all futures contract from January 1984 to November 1988 to estimate parameters, and they also test models' out-of-sample performance for 6 months out horizon. They model's accuracy decreases for longer term vs. shorter term futures contracts. They also estimate market price of convenience yield.

Schwartz et al (1997) studies the stochastic behavior of commodity prices using three mean-reversion models. **The first model** is a one-factor that assumes the commodity spot price follows the following stochastic process:

$$dS = \theta(\mu - \ln S)Sdt + \sigma dW \qquad (2.19)$$

where θ is the mean reversion factor, μ the mean value of log of spot prices, σ the volatility, and W_t is a Weiner process. By applying Ito's Lemma he gets the following Ornstein-Uhlenbeck stochastic process:

$$d\ln S = \theta(\alpha - \ln S)dt + \sigma dW \qquad (2.20)$$

where $\alpha = \mu - \frac{\sigma^2}{2\theta}$. Under the equivalent martingale measure:

$$d\ln S = \theta(\alpha^* - \ln S)dt + \sigma dW^* \qquad (2.21)$$

where dW_t^* is an increment to the Weiner process under the equivalent martingale measure and $\alpha^* = \alpha - \lambda$, where λ measures market price of risk. His **second model** is the same as the one in **Schwartz et al (1990)** which we already discussed. His **third model** extends his two-factor model by adding a mean-reversion stochastic process for interest rate:

$$dS = (r - \delta)Sdt + \sigma_1 Sdz_1^* \qquad (2.22)$$

$$d\delta = \kappa(\alpha - \delta)dt + \sigma_2 dz_2^* \qquad (2.23)$$

$$dr = a(m^* - r)dt + \sigma_3 dz_3^* \qquad (2.24)$$

with

$$dz_1^* dz_2^* = \rho_1 dt, \qquad dz_2^* dz_3^* = \rho_2 dt, \qquad dz_1^* dz_3^* = \rho_3 dt$$

where r is interest rate, α the mean value of convenience yield, a and κ speed of adjustment coefficients, m^* risk-adjusted mean short rate of the interest rate process, z_i^*'s are Wiener processes under martingale measure, σ_i's volatility of each factor, and ρ_i's correlation coefficients. In order to calibrate coefficients of the parameters of all the three sets of models, he first drives risk-neutral relationship between futures and spot prices and then uses Kalman-Filter approach to do calibration. They use weekly observations of futures prices for crude oil and copper from 1985 through 1995 in order to test their models performances. His one-factor model implies that futures prices will converge to a fixed value for longer term contracts. His two other models imply that volatility goes down for longer term contracts. *It is notable that we directly use these two studies in in Chapter 3 and also Chapter 5 of the thesis when we are pricing crude oil futures contracts*.

Ellwanger (2015) investigates "fear" in the crude oil markets. He uses a time varying disaster probabilities and disaster fears model based on data from crude oil options and futures market to address this issue. He uses the following jump diffusion process to describe dynamics of futures prices in crude oil market:

$$\frac{dF_t}{F_t} = a_t\, dt + \sigma_t\, dW_t^Q + \int_R (e^x - 1)\mu^Q(dt, dx) \tag{2.25}$$

where a_t is the drift term, and dW_t^Q is a Brownian motion, and Q is risk neutral measure, and $\mu^Q = \mu\ (dt, dx) - v_t^Q(dx)dt$ is the jump measure under Q. He then defines risk premium, FRP, for holding a long future contract and variance risk premium, VRP, for holding a long variance swap position as follows:

$$FRP_{t,T} = \frac{1}{T-1}\left[E_t^P\left(\frac{F_T - F_t}{F_t}\right) - E_t^Q\left(\frac{F_T - F_t}{F_t}\right)\right] \tag{2.26}$$

$$VRP_{t,T} = \frac{1}{T-1}\left[E_t^P(QV_{t,T}) - E_t^Q(QV_{t,T})\right] \tag{2.27}$$

where P indicates statistical measure, and

$$QV_{t,T} = \int_t^T \sigma_s^2 ds + \int_t^T \int_R x^2 \mu(ds, dx) \tag{2.28}$$

He uses WTI crude oil options and futures data to calibrate models. The results indicate that when there is smaller risk premium in the market oil futures contracts overshoot to the upside. Conversely, they undershoot to the downside when there is more risk premium in the crude oil market.

Madan et al (1998) use the Variance Gamma (VG) process to model dynamics of log stock prices and also obtain closed form solutions for the density of returns and the prices of European options. They use data on the S&P 500 futures options from January 1992 to September 1994 to calibrate three parameters of the Variance Gamma process using maximum likelihood method. The interesting point in the calibration process is to breaking down their model into two symmetric and asymmetric versions. To get the symmetric version, they set the “symmetry” parameters at 0 and compare the results with the asymmetric version for which all three parameters should be estimated. They conclude the log-price of the S&P 500 index follows a symmetric Variance Gamma process because the log likelihood values of the two versions are statistically the same. Their study is one of the main applications of the Variance Gamma process in the mathematical finance. We will be using the VG process in Chapter 4 in pricing crude oil options.

Carr et al (1999) shows how to use the fast Fourier transform in option valuation when the characteristic function of the underlying asset's return is available in closed form. Given characteristic function of the return exists. They develop an analytical expression for the Fast Fourier transform of an option value. They illustrate the approach by applying it to the Variance Gamma process using generated numbers and compare the computation speed of their algorithm. They show that their method improves speed of calculations compared to other methods. We rely on this methodology in Chapter 4 of this thesis when we value crude oil options for Variance Gamma, Normal Inverse Gaussian, and Jump-Diffusion processes.

Benth et al (2004) develops a one factor stochastic process with jump to model dynamics of spot crude oil prices, in which a pure jump follows a Levy process. Their model is an extension of the classical geometric Brownian motion and also Schwartz's (1997) classical one-factor mean-reversion model. They introduce a Levy process L_t with Levy-Ito decomposition:

$$L_t = vt + \sigma B_t + \int_{|z|<1} z\tilde{N}((0,t],dz) + \int_{|z|\geq 1} zN((0,t],dz) \qquad (2.29)$$

where B_t is a standard Brownian motion, v and $\sigma > 0$ are constants, and N is a homogeneous Poisson random measure with compensator $dt\ell(dz_t)$, where $\ell(dz_t)$ is a Levy measure, and $\tilde{N}$ is a compensated Poisson process. They then specify crude oil spot price stochastic process as

$$S_t = \beta(t)e^{X_t} \quad (2.30)$$

where $\beta(t)$ is a non-random function of time to adjust for seasonality, and

$$dX_t = a(m - X_t)d_t + dL_t \quad (2.31)$$

where $a \geq 0$, is the speed of mean-reversion and $m>0$ indicates long-term mean of the process. They pick the normal inverse Gaussian process for the Levy part, L_t. It is notable that they do not model the dynamics of spot price as the solution of a stochastic differential equation with jump. Instead they model the price process directly and argue that this will provide better model performance. They apply their models to spot Brent oil prices and conclude that their Levy process results in superior fit compared to the Gaussian model.

Crosby (2008) proposes a mean-reversion multi-factor jump-diffusion model for pricing commodity derivatives, and applies it to crude oil options. His model allows for stochastic interest rates and also generates stochastic convenience yields. A very interesting feature of his model is allowing long term futures contracts to have smaller jumps than short term futures contracts. In parameter calibration, he uses crude oil option prices on January 25, 2005.

Askari et al (2008) study behavior of daily Brent crude oil prices by modeling oil price returns as a variance gamma process and Merton (1976) jump-diffusion process. Their daily data covers period of January 2, 2002 through July 7, 2006. They limit their data to only three-month delivery prices and ignore structure of futures contracts. They conclude that crude oil prices are dominated by upward drifts and frequent jumps causing crude oil prices not settling around a mean. It is notable that their conclusion regarding unsettling around a mean is expected because their study starts from 2002 when crude oil prices were hovering around $27/barrel and ends when crude oil prices had been going up steadily to reach to highs of around $76 in July of 2006.

2.3 Variance Swap

Carr and Wu (2008) propose a direct method to calculate risk-neutral expected value of variance swap rate by the value of a particular portfolio of options. In particular, they show that the variance swap rate can be synthesized accurately by linear combination of set of option prices. They also give a definition of variance swap premium. They apply their findings to option prices on 5 stock indices including S&P 500 and Dow Jones Industrial Average, and 35 individual stocks. Their data sample starts in January 1996 and ends in February 2003. Their analysis show that there are strong negative risk premiums for the S&P and Dow indexes.

Trolle and Schwartz (2009) follow the same methodology as Carr and Wu (2008) and address the issue of variance risk premia in crude oil and natural gas prices. In order to estimate variance risk premia, they define a variance swap payoff function at the time of expiry as follows

$$[(V(t,T) - K(t,T)]L \tag{2.32}$$

With $[(V(t,T) - K(t,T)]L$ =0 at initial time, *t=0*. V is the realized annualized return on variance, K denotes the implied variance agreed at t, and L represents the notional value of the swap contract, and $K(t,T) = E_t^Q[(V(t,T)]$, where the risk-neutral measure is denoted by *Q*. They use the following relationship to calculate the value of implied variance specified in (2.32):

$$K(t,T) = \frac{2}{B(t,T)(T-t)}\left\{\int_0^{F(t,T_1)} \frac{p(t,T,T_1,X)}{X^2}dx + \int_{F(t,T_1)}^{\infty} \frac{c(t,T,T_1,X)}{X^2}dX\right\} \qquad (2.33)$$

where *B(.)* is the price of a zero-coupon bond expiring at T, and *p* and *c* are the value of put and call options both expire at T on a futures contract expiry at T_1, with strike of *k*. Finally, the V(.) calculated as

$$V(t,T) = \frac{1}{N\Delta t}\sum_{i=1}^{N} R(t_i)^2 \qquad (2.34)$$

where $R(t_i)$ is return on futures contracts and $\Delta t = t_i - t_{i-1} = 1/252$. They apply these to daily crude oil and natural gas futures and options traded on NYMEX from 1996 to 2006. Their results show that the average variance risk premia for both commodities are negative, implying that natural gas and crude oil market pay to get short implied volatilities in both markets. They do not provide an explanation as to why this is the case in these two markets.

Swishchuk (2013) derives an explicit variance swap formula and a closed form volatility swap formula for energy prices and applies the formulae to Alberta (AECO) natural gas index in risk-neutral framework. The formula for the variance swap for a mean-reverting stochastic variance model is given by

$$EV = \frac{1}{T}\int_0^T E\sigma^2(t)dt = \frac{\sigma_0^2 - L}{aT}(1 - e^{-aT}) + L \quad (2.35)$$

2.4 Summary

Our goal in this chapter was to review some of the quantitative researches related to crude oil prices.

Kuper (2002), Giot et al (2003), Cabedo et al (2003), Fan et al (2008), Gronwald (2009), and **Marimoutou et al (2006)** use time series type models such as ARCH and GARCH and try to come up with quantitative measures of risk in crude oil markets. This is the reason they also show some interest in dynamics and asymmetry of jumps in crude oil prices. They stay in P-measure space and do not move from P to Q-measure space as their objectives are risk measurements from P-measure perspective. **Gorton et al (2007)** and **Yang (2010)** also do not go from P to Q measure space but their studies try to model risk premium and convenience yield and address nature and behavior of month-on-month spreads in crude oil prices.

In contrast, **Schwartz et al (1990)**, **Schwartz (1997), Ellwanger (2015), Benth et al (2004), Crosby (2008)**, and **Askari et al (2008)** are interested in valuation and dynamics of crude oil price returns in risk-neutral framework. In fact, because our intention in this thesis is to value crude oil derivatives therefore our work could be compared to studies done by these researchers. It is notable that works done by **Madan et al (1998)** and **Carr et al (1999)** are not in crude oil markets but we discussed these two researches in this chapter because our the methodology we used in Chapter 4 is mostly based on these two papers.

Chapter Three: Valuation of Crude Oil Futures

Valuation of any contingent claims requires a model dealing with stochastic behavior of crude oil prices. This is why modeling the stochastic behavior of crude oil's spot and forward prices are the first step in valuing crude oil contingent claims. For near month contracts this might not even be so crucial, because futures contracts on crude market usually trade very actively from one through 12 months out, however for long dated contract for which liquidity dry up significantly a the valuation of claims become very important.

The general approach is to model stochastic behavior of spot prices and so forwards under a risk neutral measure. However, main difficulties arise during calibration of the parameters which are to be used for valuating contingent claims. There are two common practices in the literature to calibrate parameters of one or two factor models in crude oil markets. First, rely on spot prices and use a process derived under the equivalent martingale measure to calibrate parameters, and then use the difference between "observed" forward and "fitted forwards" in order to estimate price of risk in crude oil market. The second approach is to start with the assumption that crude oil spot prices are "*unobservable*" and use Kalman Filter and rely on observed forward prices to calibrate parameters.

However, in this research, we use WTI crude spot prices as an "observable" variable, and take an optimization approach to simultaneously calibrate all parameters of both one and two factors models. In order to calibrate parameters, we change reference point of our data sets from “date” to “time-to-expiry”. In this setup, we avoid difficulties arising because of rolling from one contract to another.

3.1 One-Factor Model

One of the most common and simplest spot processes used in commodity markets is as follows

$$dS_t = \theta(\mu - \ln S_t)S_t dt + \sigma S_t dW_t \qquad (3.1)$$

where θ is the mean reversion factor, μ the mean value of log of spot prices, σ the volatility, and W_t is a Weiner process. The log return equivalent stochastic process of (3.1) can be written as

$$d \ln S_t = \theta(\alpha - \ln S_t)dt + \sigma dW_t \qquad (3.2)$$

where $\alpha = \mu - \frac{\sigma^2}{2\theta}$. To satisfy risk neutrality, we have (Schwartz 1997):

$$d \ln S_t = \theta(\alpha^* - \ln S_t)dt + \sigma dW_t^* \tag{3.3}$$

where dW_t^* is an increment to the Weiner process and $\alpha^* = \alpha - \lambda$, where λ measures price of risk. Hence, in the risk-neutral framework, the value of a forward contract at time t, with time to expiry of T, *F(t,T)*, is given by

$$\mathrm{F(S, t)} = \mathrm{E}_t^Q[\mathrm{S(T)}] \tag{3.4}$$

where r is risk-free rate of return, E_t^Q is risk-neutral expectation operator at time t, Q is an equivalent martingale measure, and E_t is the expectation operator at time t. Hence, given equation (3.3) and (3.4), we have:

$$\begin{aligned} F(S,T) = \mathrm{E}_0^Q[\mathrm{S(T)}] &= \exp\left(\mathrm{E}_0^Q[\ln S(T)] + 1/2\, Var_0^Q[\ln S(T)]\right) \\ &= \exp\left(e^{-\theta T} \ln S(T) + (1 - e^{-\theta T})\alpha^* + \frac{\sigma^2}{4\theta}(1 - e^{-2\theta T})\right) \end{aligned} \tag{3.5}$$

And so

$$\ln F(S,T) = e^{-\theta T} \ln S(T) + (1 - e^{-\theta T})\alpha^* + \frac{\sigma^2}{4\theta}(1 - e^{-2\theta T}) \tag{3.6}$$

We use equation (3.6) to calibrate the parameters, given observed market prices on *F(S,T)*.

3.2 Two-Factor Model

By construction, the one-factor model discussed earlier implicitly assumes that the volatility of spot and futures prices are the same, and that the risk-free interest rate is the only factor affecting holding crude from one period to another in a risk neutral framework. However, in the case of crude oil the following relationship

$$F(S,t) = S_t \; E_Q\left[\exp\left(\int_t^T (r_s)ds\right)\right] \tag{3.7}$$

does not necessarily hold. In particular, in crude markets market participants cannot easily trade crude oil in the spot market against their futures positions in order to take advantage of some "*seemingly*" observed arbitrage. For example, on the expiry of crude march future contract in 2009, the contract was trading around $37/barrel while on the same day the April contract was trading around $45/barrel. It seems this presented over 250% annualized return as one could buy March contract and sells the April at the same time. However, in practice, this would require investors to take delivery of crude on during month of March, store those barrels and then deliver them back during month of April, which is only feasible for those who had physical assets such as oil trucks or storage facilities. Therefore the non-arbitrage pricing formulas (3.6) or (3.7) do not necessarily hold for crude oil because the key role of physical assets is not accounted for in that framework.

In fact, given crude is a physical asset and holding it requires an adjustment to the cost of carry, the price of a crude forward contract would be

$$\mathrm{F(S,t)} = \mathrm{S_t}\ \mathrm{E_Q}\left[\exp\left(\int_{t}^{T}(r_s-\delta_s)ds\right]\right. \tag{3.8}$$

where δ is an adjustment factor called convenience yield. This can be thought of an "*implied*" return on holding physical crude in storage. It is notable that compared to (3.7), stochastic behavior of δ impacts value of crude oil futures contracts.

In this regard, Gibson and Schwartz (1990) and Schwartz (1997) introduced a two-factor model for which the spot price S_t follows (3.1) but its rate of growth is corrected by a stochastic mean-reverting convenience yield δ_t. In particular, the two-factor model of Gibson and Schwartz is given by:

$$dS_t = (\mu - \delta_t)S_t dt + \sigma S_t dW_t^1 \tag{3.9}$$

$$d\delta_t = \theta(\alpha - \delta_t)dt + \gamma dW_t^2 \tag{3.10}$$

where W_t^1, and W_t^1 are correlated Wiener processes, $dW_t^1 dW_t^2 = \rho dt$. It is notable that depending on supply-demand conditions, convenience yields can be either positive or negative. As in Gibson and Schwartz (1990), under the equivalent martingale measure Q the stochastic differential equations (3.9) and (3.10) can be expressed as

$$dS_t = (\mu - \delta_t)S_t dt + \sigma S_t dW_t^1$$

$$d\delta_t = [\theta(\alpha - \delta_t) - \lambda]dt + \gamma dW_t^2$$

where $dW_t^1 dW_t^2 = \rho dt$, and λ is the market price of convenience yield. In this setup, ρ >0 implies that the convenience yield keeps increasing when crude spot prices goes up due to some positive shocks to the prices, and consequently we would expect a smaller mean reversion coefficient, κ, in (3.12). This is consistent with crude market dynamics as well, as in sharp price declining environment convenience yield tend to go negative as near month contracts tend to decline more than long dated ones. For example, during first quarter of 2009 when crude market was significantly over supplied and inventory levels were quite high due to a negative demand shock to crude market followed by the financial crisis, at times the near 1-month contract were trading at 20% discount to the 2-month contract, beyond a typical month on month cost of carrying physical crude. In fact, scarcity of storage capacity at the time increased value of storage optionality and so storage operators would demand significantly higher premium to put physical crude oil in storages and so convenience yield would have dipped into negative territory.

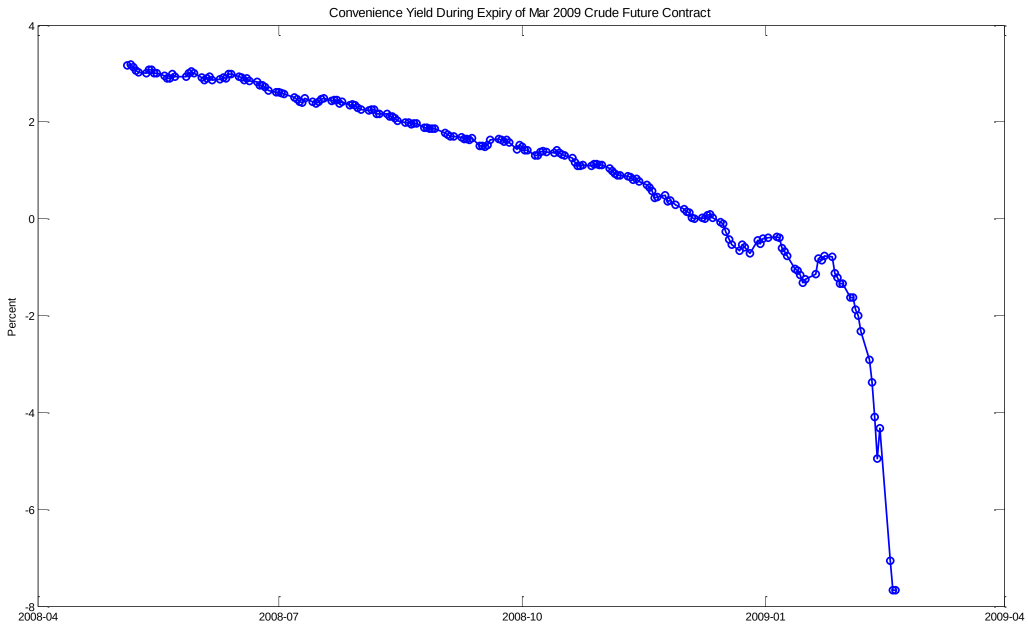


Figure 3.1: Convenience Yield During Expiry of March 2009 Crude Oil Future Contract using equation (2.18), $\delta_{T-1,T} = r_{T-1,T} - 12\ln F(S,T)/F(S,T-1)$ **(See Schwartz et al, 1990) (Data Source: Bloomberg)**

Under no-arbitrage assumption, the price of futures prices on crude oil must satisfy the following partial differential equation

$$\tfrac{1}{2}\sigma SF_{SS} + \sigma\gamma\rho SF_{SS} + \tfrac{1}{2}\gamma^2 F_{SS} + (r-\delta)SF_S + [\kappa(\theta-\delta)-\lambda]F_\delta - F_T = 0 \qquad (3.13)$$

where the solution of (3.13), $F(S,\delta,t)$, satisfies boundary condition $F(S,\delta,0) = S$ at $t = 0$ (Schwartz, 1997).

As Jamshidian and Fein (1990), and Brerksund (1991) shown, which is also cited by Schwartz (1997), the solution of (3.13) is given by

$$F(S,\delta,T) = S\exp[-\delta\frac{1-e^{-\theta T}}{\theta} + A(T)] \tag{3.14}$$

where we have

$$A(T) = (r - \alpha^* + \frac{1}{2}\frac{\gamma^2}{\theta^2} - \frac{\gamma\sigma}{\theta}) + \frac{1}{4}\gamma^2\frac{1-e^{-2\theta T}}{\theta^3} + (\alpha^*\theta + \gamma\sigma\rho - \frac{\gamma^2}{\theta})\frac{1-e^{-\theta T}}{\theta^2} \tag{3.15}$$

and $\alpha^* = \alpha - \frac{\lambda}{\theta}$.

We use equations (3.14) and (3.15) in pricing crude oil futures contract and also the calibration process.

3.3 Data and Empirical Models

In this section I describe the data. For crude oil prices I use the West Texas Intermediate (WTI) spot and futures prices as benchmark in calibrating and pricing crude oil contracts. The physical delivery point of WTI is Cushing, Oklahoma and the NYMEX oil futures contracts are settled against it. I use WTI spot and futures contracts during the period January 2000 through August 2013. The monthly futures contracts are usually traded with expiry of 1 through 60 months. However, the liquidity is relatively high for 1-month:12-month contracts but significantly dries up for longer dated contracts. Each crude contract is for 1000 barrels of crude oil delivered each day during the delivery month at Cushing. There is a sub-month forward contract called Balance-of-the-Month (BALMO) which is a contract for the rest of the month, and each day of the contract settles against WTI cash on a daily basis. This is useful for market participants to balance their existing financial or physical positions, or those who probably want to speculate on crude prices in the physical market.

3.3.1 Seasonality and Aggregation of Daily Data

Figure 3.2 shows NYMEX September crude oil prices from 2008 through 2013 in terms of trading days from 250 days to the expiry of each contract. Crude oil does not exhibit a clear seasonality in the spot and forward markets. In fact, unlike other energy commodities such as corn or natural gas which their supply or demand conditions strongly depend on seasonal supply-demand pattern, the lack of seasonal winter vs. summer factors have little impact on crude oil supply demand and that's why we observe little to no seasonality in crude oil prices. As Figure 3.2 show lack of seasonality is evident and hence we do not try to model seasonality throughout this research. However, as Figure 3.3 indicates, the relationship between crude futures contracts might show some seasonality measured by the difference between September and December contracts.

Figure 3.2: NYMEX September Crude Oil Futures Contracts from 2008 through 2013 (Data Source: Bloomberg)

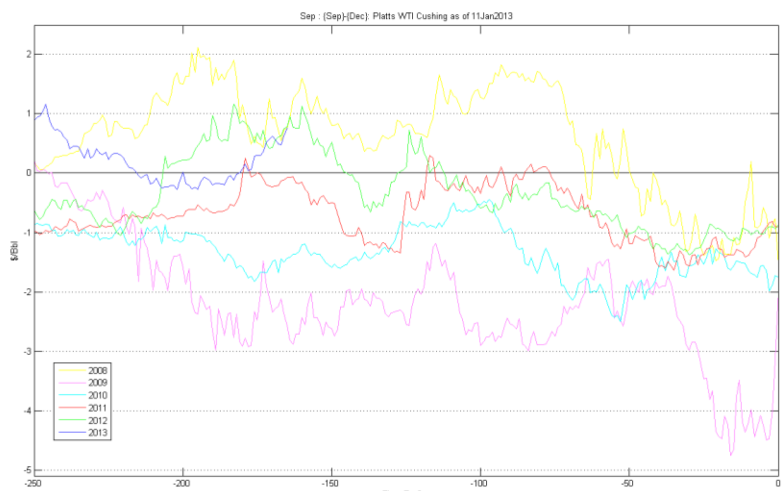


Figure 3.3: NYMEX September-December Crude Oil Spread Contracts from 2008 through 2013 (Data Source: Bloomberg)

The **Table 3.1** describes the data on futures oil contracts. The second column highlights the fact that in spite of significant volatility in crude oil price, it has been rising steadily since 2000, and at the times of big price increases the volatility goes up dramatically as it is indicated by Standard Errors (StE) of crude oil prices. The third column sheds some lights on how front vs. back dated contracts measured by 1st-12th prompts move and react to near term supply-demand information in the market. For example, before 2004-01-01 F1 was on average trading by about $2.5 over F12 and since then it averaged around (-0.5, 0.5). However, during financial crisis front month was trading over $21 discount to F12 contract, and after the crisis it has been fluctuating between -$9.9 and $11.9.

This clearly shows that averaging crude oil prices from daily to weekly or monthly data, as it is done in the literature in the past (See for example, Schwartz 1997) could eliminate significant amount of information hidden in the data and therefore should be avoided. This is the reason we only use daily settlements in this research without averaging data over time.

Table 3.1: Descriptive Information on Crude Oil Futures and Spreads

Period	Observation	Futures: Min-Max-Mean-StE	F1-F12: Range -- Mean (StE)
2000/01/01-2003/12/31	***999***	***(11.4, 37.8) -- 28.3 (5.3)***	***(-1.8, 8.4), 2.5 (1.7)***
2004/01/01-2007/12/31	***1251***	***(32.5, 145.3) -- 67.2 (24.2)***	***(-21.4, 8.7), -0.5 (3.6)***
2008/01/01-2013/08/31	***1245***	***(34.0, 113.9) -- 85.7 (15.9)***	***(-9.9, 11.9), 0.5 (4.1)***

3.3.2 Individual Contracts vs. Continuations

In modeling crude oil prices, it is important to avoid or at least correct for roll-over problem. Consider Figure 3.4 for example, which shows October-November crude oil futures spread contracts from 2004 to 2012. If one decides on September 1st, 2008 to use prompt month, F_1, or prompt i, F_i, as a proxy for crude oil futures he could buy October 2008 contract and at some point before expiry of the contract on September 22nd, 2008 he should simultaneously sell it and buy F_2, November 2008 contract, which then would be called F_1 as October 2008 rolls off the board. However, as the Figure 3.4 shows this transaction is not costless as F_1 could be trading higher or lower than F_2. Also, when roll-over took place, the price will show an artificial jump\drop right after the roll which never happened in the market. Schwartz (1997) does not correct for roll-over problem, and therefore when there is a significant difference between two consecutive contracts (e.g. October 2008 or March 2009 contracts) the finding could be impacted by the artificial jump\drop due to roll-over problem. This is why it is important to have an appropriate model design such that individual contracts feed into the model without running into roll-over problem.

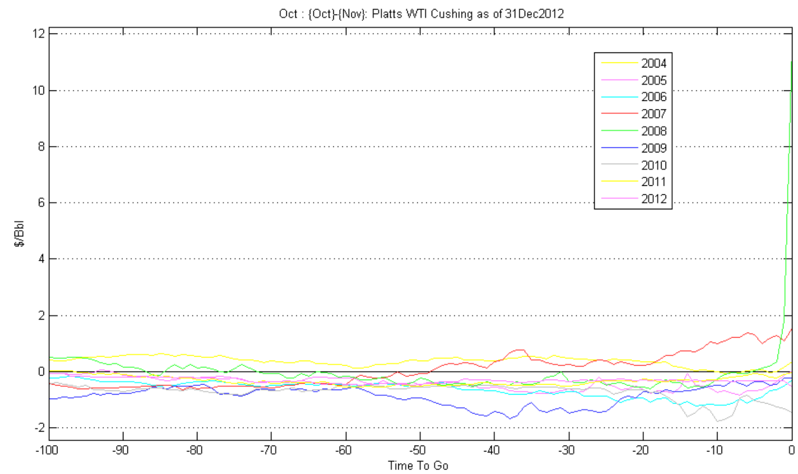


Figure 3.4: : NYMEX October-November Crude Oil Spread Contracts (Data Source: Bloomberg)

3.3.3 Empirical Models and Calibration Method

We have two sets of equations to calibrate in this section; One and two factor models. We mainly use Particle Swarm Optimization (PSO) to calibrate our models.

Equation (3.11) is the basis for one-factor models and equation (3.14) is used to calibrate two-factor models in which convenience yield is taken into account. From equation (3.11) the empirical model for the one-factor case can be written as:

$$\ln F_{ti}(S_t, T_{ti}) = e^{-\theta_j T_{ti}} \ln S_t(T_{ti}) + (1 - e^{-\theta_j T_{ti}})\alpha_j^{\,*} + \frac{\sigma_j^{\,2}}{4\theta_j}(1 - e^{-2\theta_j T_{ti}}) \qquad (3.16)$$

where F_{ti} is the model based value of forward contract for the month of *i* at time *t*, $\alpha_j^* = \alpha_j - \lambda_j$, *i = {Feb 2000, Mar 2000, ... , Dec 2013}* crude futures contracts, calendar date t $\in$ {2000-01-01:2013-08-31} where *t < expiry(i)*, and *j = 1,2,...,n.* Equation (3.16) is the first set of equations we calibrate in this section.

In order to estimate the parameters of equation (3.16) we minimize the difference between the right-hand-side of (3.16) which is implied by the model, and the left-hand-side which is observed in the market as settlement prices for futures contracts:

$$\left\{\hat{\theta}, \hat{\sigma}^2, \hat{\alpha}, \hat{\lambda}\right\} = \arg\min \sum_{t=1}^{T} \sum_{i=1}^{N} \left(\ln F_{ti} - \ln \hat{F}_{ti}(S_t, T_{ti})\right)^2 \tag{3.17}$$

where *"arg min"* **arg min** is argument of the minimum, the value of parameters in LHS for which (3.17) attains it's minimum. $\ln F_{ti}$ is observed in the market. Equation (3.17) selects the optimal parameters such that the sum of squared errors *(SSE)* between observed and the model-implied futures prices are minimized.

For the second set of empirical models that involve estimating convenience yield, equation (3.14) is used as a basis. The empirical models for the two-factor case can be written as:

$$\begin{aligned} \ln F_{ti}(S_t, T_{ti}, \delta_{ti}) = S_t \exp\{ & -\delta_{ti}\theta_j \frac{1-e^{-\theta_j T_{ti}}}{\theta_j} + (r_{ti} - \alpha_j + \frac{\lambda_j}{\theta_j} + \frac{1}{2}\frac{\gamma_j^{\,2}}{\theta_j^{\,2}} - \frac{\gamma_j \sigma_j}{\theta_j}) \\ & + \frac{1}{4}\gamma_j^{\,2} \frac{1-e^{-2\theta_j T_{ti}}}{\theta_j^{\,3}} + ([\alpha_j - \frac{\lambda_j}{\theta_j}]\theta_j + \gamma_j \sigma_j \rho_j - \frac{\gamma_j^{\,2}}{\theta_j}) \frac{1-e^{-\theta_j T_{ti}}}{\theta_j^{\,2}} \} \end{aligned} \tag{3.18}$$

where F_{ti} is the model based value of forward contract for the month of *i* at time *t*. It is notable that we use equation (2.18), $\delta_{T-1,T} = r_{T-1,T} - 12\ln F(S,T)/F(S,T-1)$ (See Schwartz et al, 1990) to get data on convenience yield as shown in Figure 3.1 for March 2009 future contract. Similar to equation (3.17) we calibrate equation (3.18) by minimizing the sum of squared errors between observed and model-implied futures prices:

$$\left\{\hat{\theta}, \hat{\sigma}^2, \hat{\alpha}, \hat{\lambda}, \hat{\gamma}, \hat{\rho}\right\} = \arg\min \sum_{t=1}^{T}\sum_{i=1}^{N}\left(\ln F_{ti} - \ln \hat{F}_{ti}(S_t, T_{ti})\right)^2 \qquad (3.19)$$

The PSO algorithm (Kennedy, *et al* 1995) is used to minimize (3.18) and (3.19). The idea behind PSO is to start by some candidate solutions (swarms) and then find an optimal solution by moving these particles around and finding their best position with respect to their own and also entire swarms' best positions. The algorithm for the PSO used in this research is as follows:

1. Set max number of particles at S and iterate for *i* = 1, ..., *S*
2. For *i=1*, initialize $g=SSE(x_1)$, where *x* is a candidate solution and g is the best solution, and SSE indicates sum of square of errors.
3. For *i* > *1* do
4. For each parameter in (3.17) and (3.19) initialize the particle's position with a uniformly distributed random vectors: $\mathbf{x}_i \sim U(\mathbf{b}_l, \mathbf{b}_u)$, where $\mathbf{b}_l$ and $\mathbf{b}_u$ are minimum and maximum for each parameter
5. Initialize each parameter's best know position to it's initial value: $\mathbf{p}_i \leftarrow \mathbf{x}_i$

6. If $SSE(pi) < SSE(\boldsymbol{g})$ update the swarm's best known position: $\mathbf{g} \leftarrow \mathbf{p}_i$
7. Initialize the particle's velocity: $\boldsymbol{v}_i \sim U(-|\boldsymbol{b}_u\text{-}\boldsymbol{b}_l|, |\boldsymbol{b}_u\text{-}\boldsymbol{b}_l|$
8. Repeat (4)-(6)
9. Pick random numbers: r_p, $r_g \sim U(0,1)$
10. For each parameter $j = 1, ..., n$ do:
11. Update the particle's velocity: $\boldsymbol{v}_{ij} \leftarrow \omega\ \boldsymbol{v}_{ij} + \varphi_p\ r_p\ (\boldsymbol{p}_{ij}\text{-}\boldsymbol{x}_{ij}) + \varphi_g\ r_g\ (\boldsymbol{g}_j\text{-}\boldsymbol{x}_{ij})$
12. Update the particle's position: $\boldsymbol{x}_i \leftarrow \boldsymbol{x}_i + \boldsymbol{v}_i$
13. Repeat (5)-(6) until a termination criteria is met

The parameters ω, φ_p, and φ_g are selected by researcher to increase the efficiency of the PSO method.

3.3.4 Empirical Results

Table 3.2 presents calibrated parameters for the one factor model (3.16) for all three periods without imposing Spot-Forward condition, by taking a *two-stage approach*. In particular, first I use equation (3.3) to calibrate parameters using spot prices without utilizing crude oil futures. In the second step we calibrate λ_j given α_j^* such that the equation (3.16) holds. In other words, forward prices are never used in calibrating parameters and only utilized once to calibrate market price of risk.

In this case, the speed of adjustment coefficient is relatively large for the first period but smaller for the second period. In all three cases market price of risk is negative as expected. It seems market participants were not willing to pay much for crude oil price risk in the first period but heading into 2008 as supply demand condition was very tight the willingness to pay for oil price risk was very high and of course declined in the third period as we entered into an over supplied market as a result of financial crisis and the recession afterwards.

Table 3.2: Calibrated Parameters without Spot-Forward Condition: One-Factor Model

Parameters	***2000/01/01-2003/12/31***	***2004/01/01-2007/12/31***	***2008/01/01-2013/08/31***
μ	***3.3148***	***3.9618***	***4.3903***
θ	***4.9837***	***0.8492***	***1.6590***
α	***3.3351***	***4.2369***	***4.4423***
λ	***-0.0422***	***-0.3681***	***-0.1296***
σ^2	***0.4674***	***0.3976***	***0.5076***
Objective Function	***349.7414***	***766.5686***	***888.8070***
Observations	***999***	***1251***	***1245***

Table 3.2 shows calibrated parameters for the model (3.16) by imposing Spot-Forward condition which is a one-stage approach in a sense that all parameters are calibrated simultaneously. The main differences in results reported in **Table 3.2** and **Table 3.3** are λ and σ^2. λ, market price of risk, in **Table 3.2** is negative implying that market participants are not compensated for the risk they take by being long crude oil given σ^2 of around 0.3976-0.5076. The **Table 3.3**, in contrast, shows higher σ^2 with positive market price of risk. Even though results in **Table 3.2** intuitively makes more sense but predicted values imply otherwise. The mean-square-errors between simulated and actual forward prices for the two versions clearly

show that calibrated parameters in **Table 3.3** results in much better simulated results than **Table 3.2**.

Table 3.3: Calibrated Parameters with Spot-Forward Condition: One-Factor Model

Parameters	***2000/01/01-2003/12/31***	***2004/01/01-2007/12/31***	***2008/01/01-2013/08/31***
μ	***2.5609***	***4.0771***	***4.2935***
θ	***0.6500***	***-0.2501***	***0.2500***
α	***2.4100***	***4.010***	***4.5098***
λ	***0.1623***	***0.1120***	***-0.3111***
σ^2	***0.99***	***0.2431***	***0.7001***
Objective Function	***100.2130***	***181.5918***	***782.6381***
Observations	***999***	***1251***	***1245***

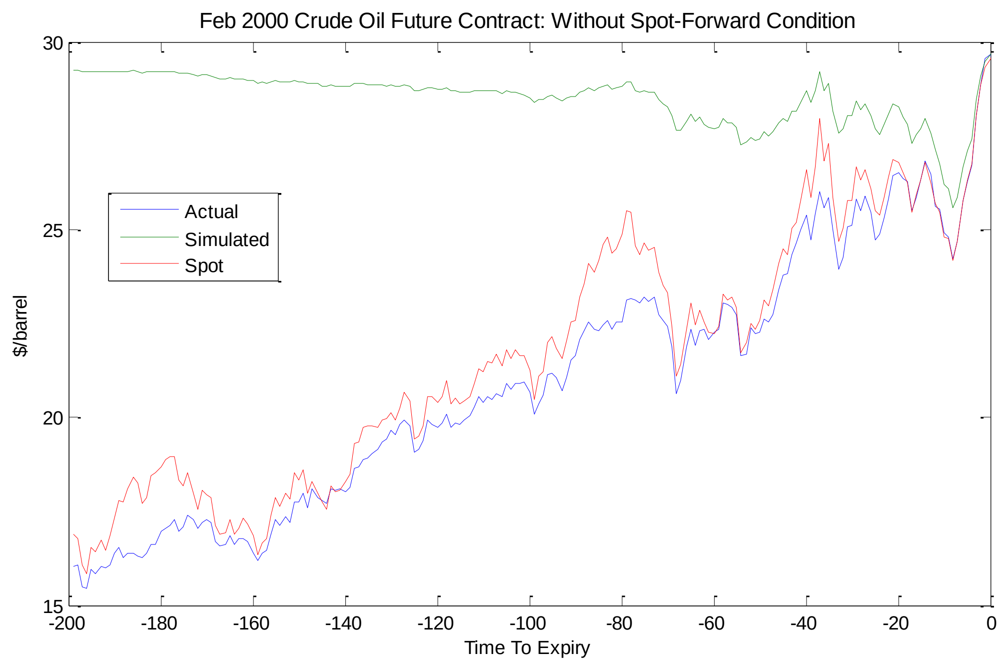


Figure 3.5: Fair Value of NYMEX Feb 2000 Crude Oil Future without Spot-Forward Condition

Figures 3.5 and 3.6 show Fair Value of Feb 2000 and Jan 2004 Futures contracts 200 days before expiry of each contract for the *two-stage approach* reported in **Table 3.2**. As it is clear, in both case the simulated crude oil prices are significantly higher than both spot and futures prices especially between 50 to 200 days to expiry.

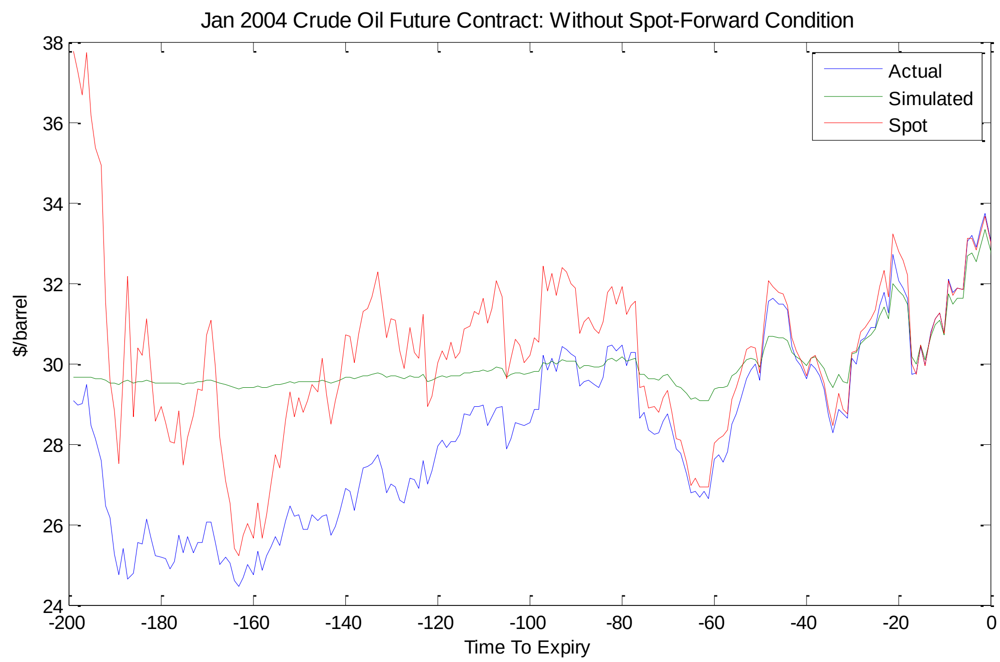


Figure 3.6: Fair Value of NYMEX Jan 2004 Crude Oil Future without Spot-Forward Condition

In order to compare the results between the two versions, now I use the calibrated parameters for the *one-stage* version presented in **Table 3.3** to simulate Feb 2000 and Jan 2004. The results for these two Futures contracts 200 days before expiry of each contract are shown in Figure 3.7 and Figure 3.8. Unlike *the two-stage approach*, calibrating all parameters at-once produces much better results. For both contracts the simulated crude oil prices are very close to actual prices and we do not see significant difference between actual and simulated values for 50 to 200 days to expiry.

Figures 3.7-3.14 compare simulated and actual futures prices for four different NYMEX crude oil prices. Similar to previous results, the one-stage approach is superior to the two-stage approach.

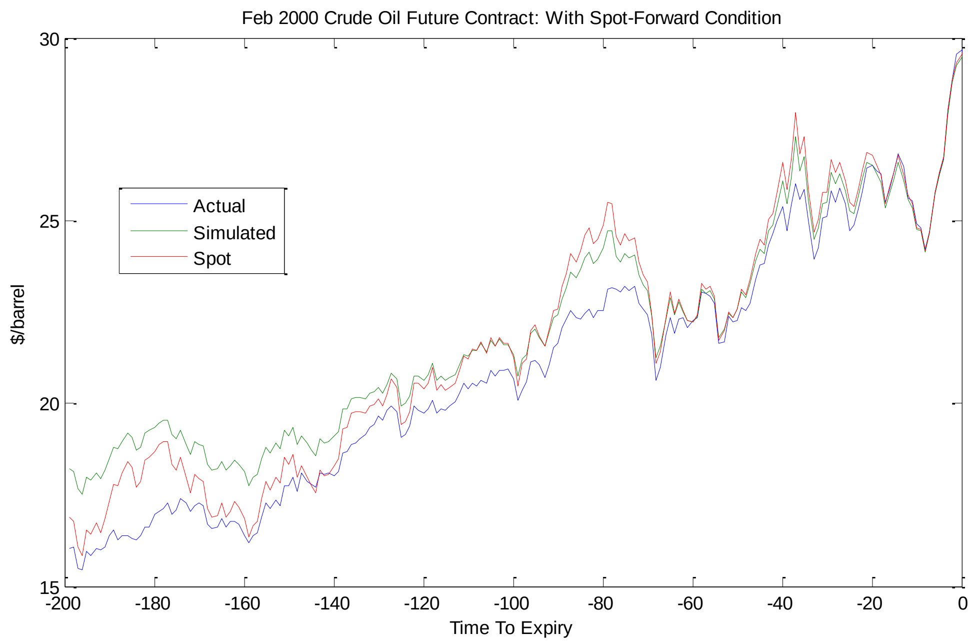


Figure 3.7: Fair Value of NYMEX Feb 2000 Crude Oil Future with Spot-Forward Condition

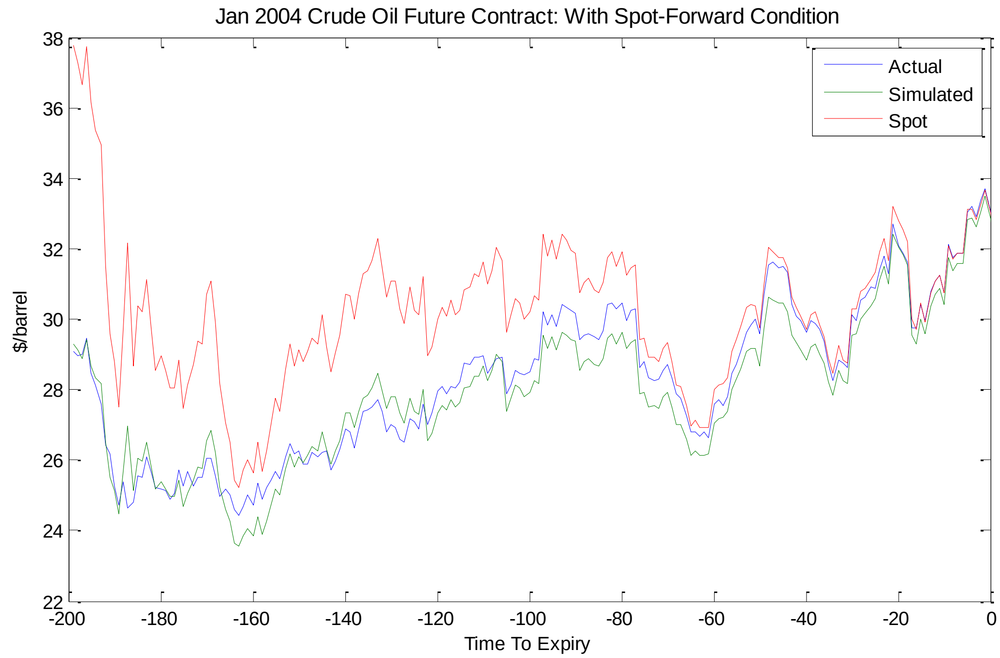


Figure 3.8: Fair Value of NYMEX Jan 2004 Crude Oil Future with Spot-Forward Condition

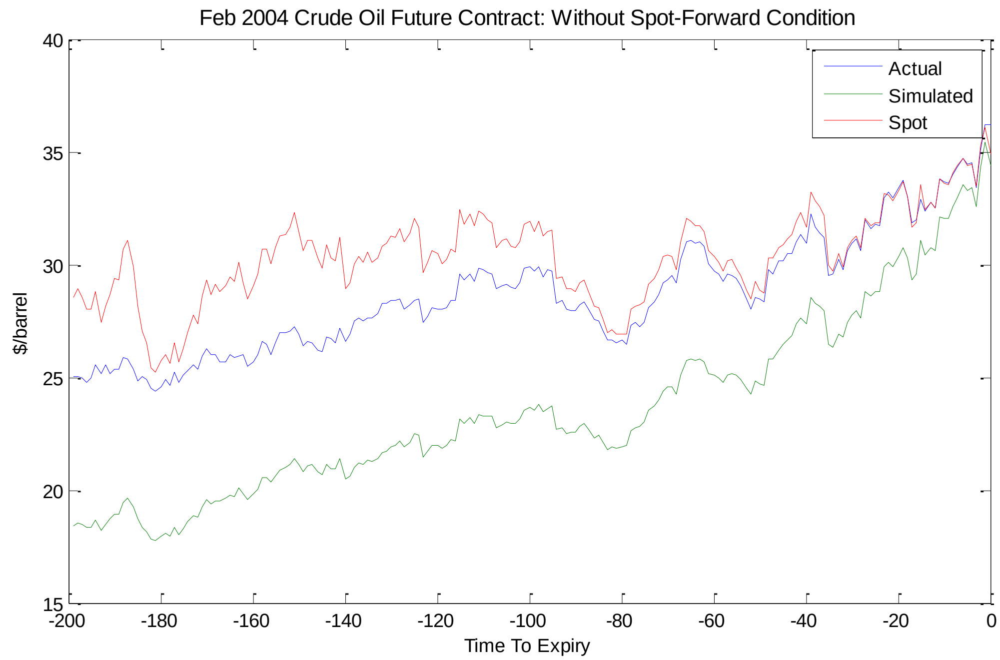


Figure 3.9: Fair Value of NYMEX Feb2004 Crude Oil Future without Spot-Forward Condition

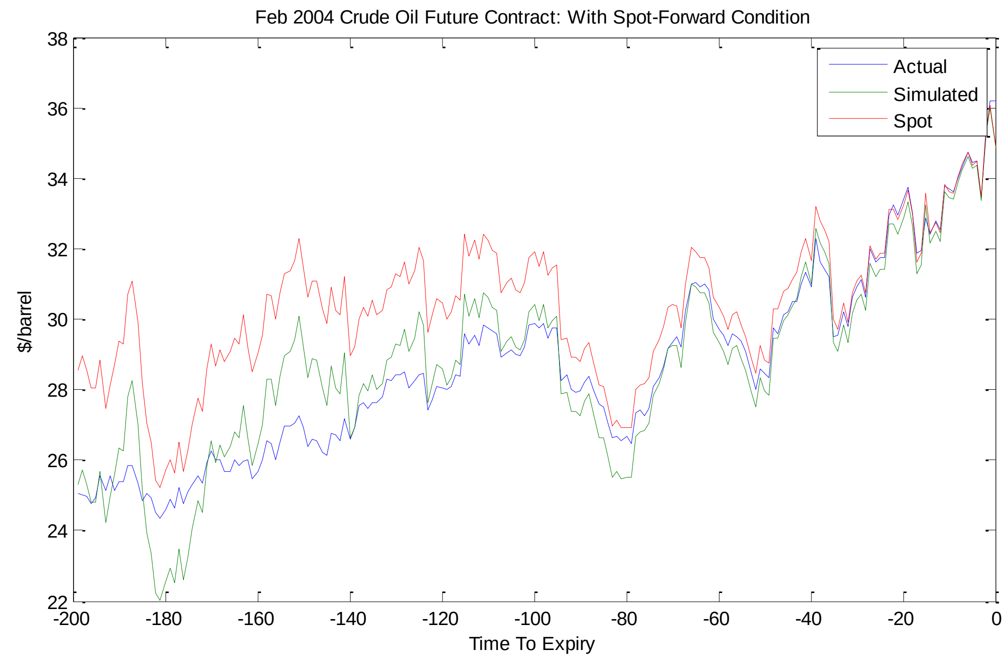


Figure 3.10: Fair Value of NYMEX Feb 2004 Crude Oil Future with Spot-Forward Condition

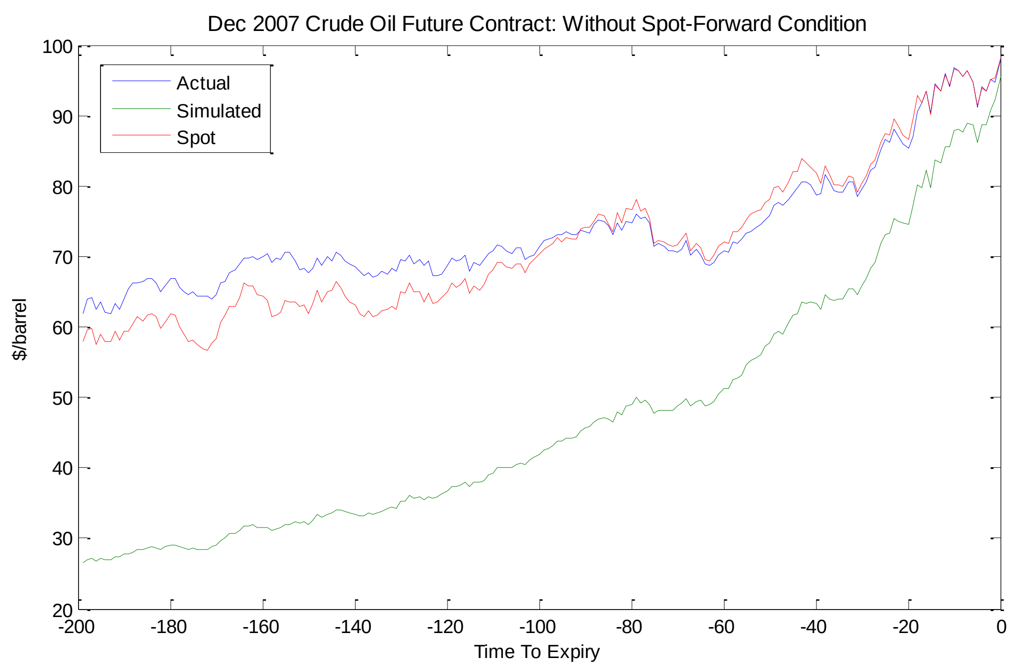


Figure 3.11: Fair Value of NYMEX Dec2007 Crude Future without Spot-Forward Condition

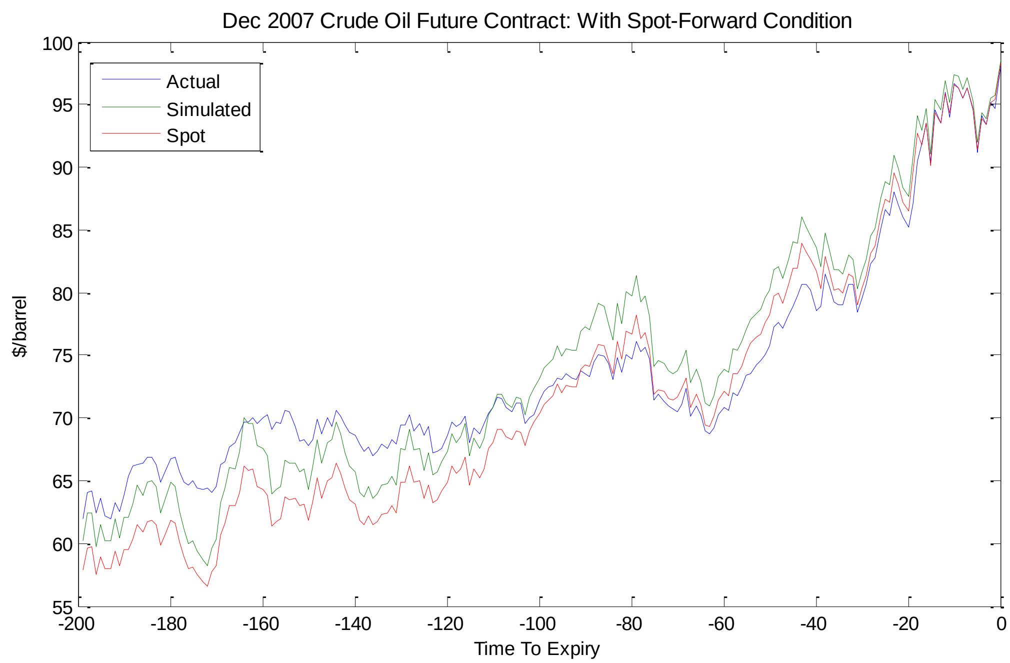


Figure 3.12: Fair Value of NYMEX Dec 2007 Crude Oil Future with Spot-Forward Condition

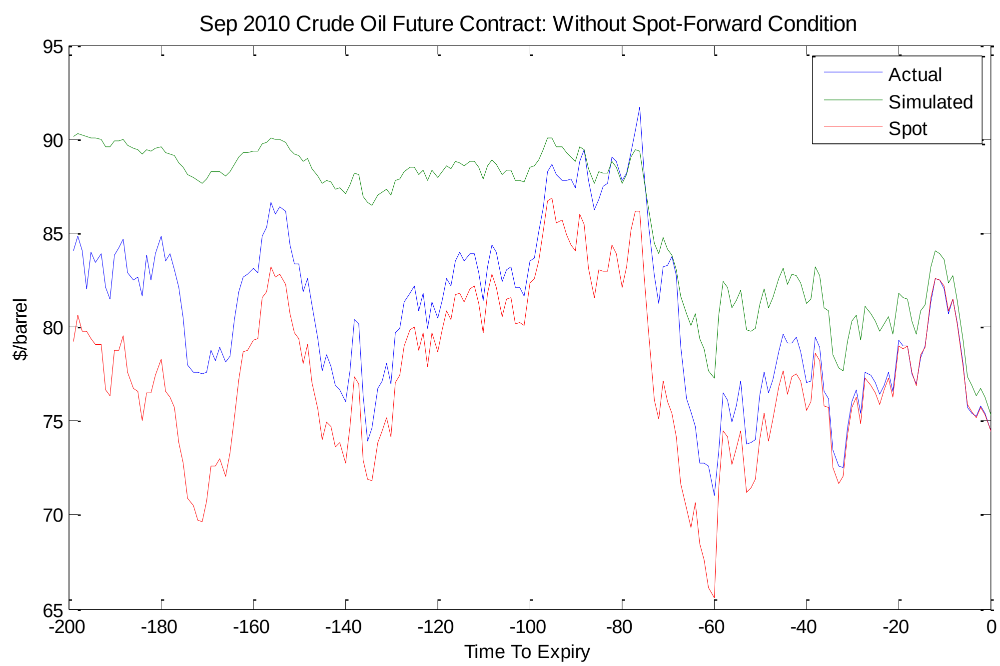


Figure 3.13: Fair Value of NYMEX Sep2010 Crude Oil Future without Spot-Forward Condition

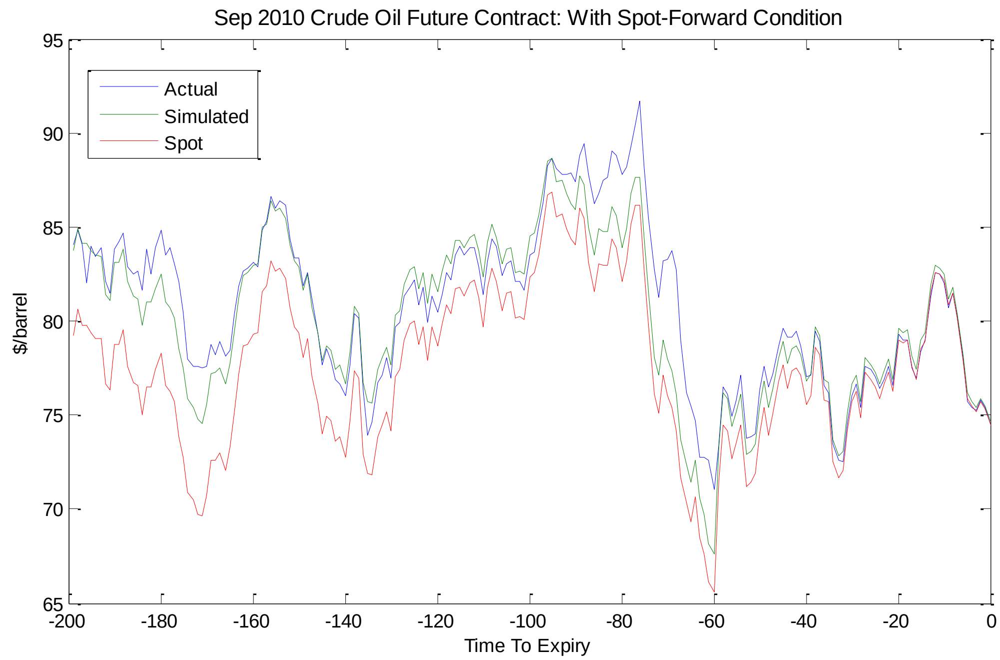


Figure 3.14: Fair Value of NYMEX Sep 2010 Crude Oil Future with Spot-Forward Condition

Table 3.4 presents calibrated parameters for the two-factor model (3.18) by minimizing objective function (3.19) using PSO algorithm. For all three periods the Spot-Forward condition is imposed and therefore this is a *one-stage approach*. The result show a very strong mean reversion coefficient, θ, of the convenience yield, δ_t, for the first two periods but relatively small one for the third period. The value of the mean convenience yield for the first, second and third period is 88.37%, 16.65%, and 123.44%, respectively. The correlation coefficient between volatility of spot prices, σ, and volatility of convenience yield, γ, is positive in all three sub periods implying that factors causing some increase in volatility of crude spot prices are probably responsible for increasing volatility of convenience yield as well.

Table 3.4: Calibrated Parameters with Spot-Forward Condition: Two-Factor Model

Parameters	*2000/01/01-2003/12/31*	*2004/01/01-2007/12/31*	*2008/01/01-2013/08/31*
μ	***2.5609***	***4.0771***	***4.2935***
θ	***2.7710***	***13.8690***	***0.1000***
α	***0.8837***	***0.1665***	***1.2344***
λ	***-2.3889***	***-2.0485***	***-0.1218***
ρ	***0.5866***	***0. 9864***	***0.5339***
σ	***0.4101***	***0.3583***	***0.3197***
γ	***4.0223***	***10.6947***	***0.5200***
Objective Function	***140.7591***	***444.2527***	***1255.726***
Observations	***999***	***1251***	***1768***

In all three cases market price of convenience yield is negative as expected. This implies that it pays to bear risk of convenience yield in crude oil market. In fact, based on equation (3.13) the negative value of convenience yield implies that an investor who is long crude oil futures contracts would ask for higher risk adjusted drift of the convenience yield $[\kappa(\theta-\delta_t)-\lambda]$ under the equivalent martingale measure. Figures 3.15-3.17 show spot, simulated and actual futures prices for three contracts. Simulated results seem satisfactory however model usually show larger error as we get further away from expiry of a contract. Figure 3.18 shows out of sample results for April 2015 contract. It is notable that this is beyond sample coverage we used to calibrate our models.

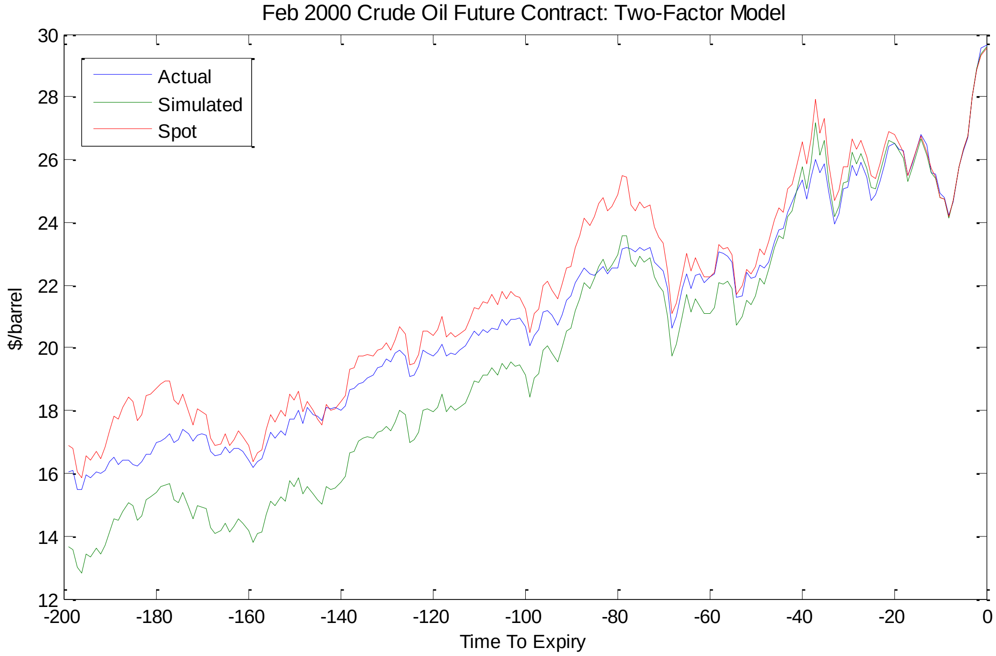


Figure 3.15: Fair Value of NYMEX Feb 2000 Crude Oil Future Using Two-Factor Model

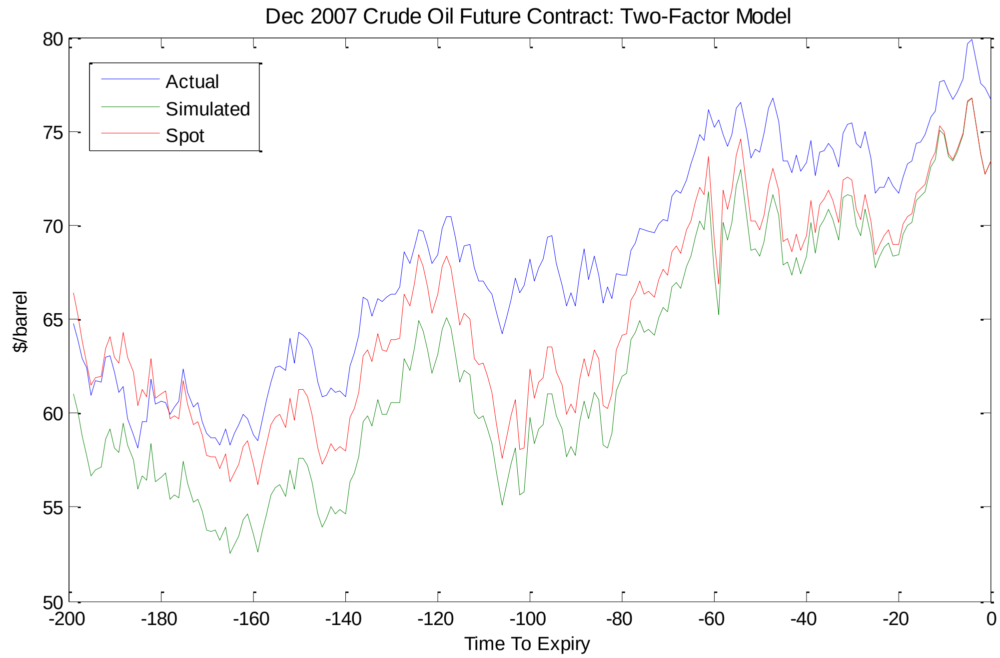


Figure 3.16: Fair Value of NYMEX Dec 2007 Crude Oil Future Using Two-Factor Model

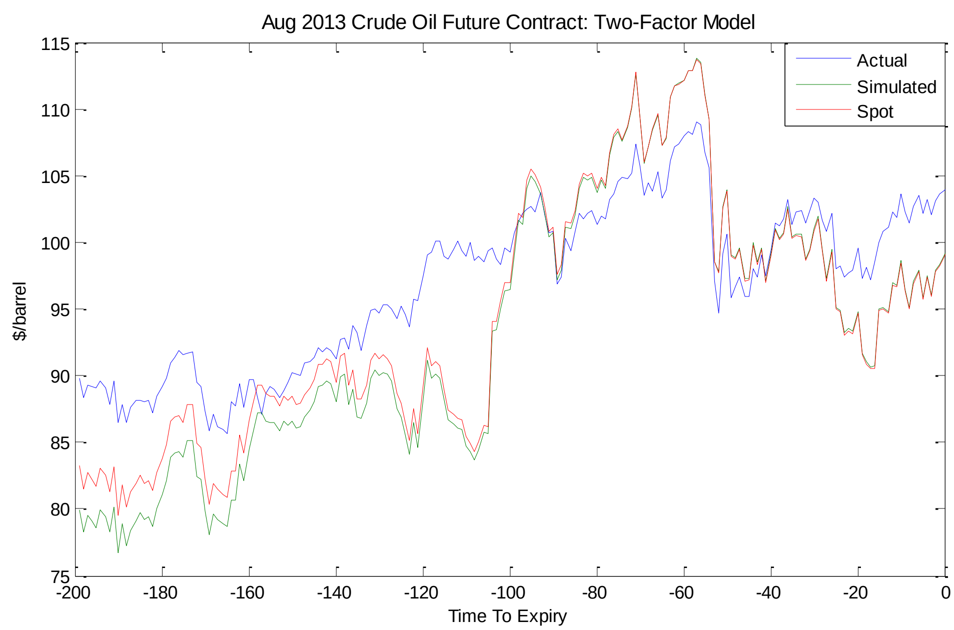


Figure 3.17: Fair Value of NYMEX Aug 2013 Crude Oil Future Using Two-Factor Model

Figure 3.18: Out-Of-Sample Fair Value April 2015 Crude Future Using Two-Factor Model

Chapter Four: Pricing Crude Oil Options Using Levy Processes

As discussed in Chapter 1, crude oil prices exhibit significant volatility over time. The distribution of returns on crude oil prices both in spot and forward markets show fat tails and skewness and they barely follow normal distribution. Figure 4.1 depicts the empirical distribution of daily returns of WTI crude oil prices from 1983:01:04 to 2015:04:21, which does not seem to follow normal distribution. In crude oil markets, tiny price movements occur with higher frequency, small and middle sized movements with lower frequency and significant changes are much more frequent than predicted by the normal law. The fat tails of the distribution indicates the tendency of the crude oil price to huge tail events in a sense that large jumps tend to happen more often than implied by normal distribution.

The Jarque-Bera test of the null hypothesis that the returns are normally distributed is rejected at the 5% significance level, implying that the deviation from normality is substantial.

Table 4.1: Descriptive Statistics of Returns on WTI Spot Crude Oil from 1990:01 through 2015:04 (Data Source: Bloomberg)

Period	Observation	Mean	StDev	Skewness	Kurtosis
1990/01/01-1992/12/31	***750***	0.0003	0.031	-1.6856	28.3106
1993/01/01-1996/12/31	***753***	0.0001	0.016	-0.4139	5.9106
1997/01/01-1999/12/31	***503***	-0.0012	0.0262	0.9451	11.6558
2000/01/01-2002/12/31	***747***	0.0006	0.0265	-0.4389	5.5271
2003/01/01-2005/12/31	***750***	0.0012	0.0241	-0.294	4.3453
2006/01/01-2008/12/31	***749***	0	0.0279	1.0834	14.507
2009/01/01-2012/12/31	***753***	0.0014	0.0258	0.1359	6.7841
2013/01/01-2015/04/30	***826***	-0.0006	0.0167	0.0406	5.7803

Table 4.1 shows descriptive statistics of returns on WTI spot crude oil for some sub-periods from 1990:01 through 2015:04. As the table shows return on crude oil prices show high Kurtosis and Skewness in all sub periods. In fact, crude oil returns not only show Skewness and Kurtosis for the whole sample but also we observe the same phenomenon in all sub-periods as well. These observations make Levy processes good candidates to model crude oil price dynamics.

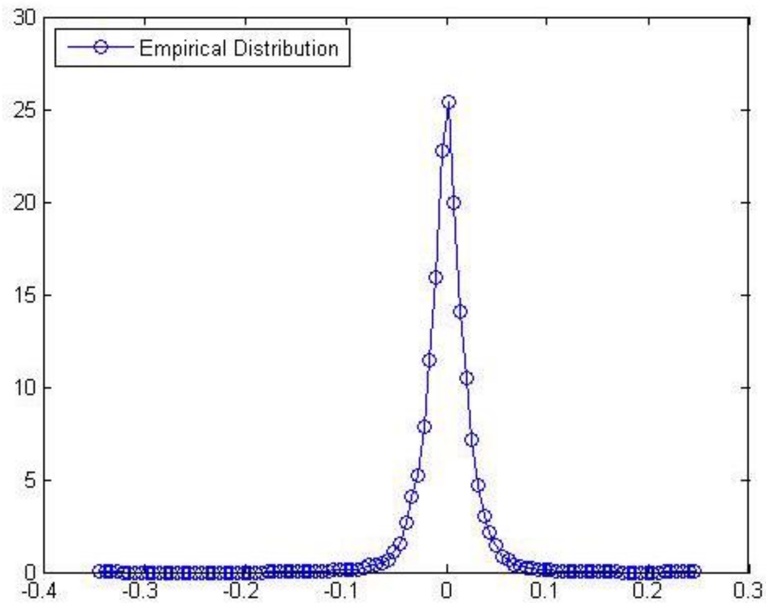


Figure 4.1: Empirical Distribution of Returns on WTI Spot Crude Oil Prices from 1983:01:04 to 2015:04:21 (Data Source: Bloomberg)

However, it is notable that simple Brownian motion with time-varying volatility could also generate a similar distribution as we have in Figure 4.1 but it does not generate the same pattern on Skewness and Kurtosis that we observe for crude oil prices as indicated in Table 4.1. For example[1] Figure 4.2 shows Empirical distribution of log price return from a Brownian motion with time varying volatility, $\sigma^2 = \exp\{\upsilon \sin(t)\}$, with υ =10 where υ is a scale factor.

[1] - I would like to thank Professor Tony Ware for providing this example.

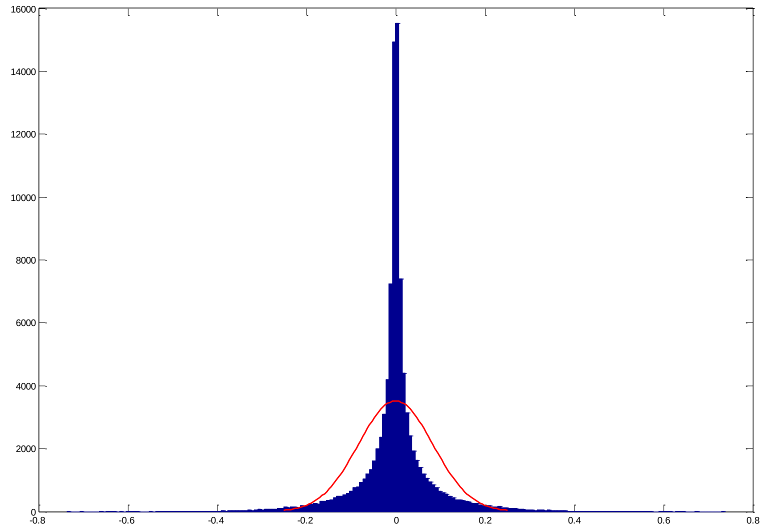


Figure 4.2: Empirical Distribution of Returns Generated from Brownian motion with time varying volatility $\sigma^2 = \exp\{\upsilon \sin(t)\}$ with scale factor υ=10

But as Table 4.2 shows, a mere time varying volatility based Brownian motion does not generate the same Kurtosis pattern as observed in crude oil price returns.

Table 4.2: Descriptive Statistics of Returns Generated from Brownian motion with time varying volatility $\sigma^2 = \exp\{\upsilon \sin(t)\}$ with scale factor υ=10

Period	Observation	Mean	StDev	Skewness	Kurtosis
1-12500	***12500***	0.0000	0.0044	0.0089	3.3982
12500-25000	***12500***	0.0000	0.0083	-0.0449	3.3559
25000-37500	***12500***	0.0001	0.0151	-0.0139	3.4009
37500-50000	***12500***	0.0001	0.0272	-0.0176	3.3358
50000-62500	***12500***	-0.0001	0.0465	-0.0448	3.3578

62500-75500	***12500***	-0.0009	0.0771	-0.0184	3.2932
75500-100000	***12500***	0.0009	0.1525	0.0161	3.5885

However, a VG or NIG process not only generates similar distribution to the one we observed in Figure 4.1 for crude oil, but it also generates pattern on Kurtosis and Skewness similar to what we have observed for crude oil price returns reported in Table 4.1. These observations motivated us to use Normal Gaussian Process (NIG), Jump Diffusion Process (JD), and Variance-Gamma Process (VG) as three Levy Processes in pricing crude oil options. Levy processes do not have these drawbacks and their tails carry heavier mass than normal distributions.

Based on review of past studies we believe this is the first application of NIG, and VG models for crude oil options on futures by calibrating their parameters using option markets. Also as mentioned in Chapter 2, Crosby (2008) has applied JDM to crude oil options on futures. This is done by Madan and Carr (1998) for stocks, where all forward option contracts have the same spot price but in the case of options on crude oil futures each option contract on a future has a different underlying price compared to other option contracts. This adds significant complexity to calibration and implementation process, as we will discuss it in detail. However, there are few studies in the literature regarding application of Levy processes to crude oil spot prices. For example, Askari and Krichene (2008) use West Texas Intermediate (WTI) crude oil spot prices from January 2, 2002 to July 7, 2006 to model oil price returns by employing Merton (1976) Jump-Diffusion and VG processes.

4.1 Levy Processes

"*A Levy process is a continuous in probability, cadlag stochastic process* $L(t), t > 0$ *with independent and stationary increments, and* $L(0) = 0$" (Gatheral, 2006, and Applebaum, 2003). Brownian motion, Poisson process, Gamma process, and Inverse Gaussian process are all examples of Levy processes.

The characteristic function of a Levy process is given by the Levy-Khintchine formula defined as follows (Applebaum, 2003):

$$\phi_t(u) = \exp\left\{t\left(iu\gamma - \frac{1}{2}u^2\sigma^2 + \int_{\Re}(e^{iux} - 1 - iux\mathrm{I}_{|x|\le 1})\mu(x)\right)\right\}, \qquad (4.1)$$

where γ is the drift parameter, and $\mu(x)$ is the Levy density. It is notable that given a Levy triplet (γ, σ^2, μ) any Levy process can be expressed as the sum of a drift term, a Brownian motion, and a jump process. This property is called Levy-Ito decomposition (Applebaum, 2003).

In this chapter we use Merton's Jump Diffusion Model (MJDM), Normal Inverse Gaussian Model (NIGM), and Variance Gamma Model (VGM), as a set of Levy based models to simulate crude oil option prices on NYMEX.

4.2 Merton's Jump Diffusion Model

Under the physical probability P, Merton's (1976) Jump Diffusion model for asset prices S(t) is specified as follows (Kou, 2008)

$$\frac{dS}{S} = \mu dt + \sigma dW + (e^{\alpha+\delta\varepsilon} - 1)dq. \qquad (4.2)$$

All three source of randomness are independent; W is the standard Brownian motion, $\varepsilon \sim N(0,1)$ that generates jump size, $e^{\alpha+\delta\varepsilon}$, α the mean of jump size, δ standard deviation of jump size, and the Poisson process

$$dq = \begin{cases} 0 & \textit{with probability} \, 1 - \lambda dt \\ 1 & \textit{with probability} \, \lambda dt, \end{cases}$$

where $\lambda > 0$ is scalar parameter.

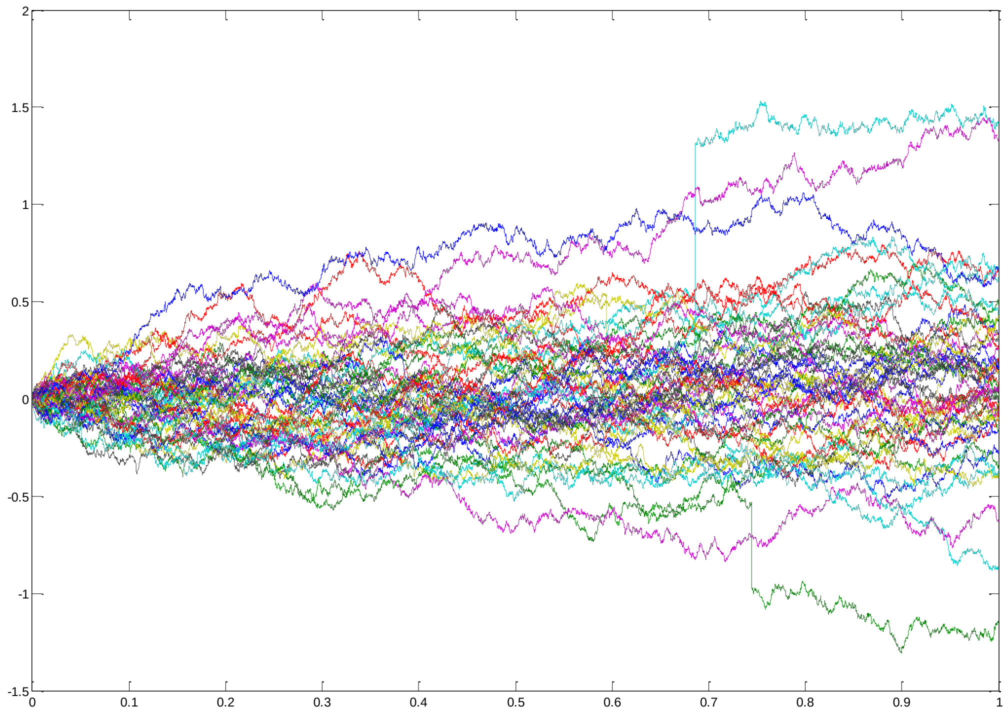


Figure 4.3: Typical Paths of a Jump-Diffusion process with $\sigma = 0.45$**,** $\delta = 0.46$**,** $\lambda = 0.33$**,**

and $\alpha = 0.12$

By applying the Levy-Khintchine formula the characteristic function of the MJDM under physical measure **P** is given by

$$\phi(u) = \exp\left\{ \mu^P - t(\frac{1}{2}u^2\sigma^2 + \lambda(e^{i\alpha u - \frac{1}{2}u^2\delta^2} - 1)) \right\}, \qquad (4.3)$$

where $\mu^P = tiu\gamma$ is the drift under P measure. The characteristic function of the MJDM under Q measure can be written as

$$\Phi(u) = \exp\left\{t(iu\mu^Q - \frac{1}{2}u^2\sigma^2 + \lambda(e^{i\alpha u - \frac{1}{2}u^2\delta^2} - 1))\right\} \tag{4.4}$$

where $\mu^Q = -\frac{1}{2}\sigma^2 - \lambda(e^{\alpha + \frac{1}{2}\sigma^2} - 1)$ is the drift under Q measure. It is notable that for crude oil options settling against spot crude oil prices the drift under Q measure would be $\mu^Q = r - \frac{1}{2}\sigma^2 - \lambda(e^{\alpha + \frac{1}{2}\sigma^2} - 1)$ however because in this research options are on futures therefore expected rate of return under Q measure will be 0.[2] Now we can simulate $\frac{dS}{S}$ given risk-neutral drift, μ^Q. Simulation of MJDM requires calibration of four parameters for volatility, σ, the jump intensity or scale parameter, λ, the mean jump size, α, and the standard deviation of the jump size, δ.

4.3 Normal Inverse Gaussian Model

The Normal Inverse Gaussian Model (NIGM) is based on the Normal Inverse Gaussian (NIG) distribution (Deville, 2008). NIG distribution's characteristic function under physical measure **P** is given by

$$\phi(u, \alpha, \beta, \delta) = \exp\left\{-\delta(\sqrt{\alpha^2 - (\beta + iu)^2} - \sqrt{\alpha^2 - \beta^2})\right\} \tag{4.5}$$

[2] - I would like to thank Professor Tony Ware for his contribution on this part.

where $u \in (-\infty, +\infty)$, $\alpha \in R$ is tail heaviness parameter, $\beta \in R$ skewness parameter, and $\delta > 0$ scale parameter.

The NIG process is given by

$$X_t = \beta\delta^2 I_t + \delta W_{I_t} \tag{4.6}$$

with $\mu^P = 0$ is the drift under P measure[3], where W is a standard Brownian motion and I_t is an Inverse Gaussian process. The characteristic function of the NIGM under Q measure is given by

$$\Phi(u) = \exp\left\{iu\mu^Q - \delta t\left(\sqrt{\alpha^2 - (\beta + iu)^2} - \sqrt{\alpha^2 - \beta^2}\right)\right\} \tag{4.7}$$

where $\mu^Q = \delta\left(\sqrt{\alpha^2 - (\beta + 1)^2} - \sqrt{\alpha^2 - \beta^2}\right)$ is the drift under Q measure. Now we can simulate asset prices

$$S_t = S_0 \exp(\mu^Q t + X_t) \tag{4.8}$$

under risk neutral measure, Q.

---

[3] - It is notable that $\Phi(u) = \exp\left\{-\delta t\left(\sqrt{\alpha^2 - (\beta + iu)^2} - \sqrt{\alpha^2 - \beta^2}\right)\right\}$ is the characteristic function under the P measure, implying that $\mu^P = 0$ and non-drift part is $-\delta t\left(\sqrt{\alpha^2 - (\beta + iu)^2} - \sqrt{\alpha^2 - \beta^2}\right)$.

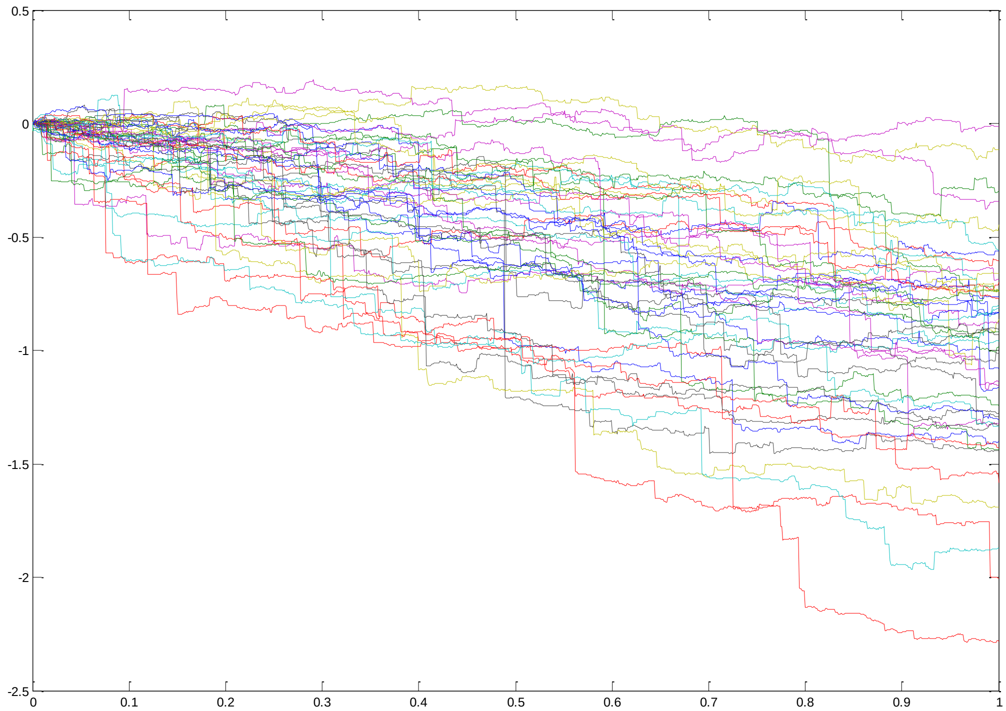


Figure 4.4: Typical Paths of a NIG process with $\alpha = 15$, $\beta = -10$, $\delta = 1$

4.4 Variance Gamma Model

The Variance Gamma Model (VGM) is based on the Variance Gamma (VG) distribution (Deville, 2008). The Gamma distribution has the following density function

$$f(x) = \frac{b}{\Gamma(a)} x^{a-1} e^{-xb} \qquad x > 0 \tag{4.9}$$

with mean a/b and variance a/b^2, where $a>0$, and $b>0$. A VG process is a process that it's increments follow VG distribution, starts at zero, and its increments are independent and stationary. The characteristic function of the VGM under physical measure **P** is given by

$$\Phi(u) = (1 - iu\theta\nu + \frac{1}{2}\sigma^2 u^2 \nu)^{-t/\nu} \qquad (4.10)$$

with $\mu^P = 0$ is the drift under P measure. θ and ν provide control over skewness and kurtosis, respectively. The distribution is negatively skews for $\theta < 0$, and larger values of ν is indication of frequent jumps and results in fatter tails.

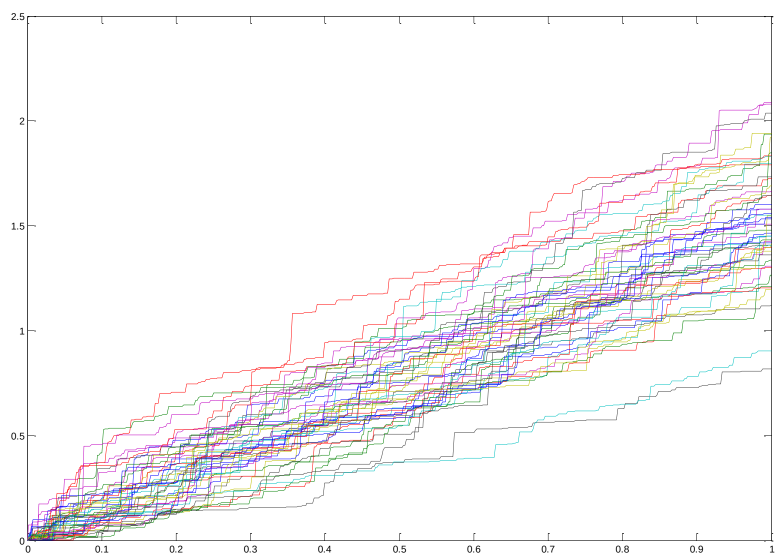


Figure 4.5: Typical Paths of a VG process with $a = 30$ and $b = 20$

The characteristic function of the VGM under Q measure is given by

$$\Phi(u) = \exp\left\{iu\mu^Q t - \frac{t}{\nu}\ln(1 - iu\theta\nu + \frac{1}{2}\sigma^2 u^2 \nu)\right\} \tag{4.11}$$

where $\mu^Q = \dfrac{\ln(1 - \theta\nu - \frac{1}{2}\sigma^2\nu)}{\nu}$ is the drift under Q measure. Now we can simulate asset prices

$$S_t = S_0 \exp(\mu^Q t + X_t) \tag{4.12}$$

under risk neutral measure, Q, and X_t is a VG process.

4.5 Application of Fast Fourier Transform For Option Pricing

In this section we use the methodology developed by Carr and Madan (1999), Chourdakis (2004), and applied by Deville (2008) to use Fourier Transform (FT) based option pricing in order value crude oil NYMEX future option prices using three models discussed so far in this chapter.

4.5.1 Fast Fourier Transform of an Option Price

Let k be the log of the option strike price, K, and $C_T(k)$ be the option price at expiry T with strike price of $\exp(k)$ as given by

$$C_T(k) = \int_k^{\infty} e^{-rT}(e^s - e^k) q_T(s) ds, \tag{4.13}$$

where $q_T(s)$ is the risk neutral density of the log price s_T. Since option price at expiry tends to S_0 as k tends to $-\infty$, the valuation function (4.13) is not square-integrable. In order to get a square-integrable function, Carr and Madan (1999) define the following modified option pricing formula

$$c_T(k) = \exp(\alpha k) C_T(k) \qquad \alpha > 0, \tag{4.14}$$

where α is called a dampening factor. The valuation model (4.14) is now a square integrable for $\forall k$. Carr and Madan specify the FT of $c_T(k)$ as

$$\psi_T(\nu) = \int_{-\infty}^{\infty} e^{-i\nu k} c_T(k) dk. \tag{4.15}$$

And then

$$C_T(k) = \frac{e^{-\alpha k}}{\pi} \int_0^{\infty} e^{-i\nu k} \psi(\nu) d\nu. \tag{4.16}$$

Carr and Madan also develop the following form for (4.15) as

$$\psi_T(\nu) = \frac{e^{-rT} \Phi_T(\nu - (\alpha + 1)i)}{\alpha^2 + \alpha - \nu^2 + i(2\alpha + 1)\nu}. \tag{4.17}$$

So the analytical expression for the value of call options can be determined by substituting (4.17) into (4.16) which can be evaluated by doing the integration.

As in Carr and Madan (1999), we use the Fast FT (FFT) algorithm to approximate (4.16). The FFT is an efficient algorithm developed by Coolery and Tukey (1965) for computing the sum

$$\sum_{j=1}^{N} e^{-i\frac{2\pi}{N}(j-1)(k-1)} x(j) \qquad \textit{for } k = 1,\ldots,N. \tag{4.18}$$

where N is number of discretization. Using the integration trapezoid rule the option formula (4.16) can be approximated as

$$C_T(k) = \frac{\exp(-\alpha k)}{\pi} \sum_{j=1}^{N} e^{-i\nu_j k} \psi_T(\nu_j)\eta, \tag{4.19}$$

where $\eta = \nu_j /(j-1)$ is the discretization step of the trapezoidal rule. The upper limit of the integration is $a = \eta N$, and k is determined as

$$k_u = -b + \lambda(u-1) \qquad \textit{for } u = 1,\ldots,N, \tag{4.20}$$

where $b = \frac{1}{2} N\lambda$. In this setup, our strikes are equally spaced between $-b$ and b . By substituting (4.20) into (4.19) and given $\eta = \nu_j /(j-1)$ we have

$$C_T(k_u) \approx \frac{\exp(-\alpha k_u)}{\pi} \sum_{j=1}^{N} e^{-i\lambda\eta(j-1)(u-1)} e^{-ib\nu_j} \psi_T(\nu_j)\eta. \tag{4.21}$$

So we can apply FFT provided that

$$\lambda\eta = \frac{2\pi}{N}. \tag{4.22}$$

It is obvious that our degree of freedom in equation (4.21) is only 2. This adds to complications of using FFT to calibrate option prices. For example, in order to have a fine grid for the integration we may choose very small discretization step but this results in relatively large strike spacing, λ . Increasing N can overcome this problem to some extent.

Choudakis (2004) adapts the methodology developed by Carr and Madan (1999) and uses the fractional FFT (FRFT) which is a much faster and more efficient algorithm developed by Bailey and Swarztrauber (1991) to calculate the sum

$$D_k(x,\alpha) = \sum_{j=0}^{N-1} e^{-i2\pi kj\alpha} x(j) \qquad \text{for } k = 0,\ldots,N-1. \tag{4.23}$$

It is notable that the FRFT reduces to standard FFT for $\alpha = \frac{1}{N}$.

The FRFT algorithm starts by defining two 2N-points vectors **y** and **z**:

$$\begin{aligned} y &= \left((x_j e^{-i\pi j^2 \alpha})_{j=0}^{N-1}, (0)_{j=0}^{N-1} \right) \\ z &= \left((e^{i\pi j^2 \alpha})_{j=0}^{N-1}, (e^{i\pi (N-j)^2 \alpha})_{j=0}^{N-1} \right) \end{aligned} \tag{4.24}$$

And so the FRFT is given by

$$D_k(x,\alpha) = (e^{-i\pi k^2 \alpha})_{k=0}^{N-1} \otimes D_k^{-1}(D_j(y) \otimes D_j(z)), \tag{4.25}$$

where $\otimes$ denotes element-by-element multiplication. In the case of FFT algorithm with the equation (4.22) the size of the discretization step and the strike spacing has to double over and over to increase accuracy of the approximation. For example, as it mentioned by Chourdakis (2004) the Carr and Madan (1999) had to set upper integration part to 1024 for a 4096 point FFT with integration step of $\eta = 1024/4096 = 0.25$, and log-strike grid step of $\lambda = 2\pi/(N\eta) = 0.006$. However, out of all these 4096 calculated option prices, only 67 will be in the 20% of ATM strike and of course close to 98% of all the calculated ones are not used. However, by using FRFT algorithm, we do not need to impose the restriction between parameters, $\lambda\eta = 2\pi/N$, specified in equation (4.22). Therefore, we can independently increase integration step, η, and the log-strike grid step, λ, without necessarily increasing number of points.

4.5.2 Data and Calibration of Parameters

In order to simulate the three Levy processes discussed in this chapter and value crude oil futures options based on these processes we need to calibrate parameters of these models. We use European-style crude oil futures options that are actively traded in NYMEX.

We obtained the option prices from Bloomberg. All data points are for the settlement date of April 24th, 2015. We collected option prices for June 2015, July 2015, September 2015, December 2015, March 2016, June 2016, and July 2016 futures contracts. This gives us time to expiries between 18 and 417 days as shown in Table 4.3. The last row in the Table shows settlements on WTI crude oil futures on the same date.

Table 4.3: WTI Crude Futures and Options Prices with Strikes (Data Source: Bloomberg)

		Days to Expiry After Market Close on April 24, 2015						
		18	52	113	205	297	387	417
		Settlements of Each Contracts						
	Strike	Jun 2015	Jul 2015	Sep 2015	Dec 2015	Mar 2016	Jun 2016	Jul 2016
WTI Crude Futures Options	50.00	7.36	9.60	11.78	13.86	14.94	15.61	15.61
	51.00	6.43	8.73	10.96	13.06	14.31	15.04	15.04
	52.00	5.53	7.89	10.15	12.28	13.38	14.29	14.29
	53.00	4.68	7.08	9.37	11.52	12.80	13.55	13.55
	54.00	3.87	6.30	8.62	10.78	11.88	12.83	12.83
	55.00	3.11	5.54	7.89	10.07	11.16	11.88	11.88
	56.00	2.44	4.84	7.19	9.37	10.65	11.19	11.19
	57.00	1.86	4.19	6.52	8.70	9.77	10.52	10.52
	58.00	1.35	3.58	5.88	8.05	9.11	9.87	9.87
	59.00	0.96	3.01	5.27	7.42	8.48	9.24	9.24
	60.00	0.67	2.51	4.69	6.81	7.87	8.63	8.63
	61.00	0.45	2.07	4.15	6.23	7.29	8.05	8.05
	62.00	0.31	1.67	3.66	5.67	6.73	7.48	7.48
	63.00	0.21	1.35	3.21	5.16	6.19	6.94	6.94
	64.00	0.15	1.08	2.79	4.67	6.00	6.43	6.43
	65.00	0.11	0.85	2.40	4.22	5.24	5.95	5.95
	66.00	0.09	0.68	2.07	3.79	5.10	5.80	5.80
	67.00	0.07	0.53	1.78	3.40	4.39	5.05	5.05
	68.00	0.06	0.42	1.53	3.05	4.00	4.64	4.64
Futures		57.15	58.90	60.50	62.03	62.98	63.57	63.68

In terms of strike prices, we picked 19 strikes for all contracts from \$50 to \$68. In order to get enough samples in In-the-Money (ITM) and Out-of-the-Money (OTM) ranges our strikes increases by increment of \$1. In this case our option prices fall in the range of $\pm 20\%$ from At-the-Money (ATM) strikes. Table 4.4 summarizes the ITM-ness and OTM-ness of each individual option prices in percentages.

Table 4.4: WTI Crude Futures and Options Prices with Money-Ness (**Data Source: Bloomberg)**

	Strike	Jun 2015	Jul 2015	Sep 2015	Dec 2015	Mar 2016	Jun 2016	Jul 2016
		Days to Expiry After Market Close on April 24, 2015						
		18	52	113	205	297	387	417
		At the Moneyness of Option Contracts						
WTI Crude Futures Options	50.00	13%	15%	17%	19%	21%	21%	21%
	51.00	11%	13%	16%	18%	19%	20%	20%
	52.00	9%	12%	14%	16%	17%	18%	18%
	53.00	7%	10%	12%	15%	16%	17%	17%
	54.00	6%	8%	11%	13%	14%	15%	15%
	55.00	4%	7%	9%	11%	13%	13%	14%
	56.00	2%	5%	7%	10%	11%	12%	12%
	57.00	0%	3%	6%	8%	9%	10%	10%
	58.00	-1%	2%	4%	6%	8%	9%	9%
	59.00	-3%	0%	2%	5%	6%	7%	7%
	60.00	-5%	-2%	1%	3%	5%	6%	6%
	61.00	-7%	-4%	-1%	2%	3%	4%	4%
	62.00	-8%	-5%	-2%	0%	2%	2%	3%
	63.00	-10%	-7%	-4%	-2%	0%	1%	1%
	64.00	-12%	-9%	-6%	-3%	-2%	-1%	-1%
	65.00	-14%	-10%	-7%	-5%	-3%	-2%	-2%
	66.00	-15%	-12%	-9%	-6%	-5%	-4%	-4%
	67.00	-17%	-14%	-11%	-8%	-6%	-5%	-5%
	68.00	-19%	-15%	-12%	-10%	-8%	-7%	-7%
Futures		57.15	58.90	60.50	62.03	62.98	63.57	63.68

In order to calibrate our parameters we take a similar approach as in Chapter 2 to recover parameters. In this case we minimize the difference between the option prices observed in the market and those generated by our models:

$$\{\hat{\theta}\} = \arg\min \sum_{i=1}^{N} \left(CM_i - CO_i(\theta)\right)^2 \qquad (4.26)$$

Equation (4.26) selects the optimal parameters such that the sum of squared errors *(SSE)* between observed and the model-implied option prices are minimized. The calibrated parameters for the three models are reported in Table 4.5.

Table 4.5: Calibrated Parameters of JDM, VG, and NIG Processes

Parameters	***JDM***	***VG***	***NIG***
θ	***NA***	***0.1810***	***NA***
δ	***0.4628***	***NA***	***2.5225***
α	***0.1247***	***NA***	***-18.531***
λ	***0.3318***	***NA***	***NA***
σ	***0.4272***	***0.4115***	***NA***
ν	***NA***	***0.0783***	***NA***
β	***NA***	***NA***	***-7.4451***

For the JDM, the volatility of the diffusion component estimated at $\sigma = 0.4272$ and is dominated by the standard deviation of the jump size component, $\delta = 0.4628$. Intensity of jump, $\lambda = 0.3318$ is high and indicates that oil prices are characterized by frequent jumps. The mean of the jump component size estimated at $\alpha = 0.1247$ indicating skewness in crude oil price returns. In the case of VG, the volatility parameter estimated at $\sigma = 0.4115$ which is slightly smaller than volatility of JDM. The tail parameter estimated at $\upsilon = 0.0783$ implying frequent jumps in crude oil returns. The θ calibrated at 0.181 indicating that crude oil prices are right-skewed, in a sense that market assigned higher probability for oil prices to go over future prices. Regarding NIG, the skewness parameter estimated at $\beta = -7.4451$ implying that the density skewed to the left, and the scale parameter calibrated at $\delta = 2.5225$. The estimated value of α implies that crude oil price returns have heavy tails.

Figure 4.6 – Figure 4.17 show results of our calibrated option prices compared to market prices, and also results generated by Black Scholes option pricing method. Our results indicate that NIGM performs well for options with strikes that are close to at the money strikes. Similar to NIGM our results indicate that we get relatively good results for near at the money options than others when we use JDM and VG models as well. For example, relative error for even 5% out of money options for NIGM does not exceed 5%.

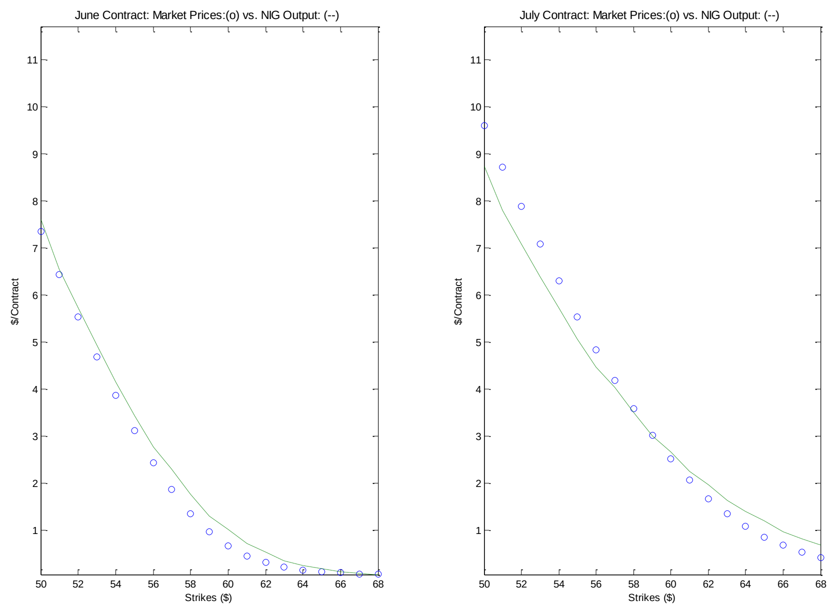


Figure 4.6: WTI June & July Market vs. NIG Based Option Prices

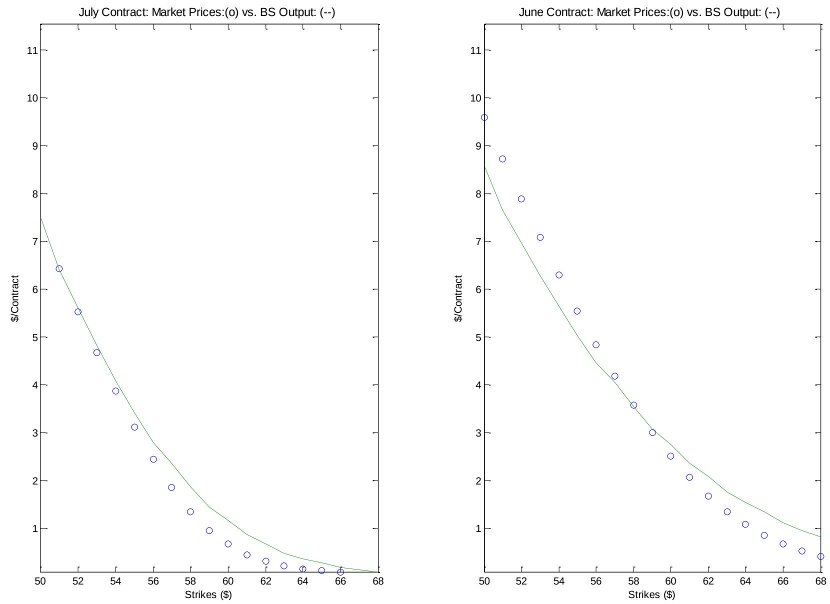


Figure 4.7: WTI June & July Market vs. BS Based Option Prices

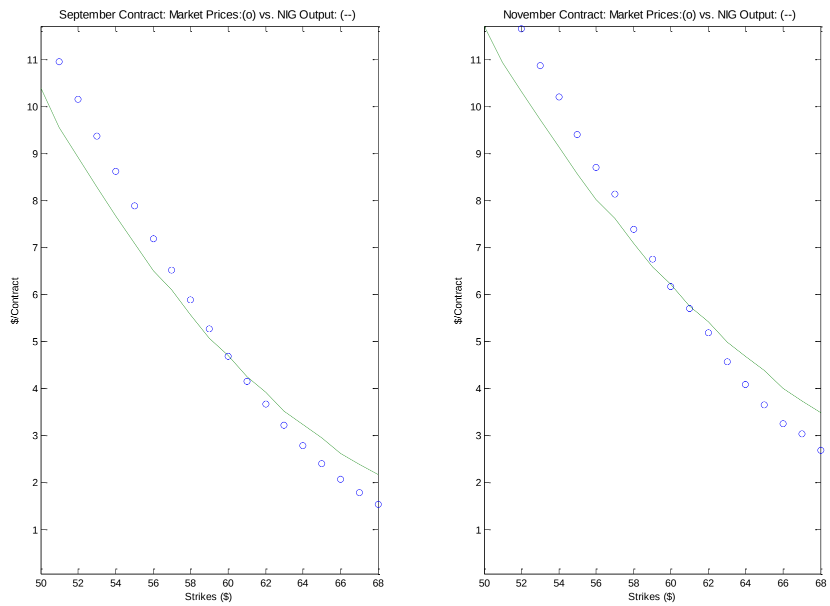


Figure 4.8: WTI September & November Market vs. NIG Based Option Prices

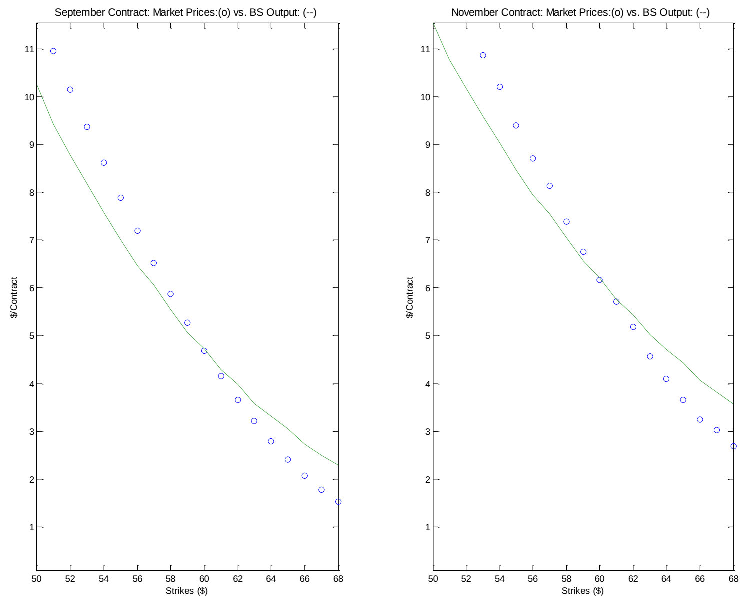


Figure 4.9: WTI September & November Market vs. BS Based Option Prices

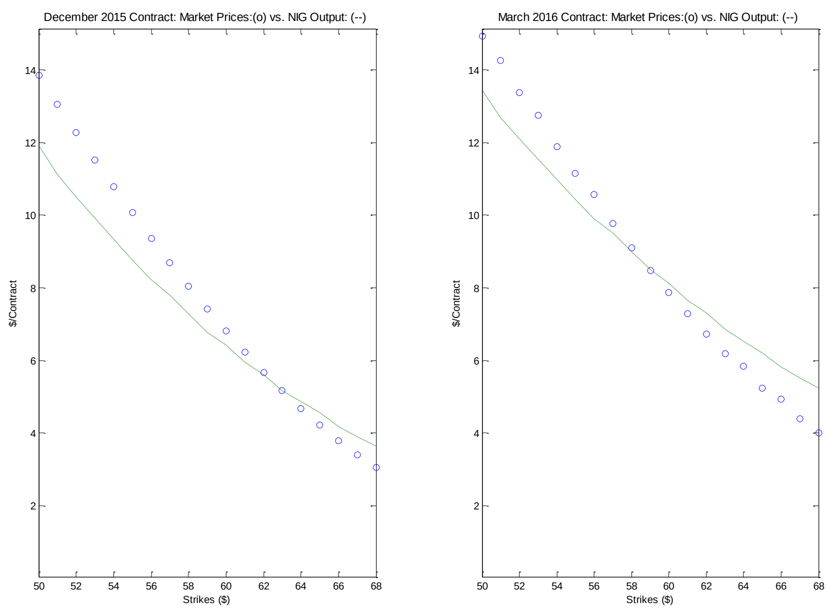


Figure 4.10: WTI December 2015 & March 2016 Market vs. NIG Based Option Prices

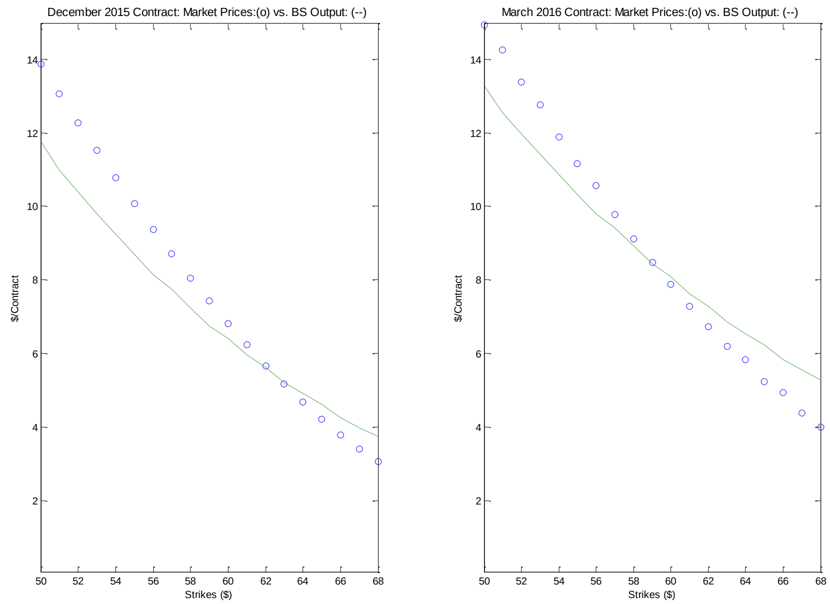


Figure 4.11: WTI December 2015 & March 2016 Market vs. BS Based Option Prices

Figure 4.12 and Figure 4.13 show results of our calibrated option prices using JDM. Similar to NIGM our results indicate that we get relatively good results for near at the money options than others. For example, relative error for even 5% out of money options for NIGM does not exceed 5%, however the error goes up significantly for out of money options.

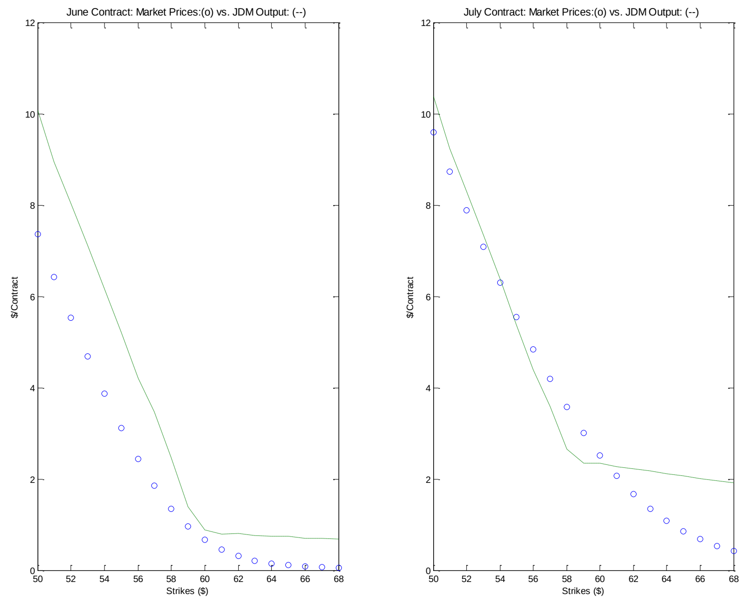


Figure 4.12: WTI June & July Market vs. JDM Based Option Prices

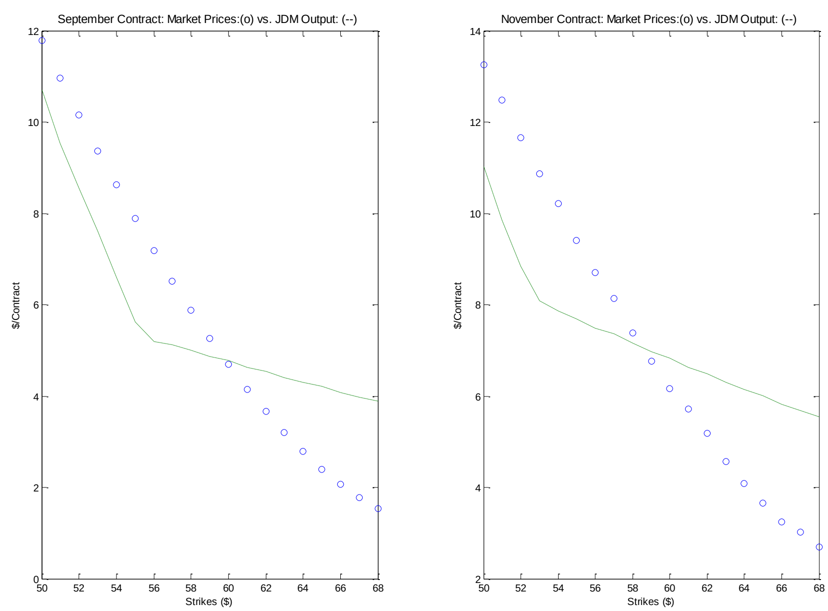


Figure 4.13: WTI September & November Market vs. JDM Based Option Prices

Figure 4.14 – Figure 4.17 show results of our calibrated option prices using VGM. Similar to NIGM and NIGM our results indicate that we get relatively good results for near at the money options than others. For example, relative error for even 5% out of money options for NIGM does not exceed 5%, but the error goes up to 20% for 10% out of money options.

We use option based calibration of MJDM and generate some sample paths for crude price returns. The results are shown in Figure 4.18**.**

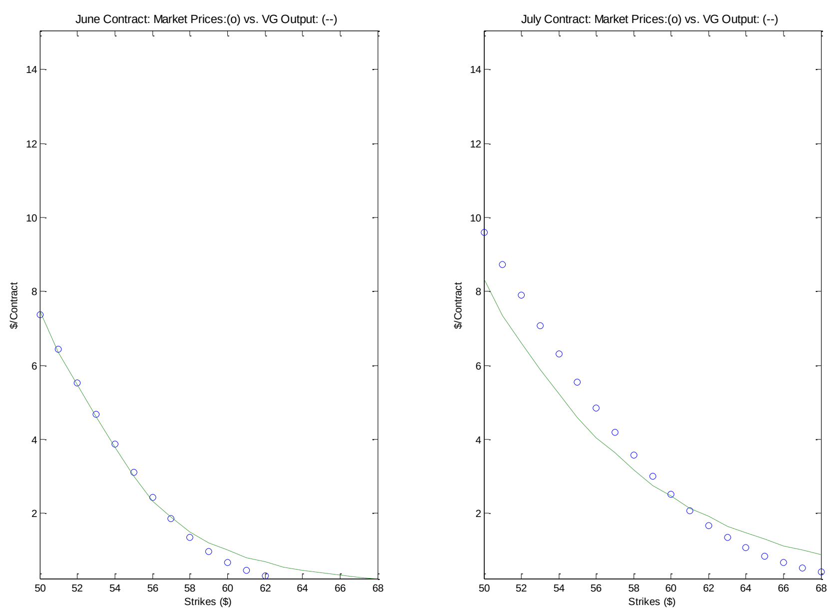


Figure 4.14: WTI June 2015 & July 2015 Market vs. VG Based Option Prices

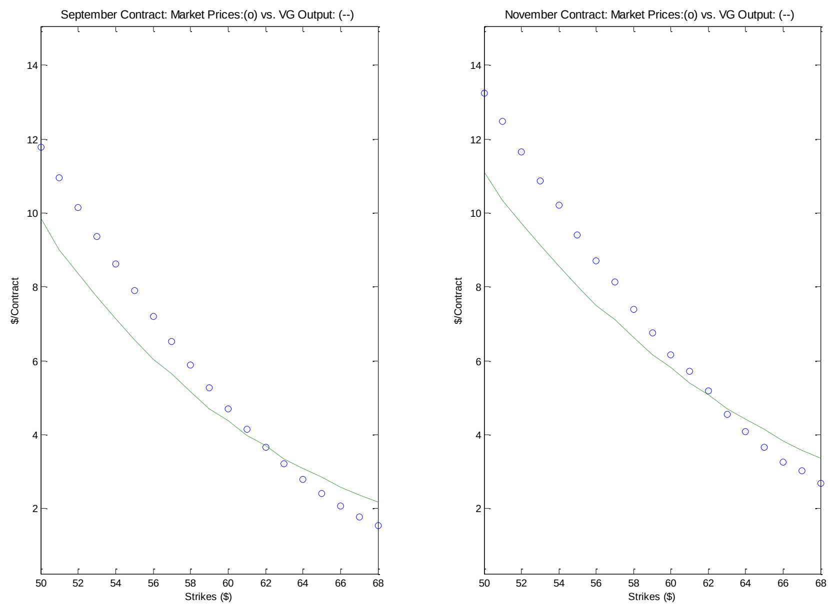


Figure 4.15: WTI September & November 2015 Market vs. VG Based Option Prices

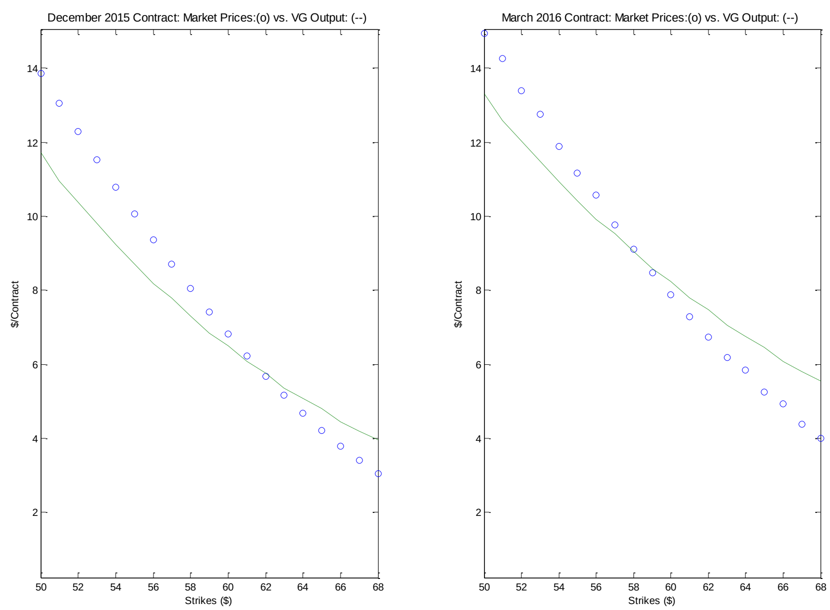


Figure 4.16: WTI December 2015 & March 2016 Market vs. VG Based Option Prices

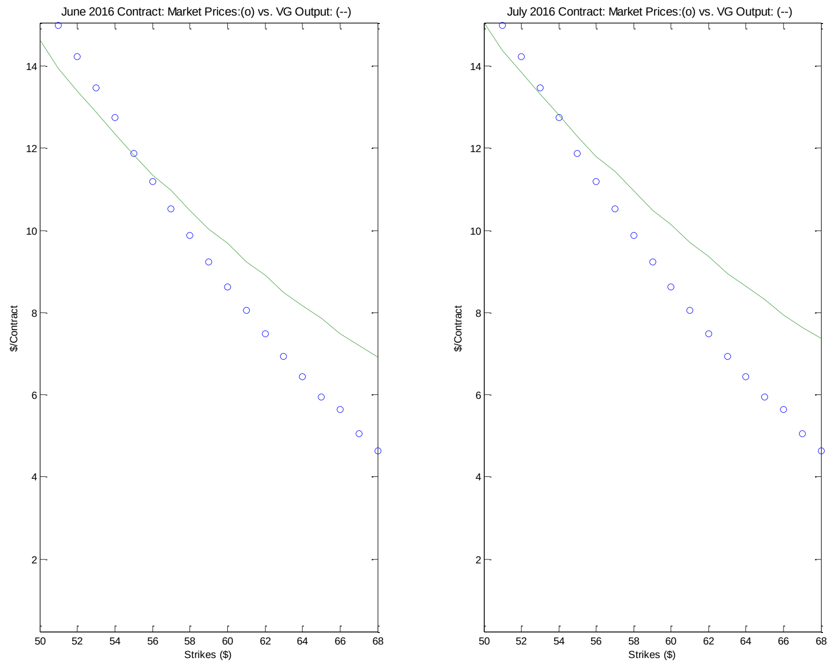


Figure 4.17: WTI June & July 2016 Market vs. VG Based Option Prices

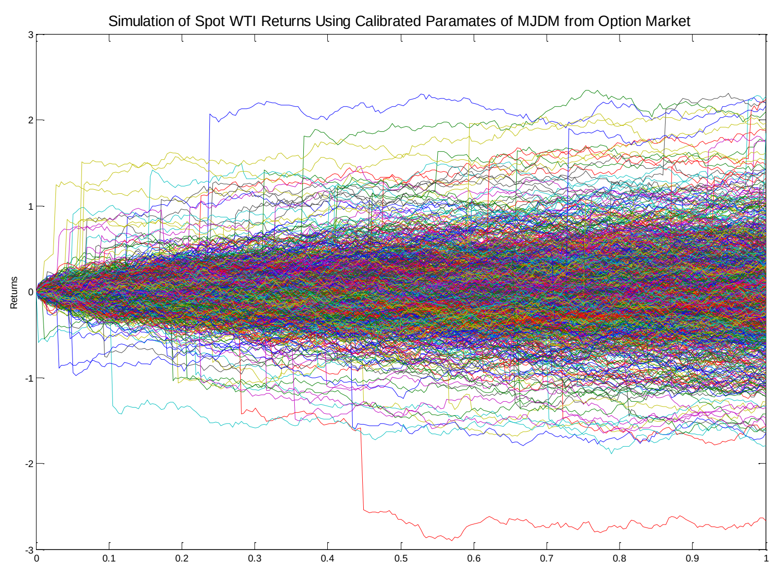


Figure 4.18: Option Market Based Simulation of MJDM for WTI Spot Returns

4.6 Summary

The distribution of returns on crude oil prices is very important for policy makers, crude oil producers and refineries. We used most recent data through April 2016 from crude oil futures and options markets to model dynamics of crude oil prices and value crude oil options. Our results indicate that crude oil prices show significant jumps that are very frequent. Crude oil price returns show skew as well. These findings are consistent across all three models we used in this research. In the case of JDM, the volatility of size of the jumps is bigger than volatility of the diffusion part. The VG process results in slightly smaller volatility than JDM. The mean of the jump component size implied by JDM, and skew parameter of VG process both indicate existence of right-skew in crude oil price returns but the NIG process implies that the density of returns are skewed to the left.

Chapter Five: Pricing Crude Oil Futures by Utilizing Storage and Refinery Data

In Chapter 3 we calculated fair values of crude oil futures contracts using a set of one-factor and two-factor models. In Chapter 3 we implicitly assumed that the log returns of crude oil prices are linearly related to their lagged values. The framework in Chapter 3 does not let a crude oil producer, refiner or other market participants to normalize fair values of crude oil futures contracts with respect to crude oil fundamental factors such as level of crude oil inventories. In addition, the framework that was set up in Chapter 3 does not help us to understand reason behind price moves, and it makes it really hard to build a bridge between say crude oil refinery outages and the price of refined products or crude oil itself.

In this chapter we take a different path to come up with fair value of crude oil futures contracts. We use WTI crude spot prices as an "observable" variable, and take an optimization approach to simultaneously calibrate all parameters of both one and two factors models. We provide a practical method to help producers, market makers and other market participants to take spot prices traded in the market as given, and price long dated crude forward contracts on a real time basis.

We then use the setup of the two-factor model and incorporate crude oil inventories and draws to the valuation process. In this case, producers and other market participants can use physical information and feed them into futures pricing methods. We presented a set of structural models to show how the dynamics of crude market works. As corollary, we use the framework to calibrate crude oil futures, crude oil spreads, and crack spreads. We performed sensitivity analysis of fair value of futures and future spreads with respect to level of inventories and storage draws.

In fact our goal in this chapter is to build a bridge between risk-neutrality and structure of the crude oil markets. This is an improvement to what we have done in Chapter 3. We provide out-of-sample results and compare them with actual data as well.

From modeling point of view we try to address three assumptions that our models in Chapter 3 were based on. For example, if inventory levels are significantly lower than a year ago, the next move is most likely to the upside rather than to the downside; so we try to address this issue. In addition, if refineries are operating on a very tight margin the size of a positive jump could be potentially significantly bigger than the size of a negative jump. In particular, factors affecting crude oil prices could potentially follow any stochastic process but it is fair to say that a shock to a risk factor that affects crude oil prices and say generates an increase in oil prices, does not mirror the decline in oil prices, if that factor had gone the other way. That is,

$$\frac{\partial r_t}{\partial x_t(up)} \neq \frac{\partial r_t}{\partial x_t(down)} \qquad (5.1)$$

Where r_t represents return on oil prices, and x_t a fundamental factor affecting prices. It is notable that x_t *(up)*, and x_t *(down)* indicate a positive and a negative change in x_t. Intuitively, suppose x_t is the level of inventories as a percentage of storage capacities in the Cushing, Oklahoma where West Texas Intermediate (WTI) is priced at, and it is 85% full. The equation (5.1) implies that if *x*t moves up from 85% to say *90%*, we would not have the same change in r_t, if *x*t had moved down to say *80%*. Furthermore, as discussed about refinery margins, the magnitudes of other factors are very important in determining the size of the next move in crude oil prices. That is,

$$\frac{\partial r_t}{\partial {x^a}_1} \neq \frac{\partial r_t}{\partial {x^b}_1} \qquad (5.2)$$

Here both x^a and $x^b \in \{X\}$, however, they refer to two different starting points *a, b*. *X* is the set of all factors affecting crude prices. To put (5.2) in perspective, the emphasis in this case is on the initial starting point of a factor affecting crude oil prices. For example, suppose x_t is the average refinery utilization in the US. The equation (5.2) indicates that the effect of a similar change in refinery utilization does not necessarily have the same impact on crude oil prices for initial points x^a =80% and x^b =75%. In fact, (5.1) highlights the asymmetric response of crude oil prices to a similar upward or downward move in a specific factor. However, (5.2) highlights an unequal impact of factors on crude oil prices if we had started from two different starting points. In the next section, a two stage optimization setup will be introduced

5.1 Two-Factor Model Augmented with Crude Oil Storage Levels and Capacity

So far we provided a practical approach in order to value a crude oil future contract given crude oil spot prices. The framework that was set up in previous section provide a very useful way of setting up a relationship between spot and forward contracts in a risk neutral framework however, it makes it really hard to see how some change in a structural factor in crude oil market such as crude oil refinery outages and hence change in level of draws could change the value of futures or futures spreads. In this chapter we use the framework to get fair value for crude forward curve as a function of level of inventories and injection withdraw from storages. In this section we utilize refinery and crude oil inputs and outputs and propose a practical setup for market participants to calculate fair value of crude oil futures and future spreads. We discuss in detail how the inventory data should be constructed, and use it to get fair value of crude oil contracts. We discuss errors and out of sample results for futures and futures spreads. Our goal in this section is to build a bridge between risk-neutrality and structure of the crude oil markets. This is an improvement to what we have done in previous section. We first present a set of structural models to show how the dynamics of crude market works. We use the framework for our out-of-sample results and error analysis.

Our aim is not forecasting fair value of contracts at all. But we want to give an answer to a set of key questions every producer deals with in any given day. What are most likely outcomes for fair value of crude oil futures contracts? And how and by how much they could change if a physical factor in crude oil market changes? As two corollaries, we also calibrate spot crude oil prices, along with refinery crack spreads. We report the out of sample results for these two models as well.

As discussed in the Chapter 3, the explicit risk-neutral relationship for a future contract in a two-factor framework can be written as

$$F(S,\delta,T) = S\exp[-\delta\frac{1-e^{-\theta T}}{\theta} + A(T)] \tag{5.3}$$

where

$$A(T) = (r - \alpha^* + \frac{1}{2}\frac{\gamma^2}{\theta^2} - \frac{\gamma\sigma}{\theta}) + \frac{1}{4}\gamma^2\frac{1-e^{-2\theta T}}{\theta^3} + (\alpha^*\theta + \gamma\sigma\rho - \frac{\gamma^2}{\theta})\frac{1-e^{-\theta T}}{\theta^2}$$

In fact, the equation (5.3) is a nonlinear function of time to expiry of a contract, volatility of the spot crude oil, interest rate, convenience yield, and a set of parameters. We can write (5.3) as an implicit function of these four factors

$$F_{tj}(S_t, \delta_{tj}, T_{tj}) = S_t z_j(\sigma_j, \delta_{tj}, r_j, T_{tj}) \tag{5.4}$$

where F_{tj} is the future contract of *j* at calendar date *t*, σ_j volatility of spot prices, r_j risk-free interest rate for contract *j*, δ_{tj} convenience yield of the contract at time t, T_{tj} time to expiry of the contract at time *t*, and z_j is some nonlinear function volatility, risk-free interest rate, and time to expiry for contract *j*. Based on theory of storage as discussed by Gao *et al* (2005) we express volatility as a function of storage full-ness and level of inventories of crude oil. In particular, we use storage full-ness in the US, V_{tj}, storage fullness at Cushing of Oklahoma, C_{tj}, and also deviation of draws from 5-year averages, D_{tj}, as determinants of volatility of crude oil spot prices, $\sigma_j(V_{tj}, D_{tj}, C_{tj})$, and convenience yield, $\delta_{tj}(V_{tj}, D_{tj}, C_{tj})$. In this case, the equation (5.4) can be written as follows:

$$F_{tj}(S_t, T_{tj}) = S_t g_j(V_{tj}, D_{tj}, C_{tj}, r_j, T_{tj}) \tag{5.5}$$

where g_j is some nonlinear function to be approximate. Before specifying a functional form for (5.5) we have to discuss how we should calculate V, D and C, in order to do sensitivity and out of sample analysis.

5.2 Crude Oil Price Determination

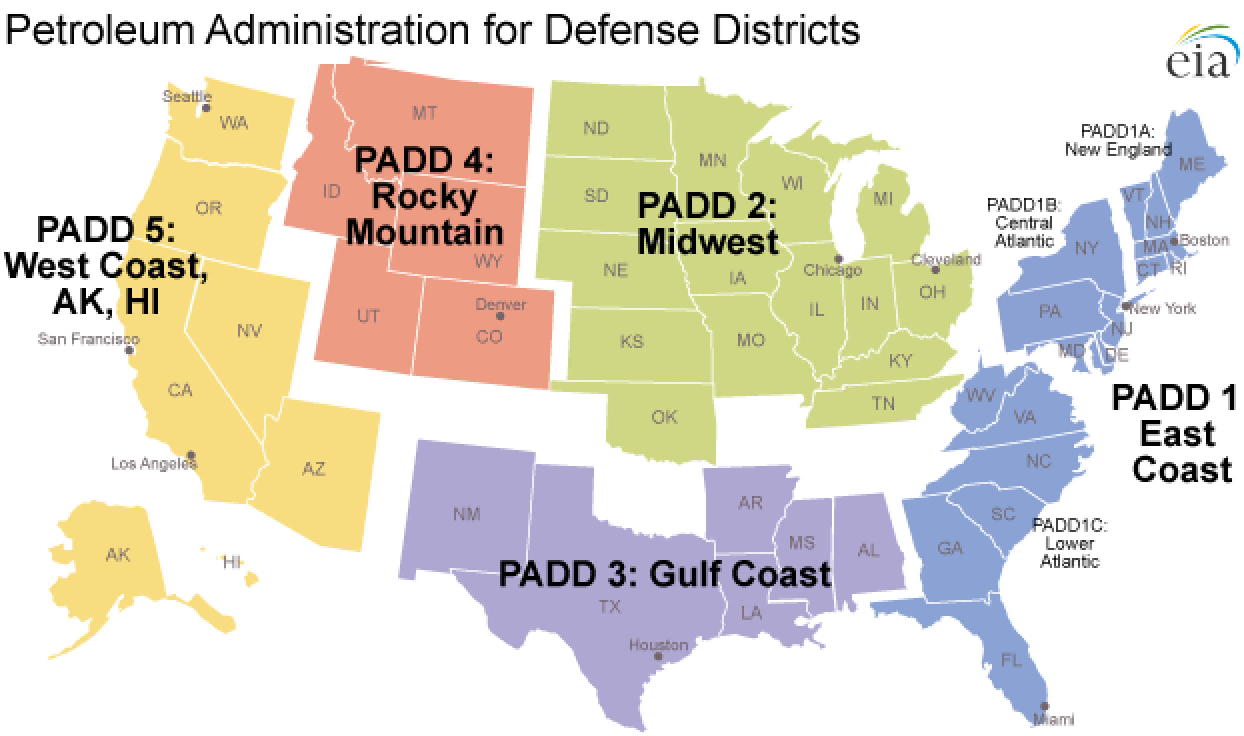
In this section we provide a framework on how crude oil prices move to balance physical crude oil market, we use the framework to lay out required steps to calculate values of required variables in order to calibrate equation (5.5) and do sensitivity and scenario analysis.

Crude oil prices are determined by interaction of futures markets with crude oil producers, refineries, and storage operators. Crude oil producers keep an eye on futures markets and made decisions on production volume depending on current and forward prices. Refineries who are responsible to refine unusable crude oil to refined products such as Gasoline, Distillate, Residual Oil, and Kerosene type jet fuel, also operate by directly interacting with crude oil and refined products forward prices. They use crude oil as input and produce refined products as outputs. They adjust their operations based on the difference between crude oil and refined product prices, known as refinery margin. Refineries cannot exactly match their level of activities with level of demand, and consequently crude oil producers also won't be able to produce according to level of activities of refineries. This is why storage operators come into play as key balancing component of crude oil and refined product markets. Refineries usually try to have an idea of say September driving season and so they keep producing more gasoline and build their gasoline inventories vs. other refined products. Before winter though, they try to build up their heating oil inventories. Storage operators charge incrementally more as level of inventories go up. In fact, as inventories build up the remaining space becomes exponentially valuable which of course should be reflected in month-on-month spreads in crude oil and refined products forward markets as discussed in section the first section.

It is notable that given significant disparity of crude oil output and availability of refineries across the US we use the PADDs division[4] in order to avoid unreasonable aggregation and also get better understanding of physical crude and refined products markets in the US. This is very crucial to modeling crude oil markets. For example, around 50% of the US refining capacity is in Gulf Coast known as PADD II region, while most of the demand for refined products comes from the East Coast, PADD I, with less than 10% refined capacity in the region. Also over 50% of US crude oil production is produced in PADD III. Given The benchmark

[4] - Petroleum Administration for Defense Districts (PADDs) were created during World War II to help allocate fuels derived from petroleum products

pricing point for NYMEX crude oil contracts are in Cushing\Oklahoma, it is important to break the US crude market down into these 5 regions.



Source: U.S. Energy Information Administration

Our structural framework of crude oil markets is setup as a two-stage optimization problem. At stage I, refinery operators and also crude oil producers optimize their operations by maximizing their profits. At stage II "*the investor*" or a risk-averse financial trader of crude oil futures contracts sets up the crude prices at time t for each delivery month j, $P^j(t)$ so that the US market balances out. We will discuss the stage II in detail. The refinery problem with n refined products becomes:

Refinery i: $\max_{f_{i1}^j,\ldots,f_{in}^j}\left(\alpha_i + p_1 f_{i1}^j + \ldots + p_n f_{in}^j - P_i^j u_i^j(f_{i1}^j,\ldots,f_{in}^j)\right)$ (5.6)

$$s.t. \quad f_{i1}^j(t) \le C_i^j(t) - g_i^E(t) - g_i^U(t) \quad (5.7)$$

with the US nation-wide "*aggregate*" constraint of

$$\sum_k f_{ik}^j + nm_i^j = \sum_k u_{ik}^j - d_i^j \quad (5.8)$$

where

$$d_i^j(t+k) = v_i^j(t+k) - v_i^j(t) \quad (5.9)$$

Where α_i is the non-crude oil costs of refinery *i*, f_{ik}^j is refined product *k* produced by refinery *i* for delivery month of *j*, $u_i^j(f_{i1}^j,\ldots,f_{in}^j)$ is the total crude utilization of refinery *i* at time *j* which is a function of each refined product, P_i^j is price of crude oil for refinery *i* for the delivery month of *j*, $C_i^j(t)$ is total name-place capacity of refinery *i* at time *t* for the delivery month of *j*, $g_i^E(t)$ is the planned or expected refinery maintenance, $g_i^U(t)$ is the unexpected outages, $d_j^i(t)$ is current draw of refined product *i* from storages as defined in (5.9), and $v_i^j(t)$ is the inventory level of refined product *i*. At this stage, refinery *i* maximizes its profit (5.6) subject to capacity constraint (5.7). The constraint ensures that total production of each refined product is not greater than the nameplate capacity of refinery after adjusting for planned and unplanned maintenances. The equality constraint (5.8) imposes the "*aggregation*" constraint, in a sense that total utilization of each refined product, $\sum_k u_{ik}^j$, including injection into the in storages, $-d_i^j$, should be equal to sum of total refined product *i* produced in the US plus net import to the country. Solving the maximization problem (5.6) results in optimal level of f_{ik}^j 's as some functions of all fuels and crude oil prices,

$$f_{ik}^{*j} = \gamma(p_1^*,..., p_n^*, p^{*j}, C_i^{*j}(t)) \tag{5.10}$$

At stage I, in addition to refineries crude oil producers also choose to produce optimal level of crude oil given crude oil prices,

$$q_i^{*j} = \chi(P^{*j}, k_1 ..., k_n) \tag{5.11}$$

The (5.11) is a mapping between crude prices and producer *i*'s profit maximizing level of crude oil production, and cost of other inputs, k_i*'s*, in the production process. It is notable that a producer's optimal level of production is only determined by crude oil and input prices. Other facts such as prices of all refined products only impact equilibrium level of crude oil prices.

At stage II "*an investor*" *i* or a risk-averse financial trader of crude oil futures contracts takes positions on the futures market if risk-reward is attractive to her. The decision the investor makes is under conditions of uncertainty. The investor has to allocate her capital K, between her margin account and crude oil contracts, F_i^j . The rate of return of the margin account, *a*, is 0 and return on crude oil future contracts is a random variable of $r = dF^j / F^j$. Intuitively if the investor expects $r < 0$ she will short the crude oil future contract *j*, and if she expects $r > 0$ she will buy the contract and go long. The investor has to maximize her expected Von Neumann-Morgenstern utility function

$$\textit{Investor } i\text{:} \qquad \max_{F^j} E\left[u(F^j(t+k) - F^j(t))b_i\right], \qquad k, j = 1,2,... \qquad (5.12)$$

where $b_i > 0$ indicates a short position and $b_i < 0$ indicates a long position. By using a linear mean-variance utility function, the investor i's maximization problem becomes

$$E\left[(F^j(j) \mid F^j(t))b_i\right] - \rho \frac{1}{2} \operatorname{var}\left[(F^j(j) \mid F^j(t))b_i\right] \qquad (5.13)$$

where *j*>*t*. By differentiating (5.13) with respect to b_i we have

$$b_i = \frac{F^j(t) - E\left[F^j(j) \mid F^j(t)\right]}{\rho \operatorname{var}\left[F^j(j) \mid F^j(t)\right]} \qquad (5.14)$$

This implies that the speculator takes a short bet on the futures crude oil market if and only if the future prices at time t, $F^j(t)$, is higher than her expectation of the value of the future price when it trades in spot market, $E\left[F^j(j) \mid F^j(t)\right]$. The speculator goes long, $b_i < 0$, vice versa.

Now the US nation-wide "*aggregate*" constraint of physical balance in the crude oil market can be written as

$$Q^j(F^j, k_1,...,k_n) + NM^j(t) = \sum_i^I \sum_k^K u_{ik}^j(F^j, k_1,...,k_n) - D^j(t) \qquad (5.15)$$

where K is the total number of refined products, and I total number of refiners. The constraint (5.15) is also the equilibrium conditions of the crude oil market in the US. The equality (5.15) indicates that the sum of all net imports and total crude oil produced by all producers in the US should be equal to the sum of all usage of crude oil by all refineries and net injection, $-D^j(t)$, into storages. Based on (5.15) the dynamic adjustment equation of crude oil market in the US can be written as (See Takayama, A., for similar setup, 1993)[5]:

$$\frac{dF^j}{dt} = h\left[\sum_i^I \sum_k^K u_{ik}^j(F^j, k_1,...,k_n) - D^j(t) - Q^j(F^j, k_1,...,k_n) - NM^j(t)\right] \qquad (5.16)$$

where *h* is a positive constant signifying the speed of adjustment of the physical crude oil market in order to balance, $Q^j(t)$ is total crude production at time *t*, $D^j(t)$ is current draw of crude from storage as defined in (27), $\sum_i \sum_k u_{ik}^j(F^j, k_1,...,k_n)$ is the total utilization of crude oil at time *t* for month *j*, $NM^j(t)$ is net crude oil imports to the US, and

$$D^j(t+k) = V^j(t+k) - V^j(t) \qquad (5.17)$$

Hence speculator i in (5.14) decides to short crude oil futures contract, $b_i > 0$, if she expects that based on (5.16), $\frac{dF^j}{dt} < 0$ in the future. Similarly, she will go long if $\frac{dF^j}{dt} > 0$.

Based on theory of storage (See for example Gao et al (2005)) we can express crude oil futures spreads as a function of level of crude oil inventories relative to storage capacities, $S^j(t)$,

$$F^j(t) - F^{j+k}(t) = f\left(\frac{V^j(t)}{S^j(t)}, \frac{V^{j+k}(t)}{S^{j+k}(t)}\right) \tag{5.18}$$

The equation (5.18) determines the relationships between forward contracts as a function of remaining capacity of storage facilities to store crude oil, in percentage term. The LHS of (5.18) is the amount storage operators get paid to keep crude in storage between month j and month $j+k$, and of course it becomes exponentially valuable as remaining capacity goes. f is a nonlinear function to be approximated. We use a polynomial functional form to approximate this function.

5.3 Empirical Models and Calibrations

The maximization problem (5.18) implies that the optimal level of f_{ik}^j 's are some function of fuel and crude oil prices,

---

[5] - We assume that market forces kicks as t increases or decreases and ultimately makes sure that this equation to converge to a fixed point.

$$f_{ik}^{*j} = \gamma(p_1^*,..., p_n^*, p^{*j}, C_i^{*j}(t)) \tag{5.19}$$

For $k = 1, 2, \ldots, n$, where *'s refer to optimal level of variables. γ is a nonlinear function to be approximated. This requires knowledge of cost structure of every refinery in the US. Given availability of data, which is not publicly available. We take the following steps to solve this two stage problem. To do so, we focus on Gasoline (GSL), Residual Fuel (RSD), Distillate (DST), and Kerosene Type Jet F.uel (KRS), and other fuels (OTH) such as Propane and Ethylene. We take 15 steps to calculate required variables and then get fair values of spot, and future crude oil prices, along with crack spread which will be defined in later:

1. Simulate refinery outages for $g_i^E(t) + g_i^U(t)$ for *i= GSL, RSD, DST, KRS, and OTH* for PADD I – PADD IV regions

2. Calculate capacity $C_i^{*j}(t)$ for *i= GSL, RSD, DST, KRS, and OTH* for PADD I – PADD IV regions

3. Calculate equilibrium level of f_{ik}^{j} for *i= GSL, RSD, DST, KRS, and OTH* for PADD I – PADD IV regions

4. Apply (1) to (2) to get refinery output in each PADD region for for *i= GSL, RSD, DST, KRS, and OTH*

5. Calculate net fuel import by each region for *i= GSL, RSD, DST, KRS, and OTH*

6. Define fuel draws from storage as $d_i^j(t)$ = (3) + (5) – (4) = $f_i^j(t) + m_i^j(t) - u_i^j(t)$

7. Calculate 5-year average fuel draws from storage, $\bar{d}_i^j(t)$, as a benchmark for a gauge of market expectation of optimal fuel draws

8. Calculate 5-year average fuel stocks, $\bar{s}_i^j(t)$, for all regions and all fuels

9. **As a corollary**, calibrate crack spreads, cs_i for *i= GSL, RSD, DST, KRS, and OTH*

$$cr_i = \eta(d_i(t) - \bar{d}_i(t), s_i(t) - \bar{s}_i(t), c_i^j(t)) \tag{5.20}$$

10. Calculate crude utilization factor, $\alpha_i^j = f_i^j(t)/u_i^j(t)$, for *i= GSL, RSD, DST, KRS, and OTH* for all 5 PADD regions

11. Given refinery output in item (4) and utilization factors of each fuel in (10) calculate crude inputs to refineries, $q^j(t)$, for all 5 regions

12. Given crude oil export from the US is prohibited by law, calculate crude oil draws from storages as $D_i^j(t) = Q^j(t) - q^j(t) - NM^j(t)$

13. Calculate 5-year average of crude oil draws from storage, $\bar{D}_i^j(t)$, as a benchmark for a gauge of market expectation of optimal crude draws

14. Calculate 5-year average crude oil inventory levels, $\bar{S}_i^j(t)$, for all regions

15. **As a corollary**, calibrate crude oil prices, P^j, as

$$P^j = \psi(D^j(t) - \overline{D}^j(t), V^j(t) - \overline{V}^j(t), c(t)) \tag{5.21}$$

and combine it with two-factor Q-measure future pricing methods discussed in Chapter 3 to build a bridge between crude oil physical variables and method of pricing crude futures contracts.

In the next section we use above steps to construct the data and then calibrate our empirical model of futures and future spreads. We also use the above algorithm to construct our variables and build out of sample values of them.

5.3.1 Data

We use weekly and monthly Petroleum & Other Liquids data sets published by the US Energy Information Administration (EIA) from 1990. The monthly data sets cover through January 31, 2015 with 300 observations but weekly data sets go through April 10, 2015 and have around 1300 observations. Our models are based on weekly data sets however we use monthly data sets if the EIA has not yet published the weekly data sets on a regional basis. In these cases we use monthly data sets to get a weekly estimate of the numbers, which is discussed in detail case by case. In terms of geographical aggregation of data in the US, it is notable that both monthly and weekly data sets are reported in 5 PADD regions that are used in this research as well. Our weekly data sets include domestic crude oil production for the US not necessarily by PADD region. Crude oil delivered to refineries, refinery utilization and their capacities in each region are reported by PADD regions. We got the refinery outages from Bloomberg which only goes back to 2011 but is reported on a daily basis and classified by fuel and PADD region. Petroleum products which include, liquefied petroleum gases, pentanes plus, aviation gasoline, motor gasoline, naphtha-type jet fuel, kerosene-type jet fuel, kerosene, distillate fuel oil, residual fuel oil, petrochemical feedstocks, special naphthas, lubricants, waxes, petroleum coke, asphalt, road oil, and still gas, are reported for each region but in this research we classify them into 5-groups for all 5 PADD regions as Total Gasoline (GSR), Kerosene-Type Jet Fuel (KRS), Distillate Fuel Oil (DST), Residual Fuel Oil (RSD), and Other Fuels (OTH). We use the same classification for imports to the US, exports from the US, draws from storage facilities, and level of stocks for the US and all 5 PADD regions. In terms of units, all data are converted to 1000 barrels (MB)/day unless stated otherwise.

We use the data and illustrate our approach with a numerical example for the week ending on 2015-05-08. **At step 1** we should simulate refinery outages. Refinery outages mostly happen because of planned shutdown, maintenance or planned upgrades of some refinery units. However, it is not necessarily limited to planned outages. Factors such as mechanical failures, weather, power failure or any other natural disaster like Hurricanes can also cause refinery outages. Obviously PADD III has more outages than any other regions because over 50% of refinery capacity in the US is located in that region. We use weekly data and build the following confidence interval for outages for the 5 PADDs regions

$$\bar{g}_i^{\,j}(t) - 1.96\left(\frac{\sigma_{w|m}}{\sqrt{n}}\right)^j \leq g_i^{\,j}(t+s) \leq \bar{g}_i^{\,j}(t) + 1.96\left(\frac{\sigma_{w|m}}{\sqrt{n}}\right)^j \qquad (5.22)$$

for i= *GSL, RSD, DST, KRS, and OTH, and j*= PADD I,…, PADD IV, where $\bar{g}_i^{\,j}(t)$ is the average of total historical outages, $\bar{g}_i\,(t) = \bar{g}_i^{E}(t) + \bar{g}_i^{U}(t)$, for the week ending at time *t*, and $\sigma_{w|m}$ the weekly standard deviation of outages in a particular month. The Figure 5.1 shows the historical data along with 95% confidence interval-based simulation going forward.

Figure 5.1: Simulated and Actual Daily Refinery Outages in 5 PADD Regions in the US

(Data Source: Bloomberg)

At step 2 we simply calculate refinery Capacity in All PADDs by adjusting operable capacity of refineries with expected outages, $C_i^j(t) - g_i^j(t)$, and for future dates $C_i^j(t+s) = C_i^j(t)$. The Figure 5.2 shows the US refinery capacity from 1990 through 2015.

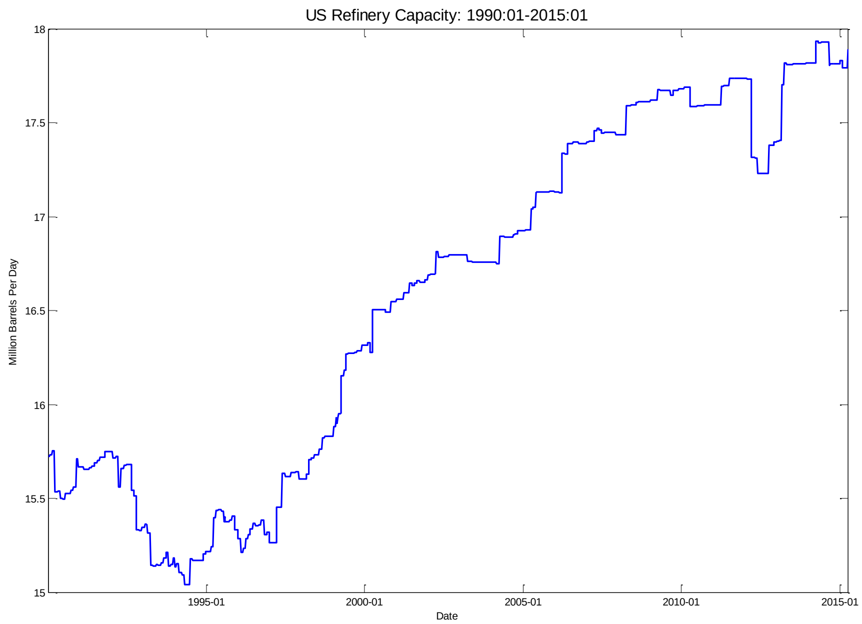


Figure 5.2: The US Refinery Capacity Since 1990 (Data Source: Bloomberg)

It is notable that most of the new additions of the refinery capacity over the last decade has been taking place in PADD III, Gulf Coast region, and the refinery capacity in other regions either have been going down or they have had very small growth compared to PADD III (EIA).

At step 3 we calculate equilibrium level of each fuel in each region. The weekly data only includes aggregated total fuel consumption for the US and each of the 5 regions. However, the Petroleum Monthly Data Sets have fuel data for all regions. For this reason, we use the monthly data and define the following weights:

$$w_i^j(t_m) = \frac{f_i^{\,j}(t_m)}{\sum_{i=1}^{5} f_i^{\,j}(t_m)} \tag{5.23}$$

where t_m indicates monthly time interval. We then use these historical weights, $w_i^j(t_m)$, and approximate weekly consumption of each fuel in each region:

$$f_i^{\,j}(t_w) = w_i^j(t_m)\sum_{i=1}^{5} f_i^{\,j}(t_w) \tag{5.24}$$

where t_w indicates monthly time interval.

At step 4 we use similar approach as above to construct weekly net import of each fuel type by region:

$$w_i^j(t_m) = \frac{ni_i^j(t_m)}{\sum_{i=1}^{5} ni_i^j(t_m)} \tag{5.25}$$

We then use these historical weights, $w_i^j(t)$, and approximate net import of each fuel in each region on a weekly basis:

$$ni_i^j(t_w) = w_i^j(t_m)\sum_{i=1}^{5} ni_i^j(t_w) \qquad (5.26)$$

The Figure 5.3 show the historical total net fuel imports by each region in the US. A significant change taking place in the PADD III where net import has dropped from close to 1 million barrels/day back in 2006 to around -2 million barrels/day, meaning that the region is now net exporter of refined products.

At step 5 we calculate the draws for the 2nd week of say May 2015. $d_i^j(t\ ) = f_i^j(t\ ) + m_i^j(t\ ) - u_i^j(t)$.

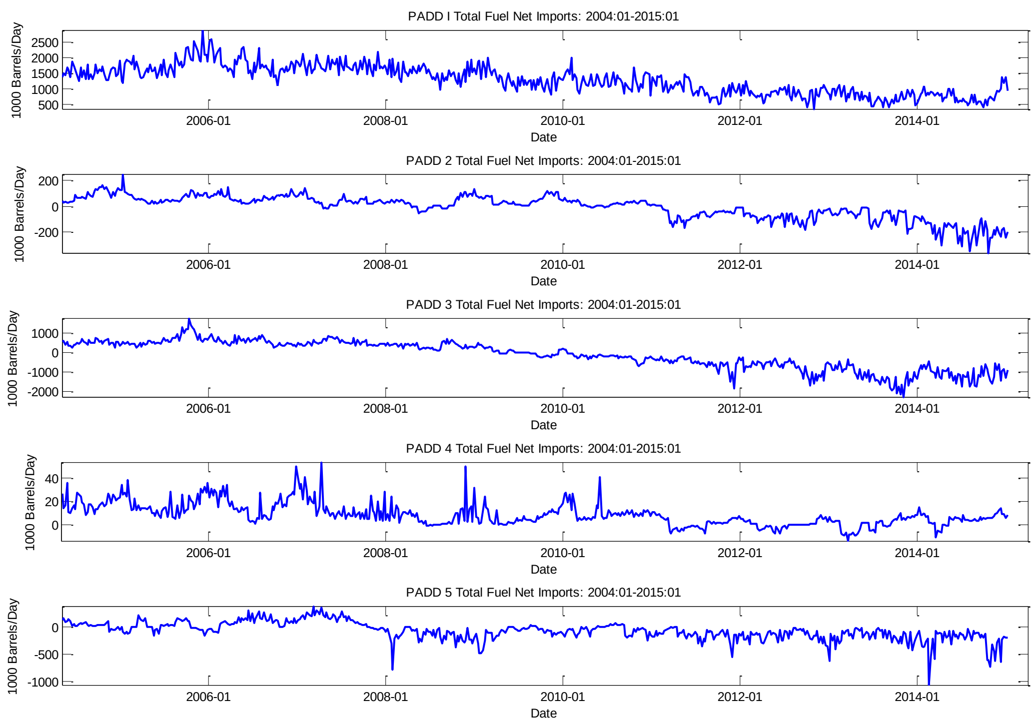


Figure 5.3: Total Net Fuel Imports by PADD region since 2004 (Source: EIA)

5.3.2 Empirical Models of F_{tj}/S_t

We use the following functional form to approximate equation (5.5):

$$\begin{aligned} F_{tj}/S_t = {} & \alpha_j + \beta_{1j}T_{tj} + \beta_{2j}T_{tj}^2 + \beta_{3j}T_{tj}^3 \\ & + \beta_{4j}C_{tj} + \beta_{5j}C_{tj}^2 + \beta_{6j}C_{tj}^3 \\ & + \beta_{7j}V_{tj} + \beta_{8j}V_{tj}^2 \qquad\qquad \textit{for } j = 1,2,\ldots,122 \\ & + \beta_{9j}D_{tj} + \beta_{10j}D_{tj}^2 \\ & + \beta_{11j}r_j \\ & + \varepsilon_{tj} \end{aligned} \tag{5.27}$$

where t is time indicator, T time to expiry, j future contract j-th, C_{tj} level of crude oil inventories relative to design capacity at Cushing of Oklahoma, V_{tj} level of crude oil inventories relative to design capacity in the US, D_{tj} crude oil withdraw from storages relative to 5-year average, r risk free interest rate, and the error term $\varepsilon_{tj} \sim N(0,\sigma^2)$. It is notable that rather than looking at data from calendar date perspective, we map all our data sets from calendar date to Time-To-Expiry format. In this way we avoid making any unreasonable assumption or adjustment when we role from one future contract to next one at the time of expiry. For example, Figure 5.4 shows structure of T_{tj}.

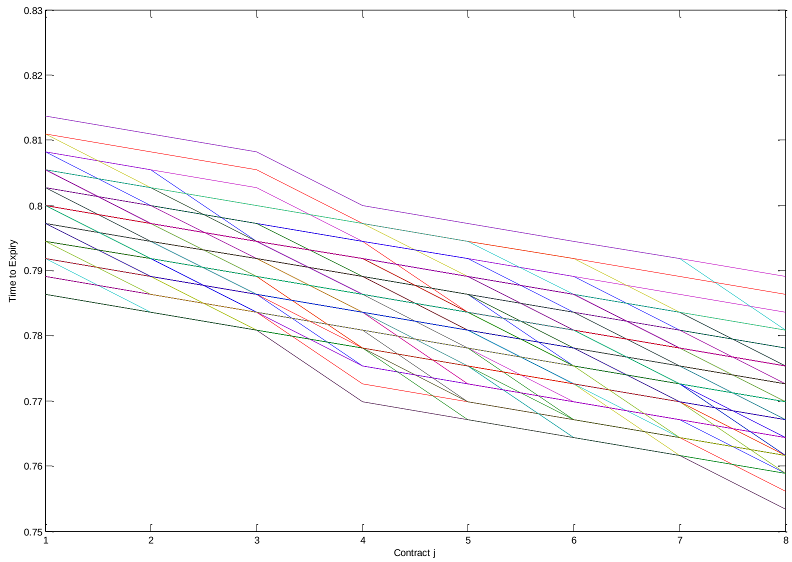


Figure 5.4: Typical Structure of the Constructed Data Sets for T_{tj} Across All Futures

We construct the data sets and line them up in terms of time to expiries of all 122 futures contracts from February 2005 through March 2015 and estimate 11 parameters for each equation. Figure 5.5 shows R2 of all 122 models with average of 76%.

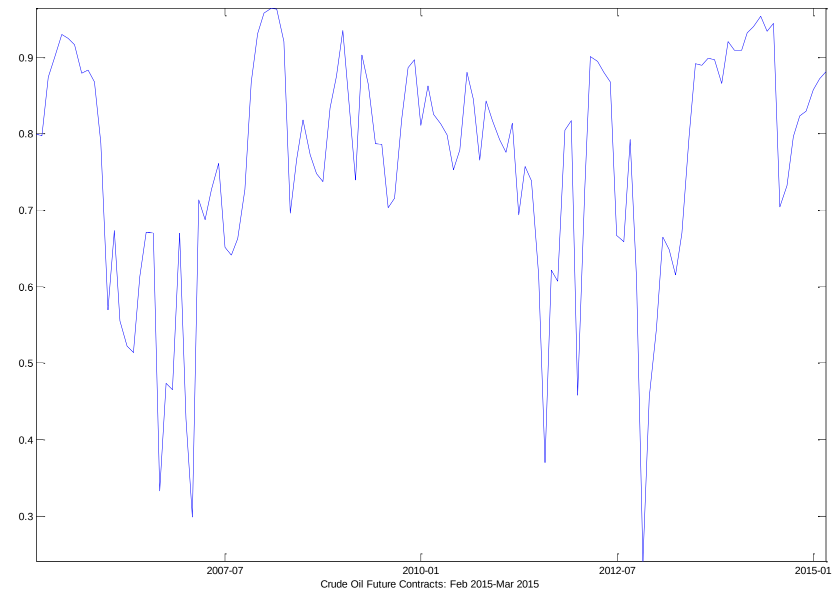


Figure 5.5: R2 of Equation (5.27) across all Futures Contracts

Figure 5.6 shows varying estimated parameters of **Equation (5.27)** across futures contracts. This finding is very important and indicates that it is not reasonable to just run one model across all contracts.

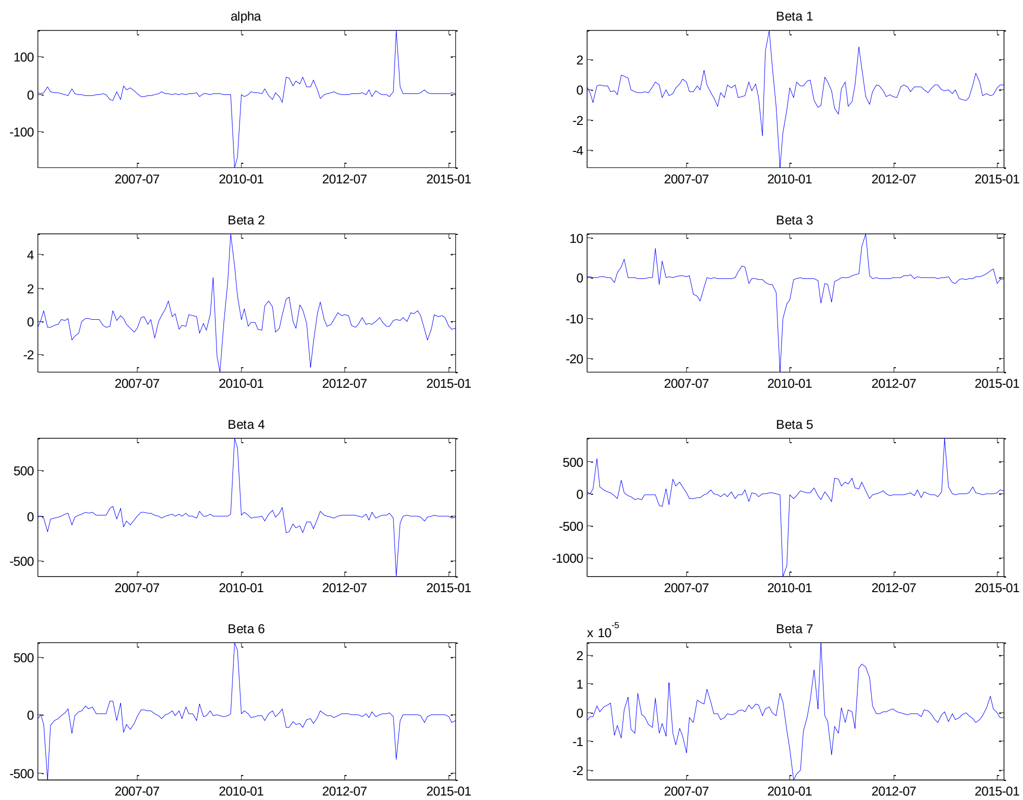


Figure 5.6: Estimated Parameters of Equation (5.27) Across All Futures Contracts

In order to run out-of-sample analysis, we left the last 60 business days of the life of future contracts out and did not use them in the parameter estimation. We then used estimated parameters up to that trading day and compared predicted values to actual values. We did the same for 30 business days in order to get an idea of model performance in terms of number of days out.

Figures 5.7 – 5.17 show out of sample forecasts of every June contract from 2005 through 2014, December 2014, and February 2015 contracts. We report both 60 days and 30 days out of sample forecasts. Because our data sets are in terms of time to expiry, these forecasts are for about 3-month and 1.5 month in terms of calendar days. Our results clearly show that model has much better out-of-sample performance for 30-days out vs. 60-days out.

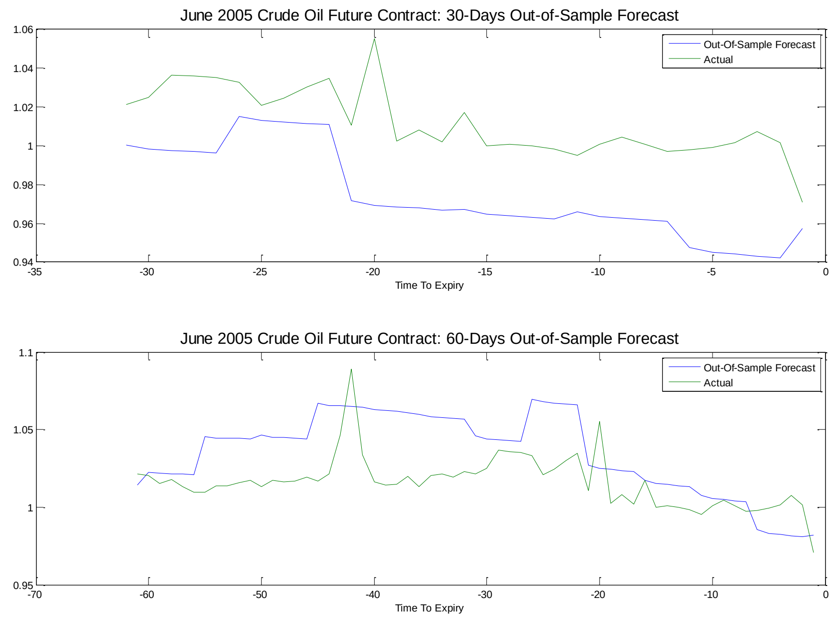


Figure 5.7: Out-of-Sample Forecasts of June 2005 Crude Oil Future

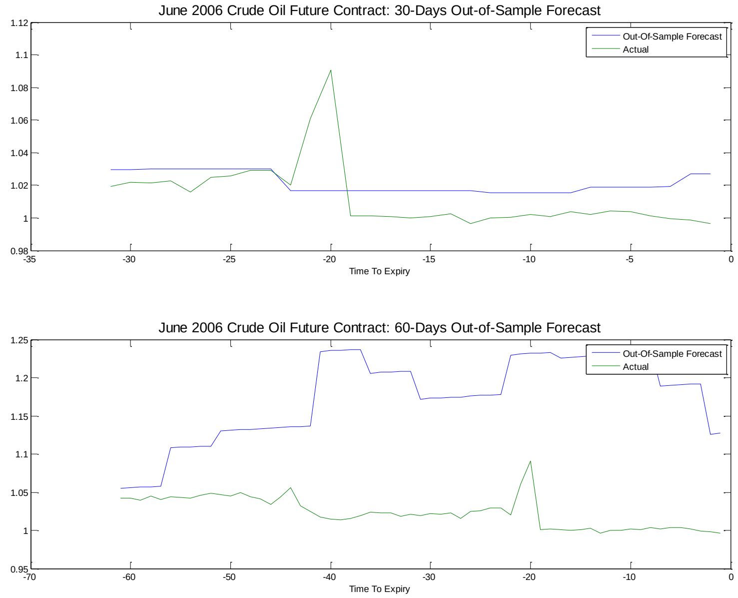


Figure 5.8: Out-of-Sample Forecasts of June 2006 Crude Oil Future

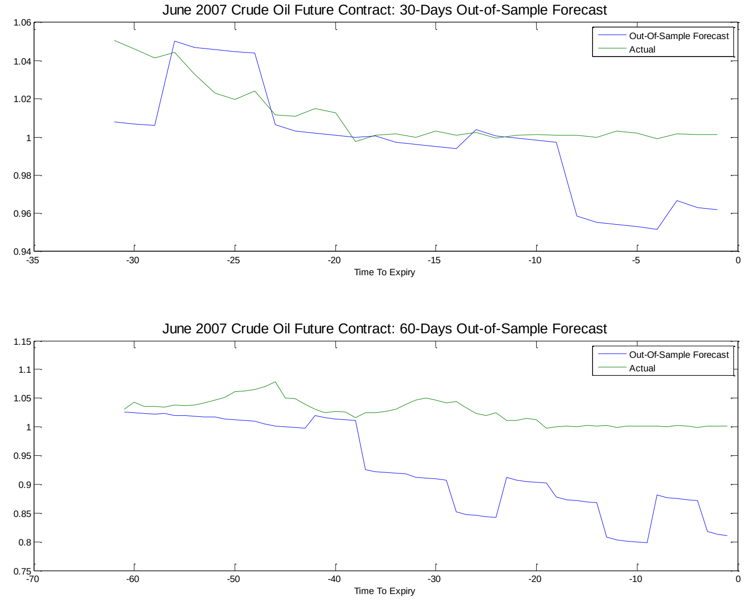


Figure 5.9: Out-of-Sample Forecasts of June 2007 Crude Oil Future

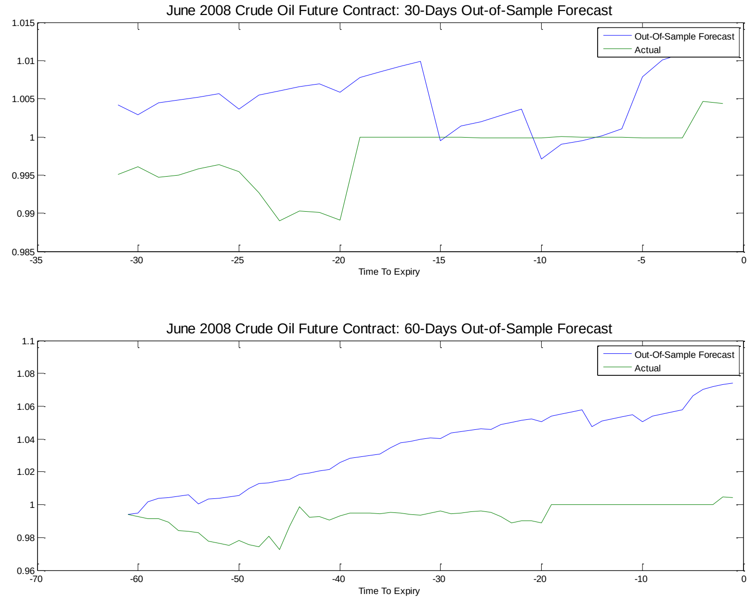


Figure 5.10: Out-of-Sample Forecasts of June 2008 Crude Oil Future

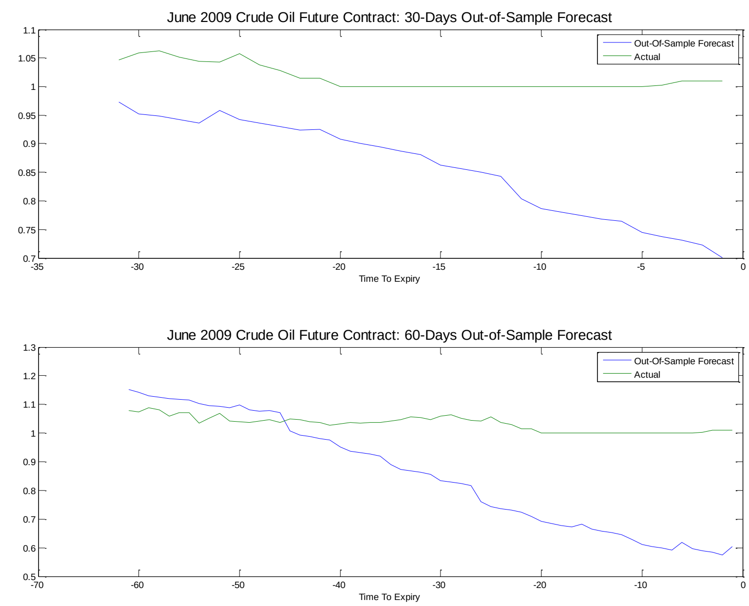


Figure 5.11: Out-of-Sample Forecasts of June 2009 Crude Oil Future

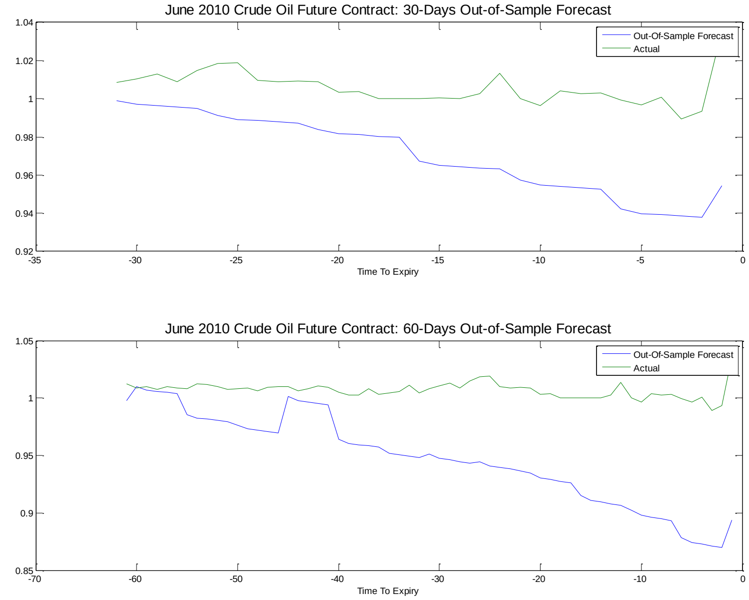


Figure 5.12: Out-of-Sample Forecasts of June 2010 Crude Oil Future

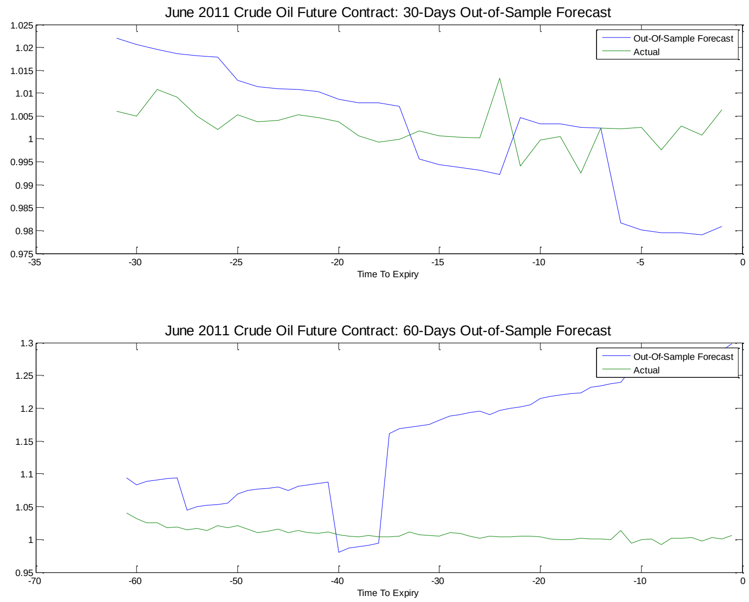


Figure 5.13: Out-of-Sample Forecasts of June 2011 Crude Oil Future

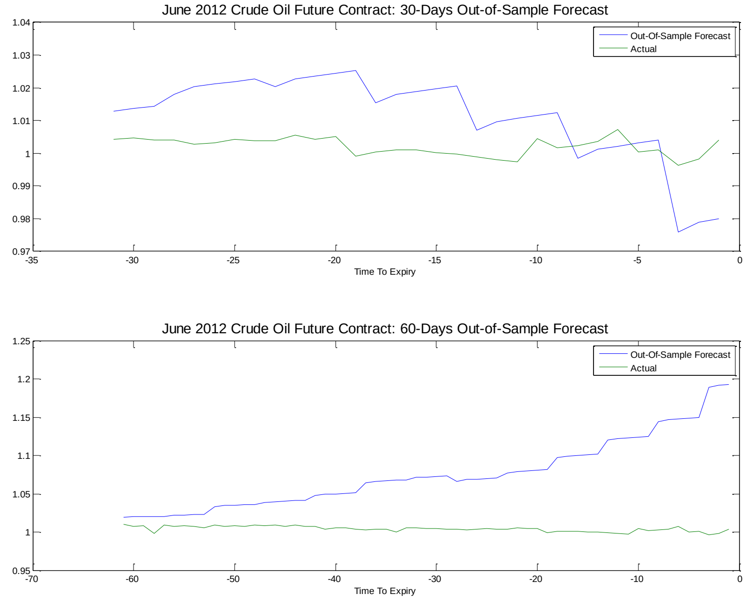


Figure 5.14: Out-of-Sample Forecasts of June 2012 Crude Oil Future

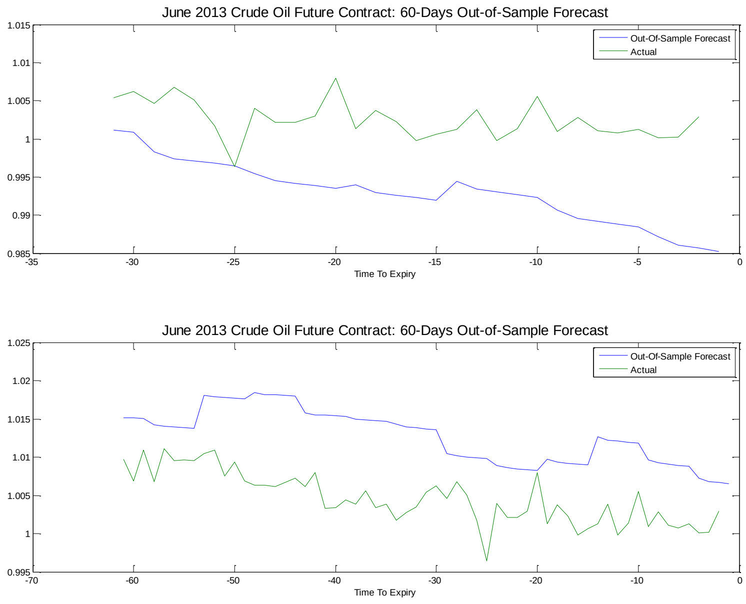


Figure 5.15: Out-of-Sample Forecasts of June 2013 Crude Oil Future

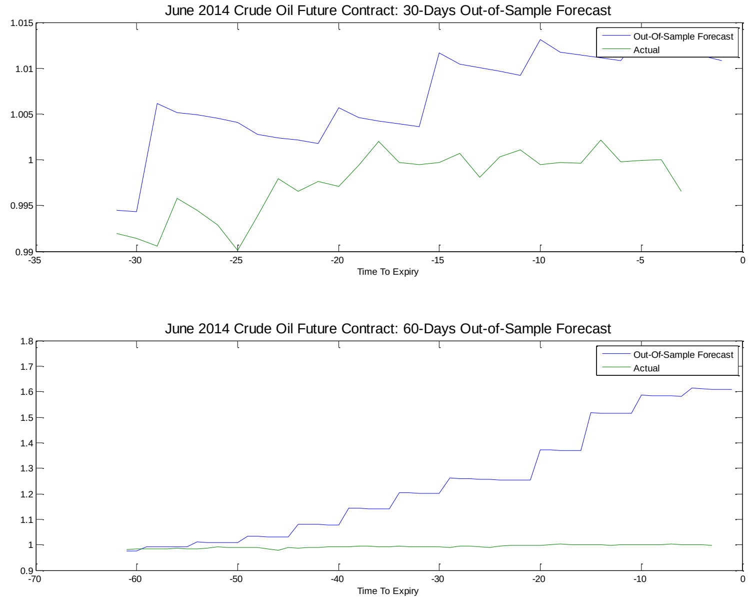


Figure 5.16: Out-of-Sample Forecasts of June 2014 Crude Oil Future

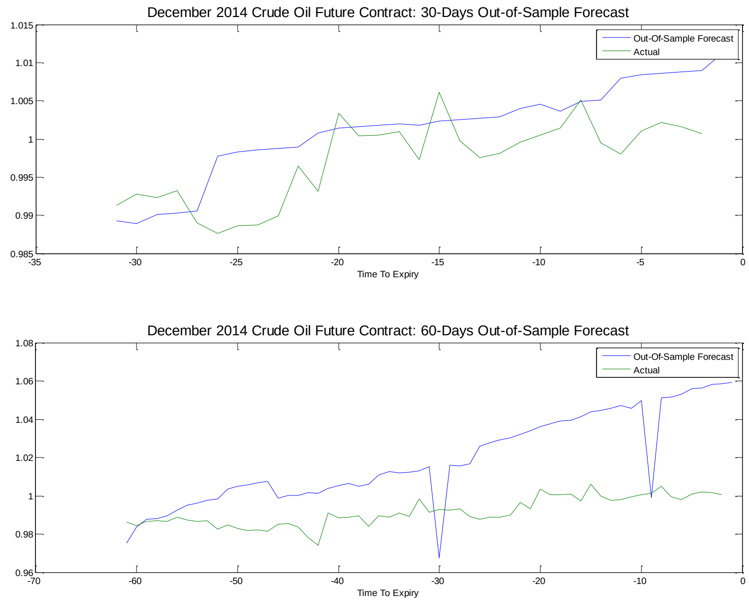


Figure 5.17: Out-of-Sample Forecasts of December 2014 Crude Oil Future

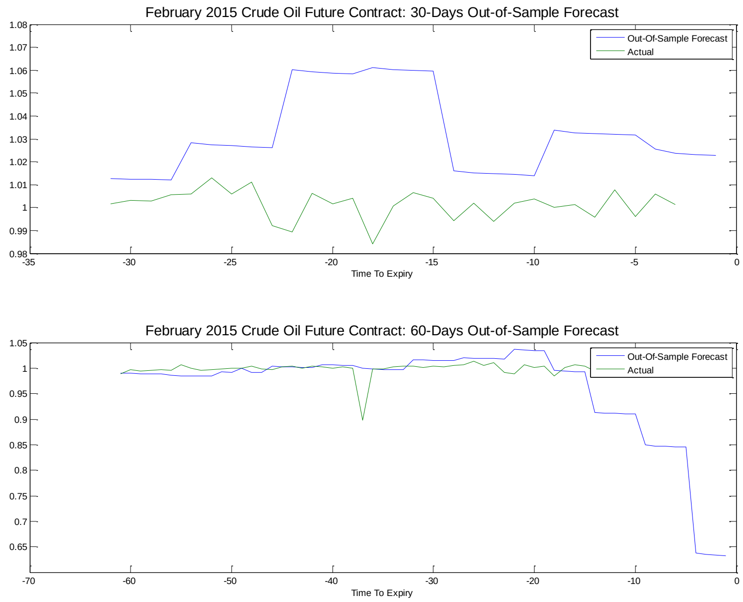


Figure 5.18 Out-of-Sample Forecasts of February 2014 Crude Oil Future

5.3.3 Empirical Models of Future Spreads

To calculate the fair value of future spreads, we use equation (5.5) and define the spread between two contracts as follows

$$\frac{F_{tj} - F_{tk}}{S_t} = \hat{f}_{tj} - \hat{f}_{tk} \qquad , for\ k > j \qquad (5.28)$$

where $\hat{f}_{tj}$ is predicted values of F_{tj}/S_t for j=1,2,…,122, $\forall t$. We switch our reference dates from time to expiry back to calendar date because the same values of time to expiry for contracts j and k do not point to the same trading day.

Figure 5.19 shows out of sample forecast of crude oil future spreads of some selected contracts.

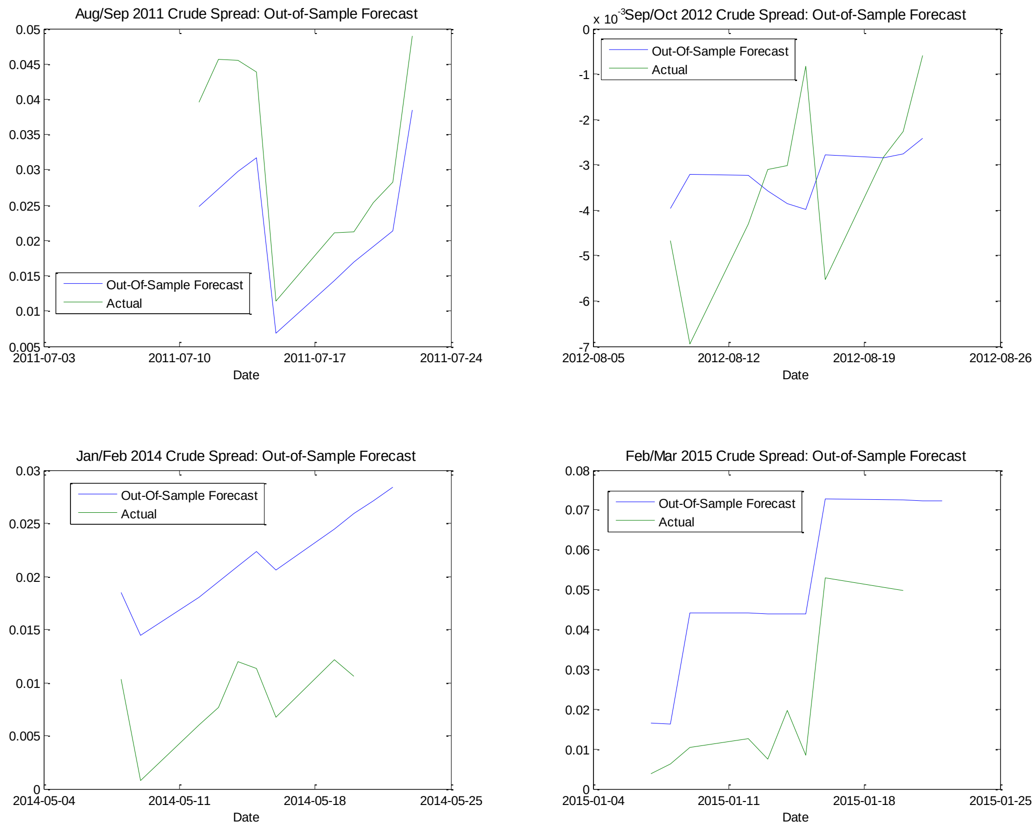


Figure 5.19 Out-of-Sample Forecasts of Crude Oil Future Spreads

For example, our model expected that the values of Feb/Mar 2015 crude oil contract to go up by 5% from early January 2015 toward the end. This is very close to what happened to actual Feb/Mar. The model shows around 2% overestimation on that particular month.

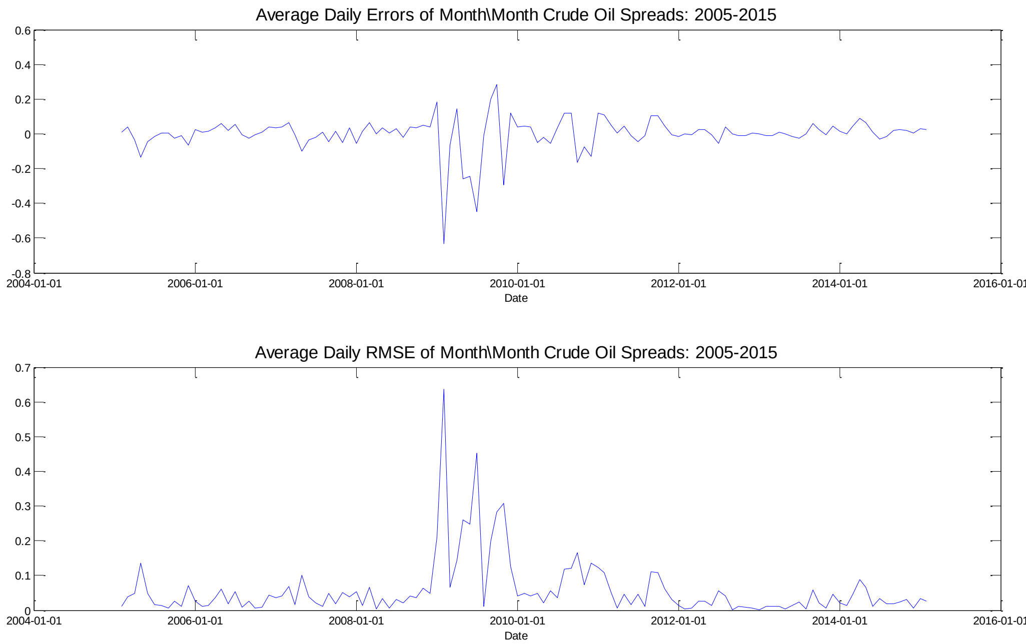


Figure 5.20 Out-of-Sample Forecast Errors and RMSE of Crude Oil Future Spreads

Figure above show average daily errors of future spreads since 2005 over 121 future spread contracts. The average error is less than 1%, however, model does not do well at all in 2009 when crude oil prices made new lows and rebound after then following by the financial crisis of 2008. The Root Mean Square Error (RMSE) of the forecasts are around 0.05 but similar spikes significantly in 2009.

5.3.4 Level of Inventories and Draws Impacts on Futures and Spreads

We use equation (5.5) to perform sensitivity analysis of fair value of crude oil future contracts with respect to change in V, C, and D. To do so we have to use the algorithm outline in the previous section otherwise the sensitivity analyses would be realistic. For example, if we want to evaluate impact of 30% drop (increase) in draws (injection) from (into) storages we might be able to keep C unchanged but the change in D should be accompanied by an equivalent change in V.

Figure 5.21 shows sensitivity analysis of fair value of February 2015 crude oil contract with respect to C, V and D. The first 3 subplots directly measure impact on each variable by keeping other variables unchanged. The last plot does the sensitivity analysis using the algorithm. Our results show that fair value of February contract relative to spot prices go up if level of inventories go down by around 1%.

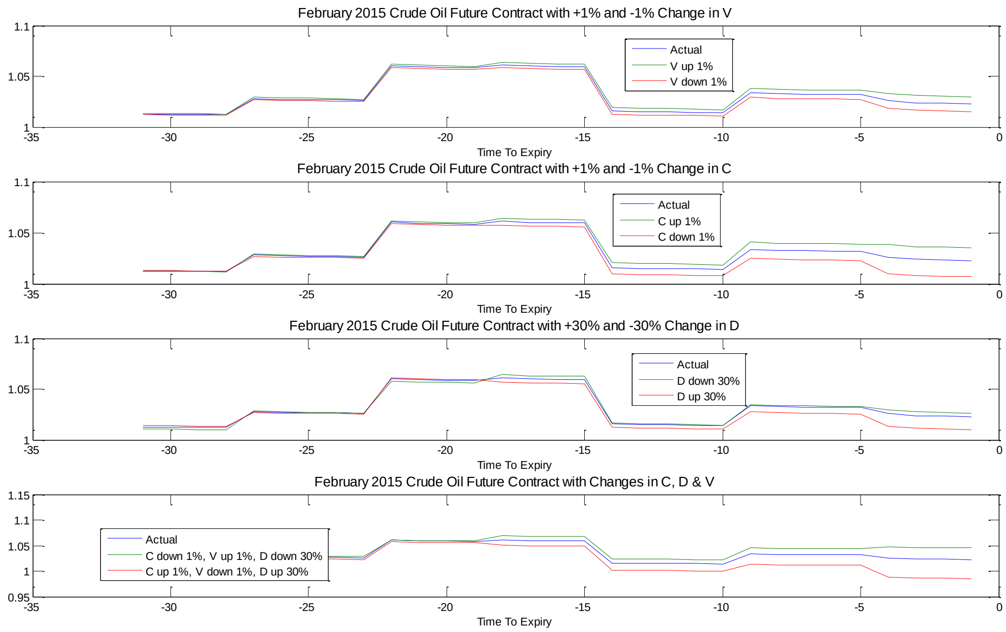


Figure 5.21 Sensitivity Analysis of Crude Oil Contracts w.r.t. C, D and V

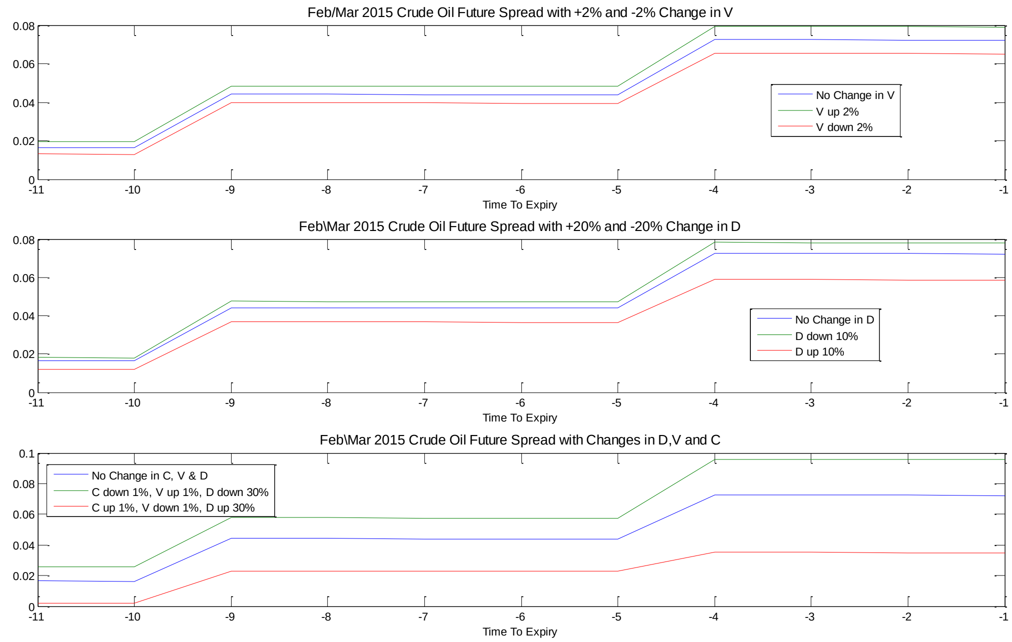


Figure 5.22: Sensitivity Analysis of Crude Oil Contracts on Combinations of C, D and V

In order run a similar sensitivity analysis on crude oil future spreads, we use equation (5.28) and evaluate changes in fair value of crude oil future spreads with respect to change in V, C, and D. Figure above shows the results for Feb/Mar 2015 spread. The third subplot is a realistic scenario generated using the algorithm discussed before. It indicates that spread between Feb and Mar should tighten by around 3% if level of inventories go down by 1% even if level of inventories at Cushing increases by 1%.

5.3.5 Corollary Results

5.3.5.1 Calibrating Crack Spread

Crack spreads (CRS) are the difference between input(s) and output(s) in refinery process. For crude oil refineries, crack spreads are a measure of the difference between crude oil and refined products prices, and are solely based on market prices. Crack spreads can be calculated either for a single refined product or for multiple refined products. The one we use in this research is the 3:2:1 crack spread measuring the price differential between crude oil NYMEX and gasoline and distillate and is calculated as follows:

$$cr_{tj} = [2\frac{42}{100}GSL_{tj} + \frac{42}{100}DST_{tj} - 3F_{tj}]\frac{1}{3} \qquad (5.29)$$

where t is time to expiry, j indicates j^{th} future contract, GSL_{tj} gasoline future prices for the delivery month of j trading at t, DST_{tj} distillate future prices for the delivery month of j trading at t, and F_{tj} crude oil future prices for the delivery month of j trading at t. It is notable that all contracts follow the crude contract, in a sense that if the front month for crude oil is September, the contracts for two other refined products should be based on September contracts as well.

Crack spreads can be positive or negative depending on supply and demand conditions of the refined products markets. In the case of unexpected maintenance in refineries we could expect that refined products rally to incentivize refineries to operate at higher capacity, therefore resulting in relatively higher crack spreads. The opposite usually takes place in an oversupply situation. Figure 5.23 shows the historical data on CRS and crude oil prices. Even though it seems the two sets of prices move in tandem but in fact there is a very weak relationship between the two. For example, the 44-days correlation between the two fluctuates between -1 and +1.

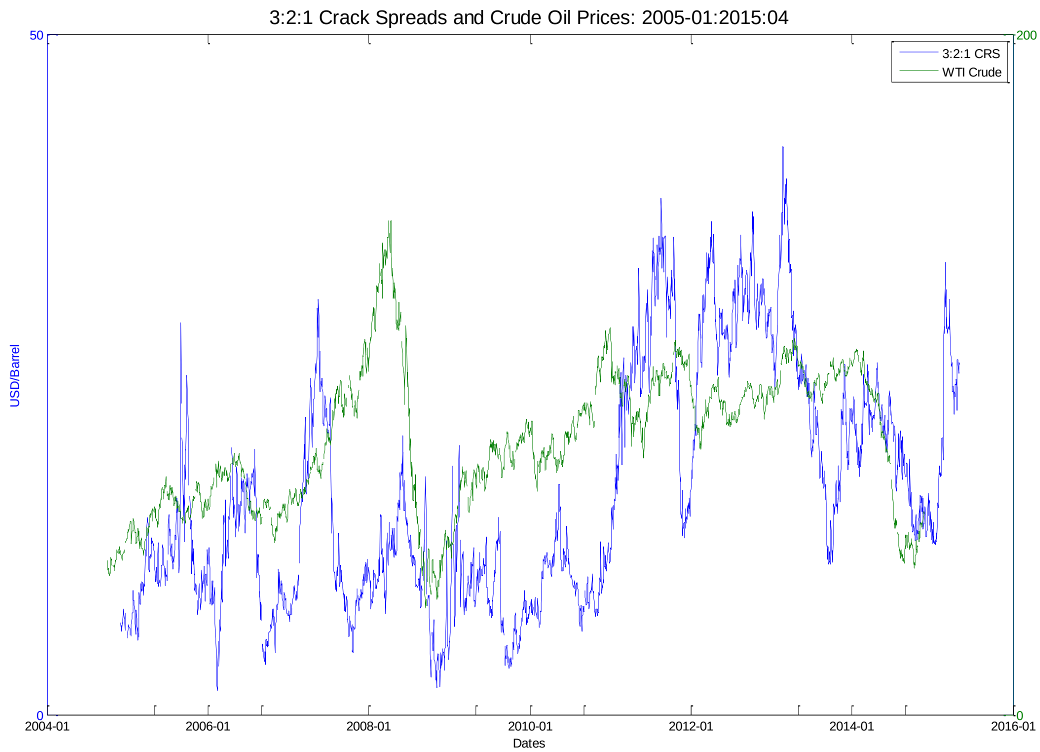


Figure 5.23: Crude Oil and CRS Prices

As specified in equation (5.20), $cr_i = \eta(d_i(t) - \overline{d}_i(t), s_i(t) - \overline{s}_i(t), c_i^j)$, we need to calculate $d_i(t) - \overline{d}_i(t)$ and $s_i(t) - \overline{s}_i(t)$ in order to calibrate equation (5.20). Figure 5.24 shows Stocks of each of the five fuels vs. 5-year averages in the US. Based on steps outlined in section 5.2 we calculate each of these variables and then calibrate crack spread, CRP.

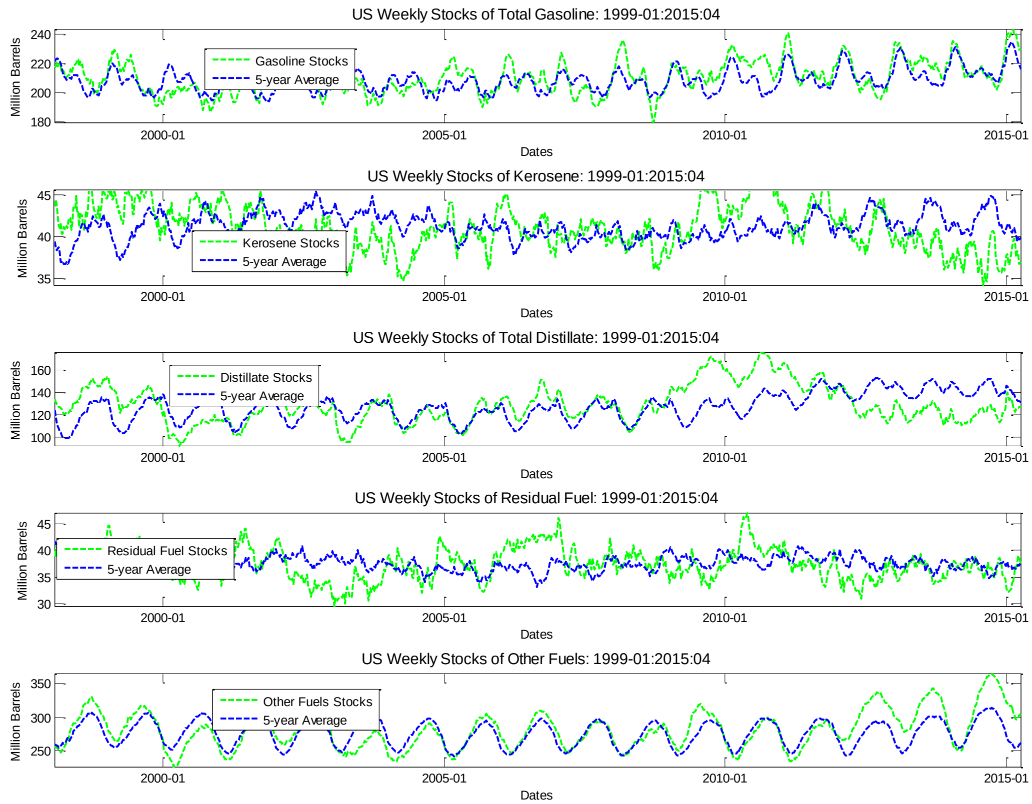


Figure 5.24: Stocks of Fuels vs. 5-year Averages

The empirical version of equation (5.20) is specified as a quadratic function of deviation from 5-year average and operable capacity of refineries in the US:

$$
\begin{aligned}
cr(t) = \beta_0 &+ \beta_1(d_{GSL}(t) - \overline{d}_{GSL}(t)) + \beta_2(d_{GSL}(t) - \overline{d}_{GSL}(t))^2 \\
&+ \beta_3(d_{RSD}(t) - \overline{d}_{RSD}(t)) + \beta_4(d_{RSD}(t) - \overline{d}_{RSD}(t))^2 \\
&+ \beta_5(d_{DST}(t) - \overline{d}_{DST}(t)) + \beta_6(d_{DST}(t) - \overline{d}_{DST}(t))^2 \\
&+ \beta_7(d_{KRS}(t) - \overline{d}_{KRS}(t)) + \beta_8(d_{KRS}(t) - \overline{d}_{KRS}(t))^2 \\
&+ \beta_9(d_{OTH}(t) - \overline{d}_{OTH}(t)) + \beta_{10}(d_{OTH}(t) - \overline{d}_{OTH}(t))^2 \\
&+ \beta_{11}(s_{GSL}(t) - \overline{s}_{GSL}(t)) + \beta_{12}(s_{GSL}(t) - \overline{s}_{GSL}(t))^2 \\
&+ \beta_{13}(s_{RSD}(t) - \overline{s}_{RSD}(t)) + \beta_{14}(s_{RSD}(t) - \overline{s}_{RSD}(t))^2 \\
&+ \beta_{15}(s_{DST}(t) - \overline{s}_{DST}(t)) + \beta_{16}(s_{DST}(t) - \overline{s}_{DST}(t))^2 \\
&+ \beta_{17}(s_{KRS}(t) - \overline{s}_{KRS}(t)) + \beta_{18}(s_{KRS}(t) - \overline{s}_{KRS}(t))^2 \\
&+ \beta_{19}(s_{OTH}(t) - \overline{s}_{OTH}(t)) + \beta_{20}(s_{OTH}(t) - \overline{s}_{OTH}(t))^2 \\
&+ \beta_{21}c(t) + \beta_{22}c^2(t) + \beta_{23}c^3(t) \\
&+ \beta_{24}t + \beta_{25}t^2(t) + \varepsilon(t)
\end{aligned} \qquad (5.30)
$$

where $\varepsilon(t)$ is normally distributed with 0 mean and variance σ^2, t indicates calendar date, d draw from storage, s storage levels, and $\overline{d}$ draw over 5-year averages, $\overline{s}$ storage levels over 5-year average. GSL stands for gasoline, RSD for residual fuel, DST for distillate, KRS for Kerosene, and OTH indicates other fuels. The data frequency is weekly. We employed Ordinary Least Squares (OLS) to estimate the parameters of equation (5.30). The estimated parameters along with their t-statistics are summarized in Table 5.1. The R^2 of the regression is 66%. The F-statistics is 36. It is notable that we performed Augmented Dicky Fuller test (ADF) of existence of unit root in the residuals of the equation (5.30) and the null of existence of unit root is reject at 5% critical level indicating that the equation (5.30) is cointegrated.

Table 5.1: Crack Spread Parameters estimates with t-stats and Pr($\beta_i = 0$) = 0 for all *i*'s.

Param.	**Estimate**	**Std. Error**	**t-Stat**	**Prob**	**Param.**	**Estimate**	**Std. Error**	**t-Stat**	**Prob**
β_0	-397.48	559	-0.71	0.4772	β_{13}	-0.00029	0.000047100	-6.12	0

β_1	-0.00037	0.000294000	-1.27	0.2041	β_{14}	0.00000	0.000000003	2.16	0.0314
β_2	0.00000	0.000000233	-0.63	0.5295	β_{15}	0.00068	0.000131000	5.21	0
β_3	0.00011	0.000107000	1.01	0.3128	β_{16}	0.00000	0.000000022	-7.56	0
β_4	0.00000	0.000000031	-1.13	0.2592	β_{17}	-0.00006	0.000035900	-1.68	0.0939
β_5	-0.00050	0.000142000	-3.52	0.0005	β_{18}	0.00000	0.000000001	3.24	0.0013
β_6	0.00000	0.000000053	-1.99	0.0469	β_{19}	-0.00048	0.000105000	-4.60	0
β_7	-0.00022	0.000111000	-1.94	0.0524	β_{20}	0.00000	0.000000018	2.84	0.0047
β_8	0.00000	0.000000033	-2.07	0.0393	β_{21}	13.94041	20.061860000	0.69	0.4875
β_9	0.00039	0.000230000	1.70	0.0904	β_{22}	-0.15804	0.239430000	-0.66	0.5095
β_{10}	0.00000	0.000000148	-0.98	0.3269	β_{23}	0.00060	0.000950000	0.63	0.5278
β_{11}	0.00001	0.000031500	0.33	0.7406	β_{24}	-0.01706	0.012970000	-1.32	0.1891
β_{12}	0.00000	0.000000001	-6.08	0	β_{25}	0.00010	0.000022500	4.52	0

Figure 5.25 shows actual and model results for crack spread for in-the-sample and out-of-sample periods. The mean absolute error of the forecast is 4.65 and the Theil Inequality Coefficient is only 0.17 indicating that the model performance is relatively satisfactory.

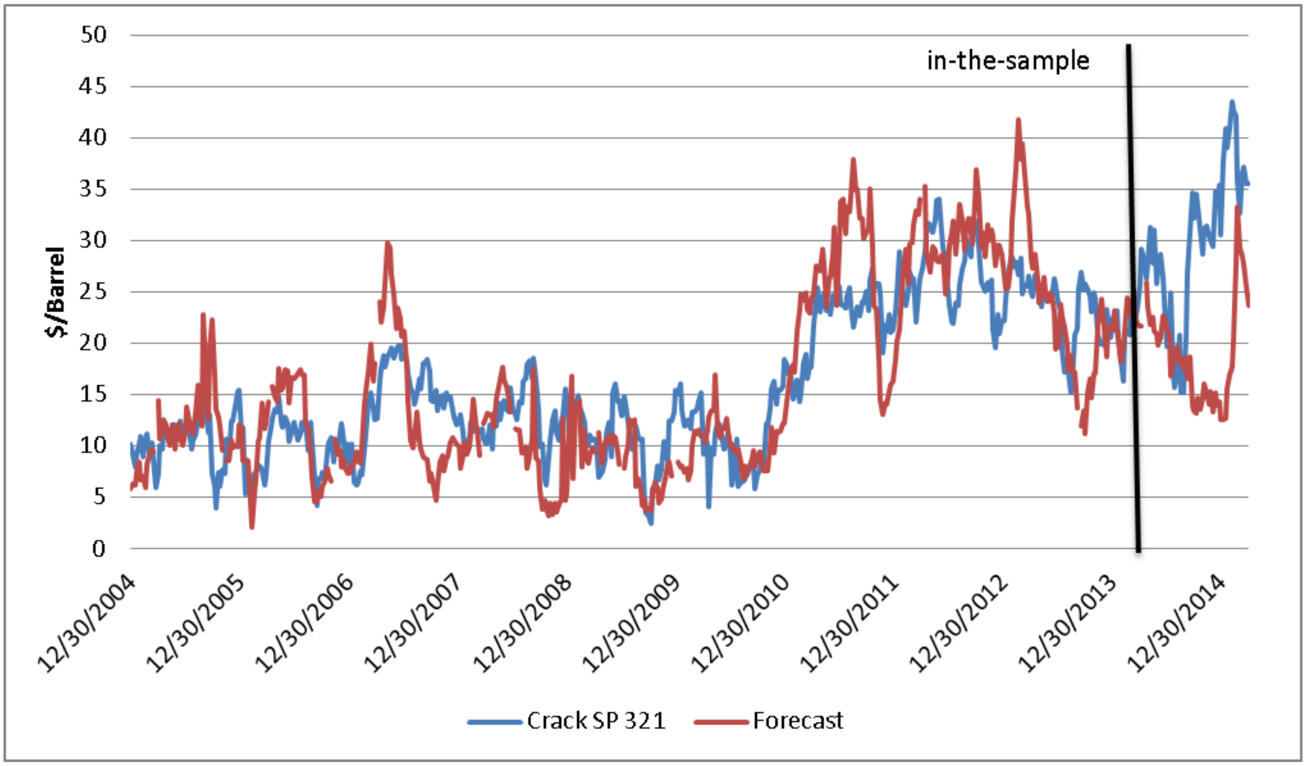


Figure 5.25: Actual vs. Predicted Crack Spread Prices for the US

The out-of-sample starts from October 1, 2014 and runs through April 1^{st}, 2015.

5.3.5.2 Calibrating Spot Crude Oil Prices

In this section we calibrate equation (5.21) by a square functional form similar to the one we used in the previous section for the crack spread:

$$P^j = \psi(D^j(t) - \overline{D}^j(t), V^j(t) - \overline{V}^j(t), c(t)),$$

where ψ is a quadratic functional form, $D^j(t)-\overline{D}^j(t)$ is the deviation of crude oil draws from inventories vs. 5-year average for period *j*, *t* indicates calendar date, and $V^j(t)-\overline{V}^j(t)$ is the deviation of level of crude oil inventories from 5-year averages.

As Figure 5.26 shows the deviation of crude oil inventories from 5-year average vary significantly across five PADD regions. This is of course expected because PADD 2 and PADD 3 have most of the above-ground storage capacity in the US. Figure 5.27 shows the data on $V^j(t)-\overline{V}^j(t)$.

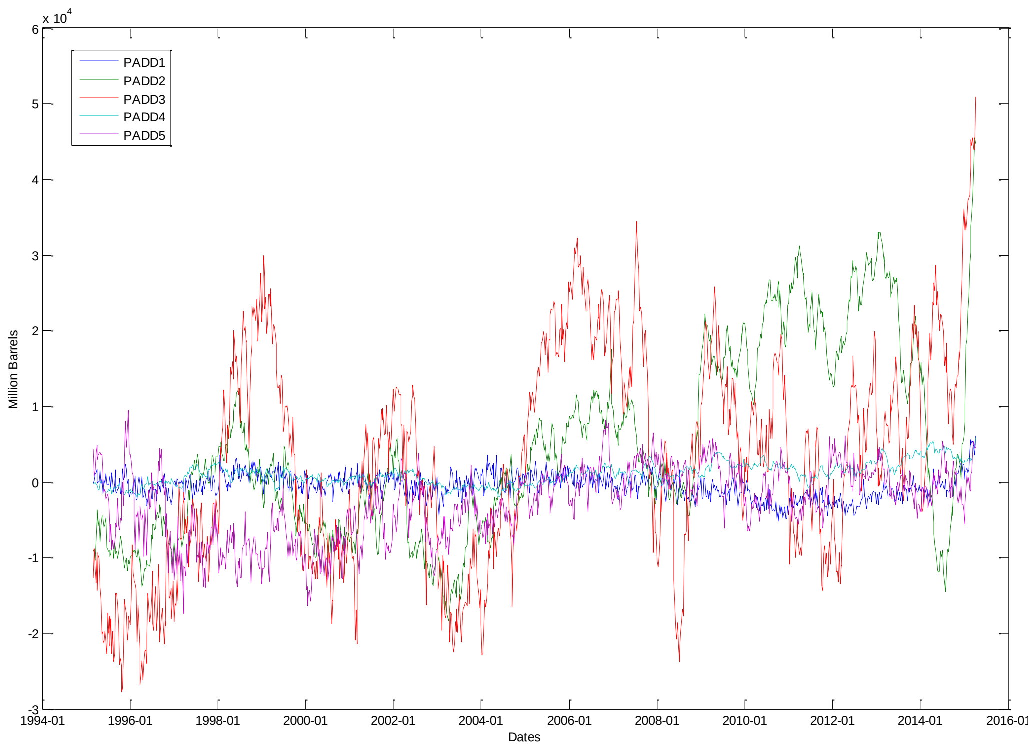


Figure 5.26: Deviation of PADD1-PADD5 Crude Oil Inventories in the US vs. 5-yr Average: 1996-01:2015:04

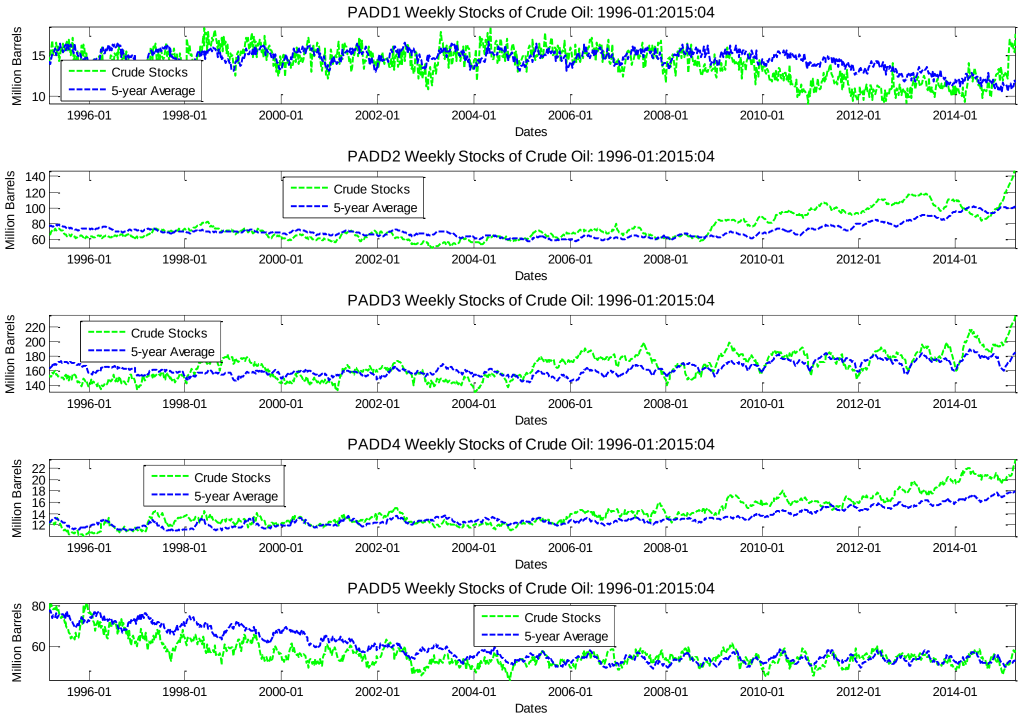


Figure 5.27: PADD1-PADD5 Crude Oil Inventories vs. 5-yr Average: 1996-01:2015:04

Because of regional disparities between 5 PADDs, we calibrate the model for the US by breaking down the level of inventories, draws and also operable refinery capacities by each PADD region. The empirical version of equation (5.21) is specified as follows:

$$
\begin{aligned}
WTI(t) = \beta_0 &+ \beta_1(D_{P1}(t) - \overline{D}_{P1}(t)) + \beta_2(D_{P1}(t) - D_{P1}(t))^2 \\
&+ \beta_3(D_{P2}(t) - \overline{D}_{P2}(t)) + \beta_4(D_{P2}(t) - \overline{D}_{P2}(t))^2 \\
&+ \beta_5(D_{P3}(t) - \overline{D}_{P3}(t)) + \beta_6(D_{P3}(t) - \overline{D}_{P3}(t))^2 \\
&+ \beta_7(D_{P4}(t) - \overline{D}_{P4}(t)) + \beta_8(D_{P4}(t) - \overline{D}_{P4}(t))^2 \\
&+ \beta_9(D_{P5}(t) - \overline{D}_{P5}(t)) + \beta_{10}(D_{P5}(t) - \overline{D}_{P5}(t))^2 \\
&+ \beta_{11}(S_{P1}(t) - \overline{S}_{P1}(t)) + \beta_{12}(S_{P1}(t) - \overline{S}_{P1}(t))^2 \\
&+ \beta_{13}(S_{P2}(t) - \overline{S}_{P2}(t)) + \beta_{14}(S_{P2}(t) - \overline{S}_{P2}(t))^2 \\
&+ \beta_{15}(S_{P3}(t) - \overline{S}_{P3}(t)) + \beta_{16}(S_{P3}(t) - \overline{S}_{P3}(t))^2 \\
&+ \beta_{17}(S_{P4}(t) - \overline{S}_{P4}(t)) + \beta_{18}(S_{P4}(t) - \overline{S}_{P4}(t))^2 \\
&+ \beta_{19}(S_{P5}(t) - \overline{S}_{P5}(t)) + \beta_{20}(S_{P5}(t) - \overline{S}_{P5}(t))^2 \\
&+ \beta_{21}c(t) + \beta_{22}c^2(t) + \beta_{23}TR + \varepsilon(t)
\end{aligned}
\tag{5.31}
$$

where $\varepsilon(t)$ is normally distributed with 0 mean and variance σ^2, t indicates calendar date, and TR is a time trend. D indicates draw from crude oil storages, S level of crude oil inventories, $\overline{D}$ 5-year average of crude oil withdraw from storages, $\overline{S}$ 5-year average of crude oil inventories, and c is the level of inventories at Cushing, Oklahoma. P1 through P5 indicate PADDI through PADD V regions in the US. The data frequency is weekly. We employed Ordinary Least Squares (OLS) to estimate the parameters of equation (5.31). The estimates of parameters and their t-statistics are summarized in Table 5.2 The R^2 of the regression is 68%. The F-statistics is 48. It is notable that we performed Augmented Dicky Fuller test (ADF) of existence of unit root in the residuals of the equation (5.31) and the null of existence of unit root is reject at 5% critical level indicating that the equation (5.31) is cointegrated.

Table 5.2: Crude Oil Parameter estimates with t-statistics and Pr($\beta_i = 0$) = 0 for all *i*'s

Param.	Estimate	Std. Error	t-Stat	Prob	Param.	Estimate	Std. Error	t-Stat	Prob
β_0	-133.9755	139	-0.96	0.336	β_{12}	0.00004	0.000190000	0.20	0.8432
β_1	-0.00212	0.000642000	-3.29	0.0011	β_{13}	-0.00166	0.000386000	-4.29	0
β_2	-0.00067	0.000140000	-4.79	0	β_{14}	-0.00019	0.001707000	-0.11	0.9113
β_3	-0.00085	0.000078600	-10.82	0	β_{15}	-0.00092	0.000379000	-2.43	0.0154
β_4	0.00609	0.001273000	4.78	0	β_{16}	0.00000	0.000000328	-1.36	0.1737
β_5	0.00110	0.000217000	5.08	0	β_{17}	0.00000	0.000000036	-0.11	0.913
β_6	0.00000	0.000000179	-2.73	0.0066	β_{18}	0.00000	0.000000128	-3.09	0.0021
β_7	0.00000	0.000000005	2.03	0.0426	β_{19}	0.00000	0.000003110	-0.03	0.9726
β_8	0.00000	0.000000004	4.19	0	β_{20}	0.00000	0.000000157	-0.85	0.3983
β_9	0.00000	0.000000285	-5.93	0	β_{21}	4.67520	3.229116000	1.45	0.1483
β_{10}	0.00000	0.000000054	-7.31	0	β_{22}	-0.02715	0.018693000	-1.45	0.147
β_{11}	0.00005	0.000580000	0.08	0.9354	β_{23}	0.08074	0.006686000	12.08	0

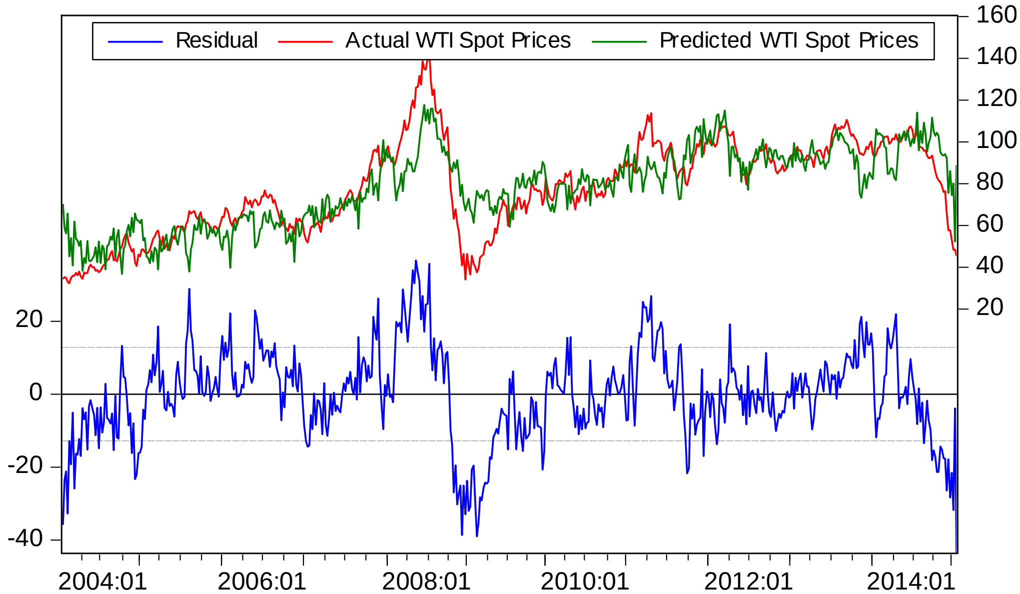


Figure 5.28: Actual, Model Estimate and Residual of Calibrating WTI Prices

Figure 5.28 shows actual and model results for calibrating equation (5.31). The mean absolute error of the forecast is 9.5 and the Theil Inequality Coefficient is only 0.07 indicating that the model performance is relatively satisfactory.

5.4 Summary

In this chapter we focused on coming up with fair value of crude oil futures contracts in any given day after spot prices settled every day, and also given level of inventories and draws. We carefully constructed our data sets by changing reference of data from calendar date to time to expiry of each future contract. In this way we ensured not to jump from one contract to the next one by ignoring how and when contracts role. We provided a practical method which could help producers, market makers and other market participants to take spot prices traded in the market as given, and price long dated crude forward contracts on a real time basis.

We then used the setup of the two-factor model and used it to add inventory and draws information to the valuation process. In this case, producers and other market participants can use physical information and feed them into futures pricing methods. We presented a set of structural models to show how the dynamics of crude market works. We used the framework to calibrate crude oil futures, and crude oil spreads. We performed sensitivity analysis of fair value of futures and future spreads with respect to level of inventories and storage draws.

Chapter Six: Variance Risk Premium in Crude Oil Prices

6.1 Introduction and Literature Review

In this chapter we investigate the variance risk premia in crude oil futures market. Approximating the risk premia in crude oil futures contracts is very important as existence of a consistent premia would provide potential hedging and trading opportunities for market participants. Previous studies show that it pays to short variance swaps in crude oil future; in other words, they show existence of negative excess return in crude oil futures. However, our results show that there is no variance risk premium in crude oil futures all the time, and the strategy only pays in "normal" market situations but as soon as an unexpected event in crude oil market takes place the strategy loses very significantly. It seems the negative excess return exists because some market participants pay to protect them at the time of crisis. Our results are based on a realistic experiment we have designed for this purpose. Unlike sample data of other studies that do not include a rare event, we use daily NYMEX option settles from August 2013 through April 2015 which covers a 50% drop in crude oil prices in a short period of time.

Variance Swap is an instrument commonly used by market participants to trade future realized volatility versus market implied option volatility. It helps market participants to bet weather or not realized variance will match the current implied variance. At the time of expiry, a swap payment is made if the realized one was not the same as the initial implied variance.

There are not many studies on variance risk premium in crude oil prices, and most studies on this subject are done for stocks and bonds. However, pretty much every study on every asset class in general, and crude oil in particular have concluded that there are negative excess returns in variance swaps in these asset classes. For example, Carr and Wu (2008) investigate variance risk premiums in S&P 500 and 100 indexes, and Dow Jones Industrial Averages and show that there is significant negative excess return in these assets. Their data sample starts from January 1996 and ends in February 2003. Trolle and Schwartz (2009) follow similar approach as Carr and Wu (2008) and show that the variance risk premia for crude oil and natural gas are both negative, implying shorting variance on these assets are a viable strategy. Their crude oil option data is from January 2, 1996 through November 30, 2006. They get the same results for S&P 500 index as well. Bollerslev et all (2009) study behavior of expected stock returns and variance risk premia and try to establish a relationship between variance risk premium and stock market returns in post-1990 through December 2007. They show that returns are predictable by the variance risk premium. One of the main drawbacks of all these studies is their sample coverage. All these studies only cover pre-2008 financial crisis, and do not include any major risky event in the financial markets.

Understanding the existence of consistent negative excess return on variance swaps in these assets was the main motivation behind this research. The absence of a major rare event in the sample data sets of previous studies also reinforced the idea. This was always a puzzle in our minds, in a sense that how and why market could misprice an instrument all the time? We solely focused on crude oil market and designed a very realistic and real-world type experiment by building forward-type curves on crude oil options from August 2013 through April 2015 with total of about 3100 option contracts on those futures contracts. Our sample does include a big event in crude oil markets which came as a shock by about 50% drop in crude oil prices in over a six month period. Our findings are different from other studies on crude oil Variance Swap. Our results indicate that the negative excess return is not an opportunity at all because as soon as turmoil shows up in crude oil markets that strategy results in very substantial hefty losses. It seems variance swap was consistently higher than realized variances through August 2014. However, as soon as crude oil prices started to go down significantly as of late July 2014, the realized variances significantly exceeded variance swaps. This is a very interesting observation. This is probably the reason why negative variance premium exists most of the time. Market participants tend to pay some premium to protect themselves from unexpected events in crude oil market. It seems the so-called strategy of shorting variance risk premium pays off in “normal” situations but the strategy gives all back “at-once” as soon as a very unlikely outcome takes place

6.2 Model Setup

The setup and the theoretical definitions in this paper are based on Carr and Wu (2008) and Trolle and Schwartz (2009). The payoff function at expiry, T, is given by

$$[(V(t,T)-K(t,T)]L \tag{6.1}$$

where $K(t,T)$ denotes the implied variance agreed at time t, L represents the notional value of the swap contract, V is the realized annualized return on variance which can be defined in a discrete setting as

$$V(t,T)=\frac{1}{N\Delta t}\sum_{i=1}^{N}\left[\log(\frac{F(t_i,T_1)}{F(t_{i-1},T_1)})\right]^2 \tag{6.2}$$

where $T\le T_1$, Δt is the time step, and N the number of trading days between inception of the trade and the expiry of the contract, $F(t,T_1)$.

At the inception of the trade, $t=t_0$, we have $[(V(t,T)-K(t,T)]L$ =0. It is notable that

$$K(t,T)=E_t^Q[(V(t,T)] \tag{6.3}$$

where the risk-neutral measure is denoted by Q.

Following Trolle and Schwartz (2009), given the futures price process is continuous, $K(t,T)$ is calculated as – which is exact,

$$K(t,T)=\frac{2}{B(t,T)(T-t)}\left\{\int_0^{F(t,T_1)}\frac{p(t,T,T_1,x)}{k^2}dx+\int_{F(t,T_1)}^{\infty}\frac{c(t,T,T_1,x)}{k^2}dx\right\} \tag{6.4}$$

Where $K(t,T)$ is the value of implied variance defined in (6.3), $F(t,T_1)$ is the price of forward contract, $B(t,T)$ is the price of a zero-coupon bond expiring at T, and *p* and *c* are the value of put and call options both expire at T on a futures contract expiry at T_1, with strike of *k*. Intuitively, this builds a weight in inverse proportion to the square of strikes which will makes the fair value of implied variance independent of directional moves in the underlying.

6.3 Data and Implementation

We use Bloomberg to get data on WTI crude oil NYMEX futures and options contracts. Because option prices were only available from Aug 2013 contract, we use the daily data on both futures and options to implement and calculate variance swap for crude oil. We used 21 options and underlying futures contracts starting from Aug 2013 through April 2015. On each option type we collect around 70 contracts with varying strikes on every underlying future contract. For example, for September 2014 we have to collect historical settlements of 78 contracts for crude oil September call options with strikes starting from 40 through 117, and also 78 contracts for its put options. Our final data sets have around 3100 call and put option contracts with different strikes on 21 crude oil futures contracts.

In order to calculate the fair value of strikes on variance swap defined in (6.4) we designed our experiment as such so that it is as realistic as possible.

1. Pick a crude oil future contract
2. A typical crude oil option on a future contract say August expires sometime in mid-July, take position on these option contracts 45 days before expiry of the option

3. Remove non trading days from the sample of historical data and calculate business days between inception date and exit date
4. On inception date, pick the underlying price of the corresponding crude oil future contract, $F(t,T_1)$
5. Create portfolio of call options

$$c(t,T,T_1,x), \qquad x = y+\$1, y+\$2,..., y+\$10$$

where y is strike of the at the money option.

6. Create portfolio of put options

$$p(t,T,T_1,x), \qquad x = y-\$1, y-\$2,..., y-\$10$$

7. Use equation (6.4) to calculate fair value of strike on variance swap, K
8. Use equation (6.2) to calculate V
9. Calculate pay-off on notional value of \$100, $[(V(t,T)-K(t,T)]L$
10. Calculate excess return, $\log[V(t,T)/K(t,T)]$

6.4 Model Results

The Table 6.1 provides information on entry dates and corresponding crude oil prices of each of the forward contracts at the time of taking Variance Swap position on each future crude oil contract.

Table 6.1: Inception Dates, Exit Dates and Crude Price on Inception Date

Contract	Trade Date	Exit Date	Future Price
Aug-13	29-May-13	16-Jul-13	93.35
Sep-13	27-Jun-13	16-Aug-13	96.89
Oct-13	29-Jul-13	16-Sep-13	103.90
Nov-13	29-Aug-13	16-Oct-13	108.15
Dec-13	27-Sep-13	14-Nov-13	102.34
Jan-14	29-Oct-13	16-Dec-13	98.38
Feb-14	29-Nov-13	16-Jan-14	93.01
Mar-14	27-Dec-13	14-Feb-14	100.39
Apr-14	29-Jan-14	14-Mar-14	96.88
May-14	27-Feb-14	16-Apr-14	101.62
Jun-14	27-Mar-14	16-May-14	100.53
Jul-14	29-Apr-14	16-Jun-14	100.59
Aug-14	29-May-14	16-Jul-14	102.78
Sep-14	27-Jun-14	14-Aug-14	105.04
Oct-14	29-Jul-14	16-Sep-14	99.73
Nov-14	29-Aug-14	16-Oct-14	95.02
Dec-14	29-Sep-14	14-Nov-14	93.49
Jan-15	29-Oct-14	16-Dec-14	81.96
Feb-15	27-Nov-14	16-Jan-15	73.76
Mar-15	29-Dec-14	16-Feb-15	54.03
Apr-15	29-Jan-15	16-Mar-15	45.30

For example, as for the Apr-15 NYMEX future contract we take a variance swap position on January 29, 2015 when crude oil price on that contract was 45.3. We calculate the fair value of the strike based on information provided in Table 6.2 and 7.3 and steps outlined in the previous section. We close this position on March 16, 2015. In fact, we have done this for all 21 futures contracts from August 2013 through April 2015. It is notable that there is usually two weeks overlap between two consecutive positions. For example, at the time when we initiated variance swap position for April 2015 contract we still had a variance swap position on March 2015 contract that going to get closed on February 16, 2015.

Table 6.2: Strikes of Out-of-the-Money Put Options Contracts on Inception Date

Contract					Strikes of Out-of-the-Money Puts						
Aug-13	$ 93.0	$ 92.0	$ 91.0	$ 90.0	$ 89.0	$ 88.0	$ 87.0	$ 86.0	$ 85.0	$ 84.0	$ 83.0
Sep-13	$ 96.0	$ 95.0	$ 94.0	$ 93.0	$ 92.0	$ 91.0	$ 90.0	$ 89.0	$ 88.0	$ 87.0	$ 86.0
Oct-13	$ 103.0	$ 102.0	$ 101.0	$ 100.0	$ 99.0	$ 98.0	$ 97.0	$ 96.0	$ 95.0	$ 94.0	$ 93.0
Nov-13	$ 108.0	$ 107.0	$ 106.0	$ 105.0	$ 104.0	$ 103.0	$ 102.0	$ 101.0	$ 100.0	$ 99.0	$ 98.0
Dec-13	$ 102.0	$ 101.0	$ 100.0	$ 99.0	$ 98.0	$ 97.0	$ 96.0	$ 95.0	$ 94.0	$ 93.0	$ 92.0
Jan-14	$ 98.0	$ 97.0	$ 96.0	$ 95.0	$ 94.0	$ 93.0	$ 92.0	$ 91.0	$ 90.0	$ 89.0	$ 88.0
Feb-14	$ 93.0	$ 92.0	$ 91.0	$ 90.0	$ 89.0	$ 88.0	$ 87.0	$ 86.0	$ 85.0	$ 84.0	$ 83.0
Mar-14	$ 100.0	$ 99.0	$ 98.0	$ 97.0	$ 96.0	$ 95.0	$ 94.0	$ 93.0	$ 92.0	$ 91.0	$ 90.0
Apr-14	$ 96.0	$ 95.0	$ 94.0	$ 93.0	$ 92.0	$ 91.0	$ 90.0	$ 89.0	$ 88.0	$ 87.0	$ 86.0
May-14	$ 101.0	$ 100.0	$ 99.0	$ 98.0	$ 97.0	$ 96.0	$ 95.0	$ 94.0	$ 93.0	$ 92.0	$ 91.0
Jun-14	$ 100.0	$ 99.0	$ 98.0	$ 97.0	$ 96.0	$ 95.0	$ 94.0	$ 93.0	$ 92.0	$ 91.0	$ 90.0
Jul-14	$ 100.0	$ 99.0	$ 98.0	$ 97.0	$ 96.0	$ 95.0	$ 94.0	$ 93.0	$ 92.0	$ 91.0	$ 90.0
Aug-14	$ 102.0	$ 101.0	$ 100.0	$ 99.0	$ 98.0	$ 97.0	$ 96.0	$ 95.0	$ 94.0	$ 93.0	$ 92.0
Sep-14	$ 105.0	$ 104.0	$ 103.0	$ 102.0	$ 101.0	$ 100.0	$ 99.0	$ 98.0	$ 97.0	$ 96.0	$ 95.0
Oct-14	$ 99.0	$ 98.0	$ 97.0	$ 96.0	$ 95.0	$ 94.0	$ 93.0	$ 92.0	$ 91.0	$ 90.0	$ 89.0
Nov-14	$ 95.0	$ 94.0	$ 93.0	$ 92.0	$ 91.0	$ 90.0	$ 89.0	$ 88.0	$ 87.0	$ 86.0	$ 85.0
Dec-14	$ 93.0	$ 92.0	$ 91.0	$ 90.0	$ 89.0	$ 88.0	$ 87.0	$ 86.0	$ 85.0	$ 84.0	$ 83.0
Jan-15	$ 81.0	$ 80.0	$ 79.0	$ 78.0	$ 77.0	$ 76.0	$ 75.0	$ 74.0	$ 73.0	$ 72.0	$ 71.0
Feb-15	$ 73.0	$ 72.0	$ 71.0	$ 70.0	$ 69.0	$ 68.0	$ 67.0	$ 66.0	$ 65.0	$ 64.0	$ 63.0
Mar-15	$ 54.0	$ 53.0	$ 52.0	$ 51.0	$ 50.0	$ 49.0	$ 48.0	$ 47.0	$ 46.0	$ 45.0	$ 44.0
Apr-15	$ 45.0	$ 44.0	$ 43.0	$ 42.0	$ 41.0	$ 40.0	$ 39.0	$ 38.0	$ 37.0	$ 36.0	$ 35.0

Table 6.3: Strikes of Out-of-the-Money Call Options Contracts on Inception Date

Contract	Strikes of Out-of-the-Money Calls										
Aug-13	$ 94.0	$ 95.0	$ 96.0	$ 97.0	$ 98.0	$ 99.0	$ 100.0	$ 101.0	$ 102.0	$ 103.0	$ 104.0
Sep-13	$ 97.0	$ 98.0	$ 99.0	$ 100.0	$ 101.0	$ 102.0	$ 103.0	$ 104.0	$ 105.0	$ 106.0	$ 107.0
Oct-13	$ 104.0	$ 105.0	$ 106.0	$ 107.0	$ 108.0	$ 109.0	$ 110.0	$ 111.0	$ 112.0	$ 113.0	$ 114.0
Nov-13	$ 109.0	$ 110.0	$ 111.0	$ 112.0	$ 113.0	$ 114.0	$ 115.0	$ 116.0	$ 117.0	$ 118.0	$ 119.0
Dec-13	$ 103.0	$ 104.0	$ 105.0	$ 106.0	$ 107.0	$ 108.0	$ 109.0	$ 110.0	$ 111.0	$ 112.0	$ 113.0
Jan-14	$ 99.0	$ 100.0	$ 101.0	$ 102.0	$ 103.0	$ 104.0	$ 105.0	$ 106.0	$ 107.0	$ 108.0	$ 109.0
Feb-14	$ 94.0	$ 95.0	$ 96.0	$ 97.0	$ 98.0	$ 99.0	$ 100.0	$ 101.0	$ 102.0	$ 103.0	$ 104.0
Mar-14	$ 101.0	$ 102.0	$ 103.0	$ 104.0	$ 105.0	$ 106.0	$ 107.0	$ 108.0	$ 109.0	$ 110.0	$ 111.0
Apr-14	$ 97.0	$ 98.0	$ 99.0	$ 100.0	$ 101.0	$ 102.0	$ 103.0	$ 104.0	$ 105.0	$ 106.0	$ 107.0
May-14	$ 102.0	$ 103.0	$ 104.0	$ 105.0	$ 106.0	$ 107.0	$ 108.0	$ 109.0	$ 110.0	$ 111.0	$ 112.0
Jun-14	$ 101.0	$ 102.0	$ 103.0	$ 104.0	$ 105.0	$ 106.0	$ 107.0	$ 108.0	$ 109.0	$ 110.0	$ 111.0
Jul-14	$ 101.0	$ 102.0	$ 103.0	$ 104.0	$ 105.0	$ 106.0	$ 107.0	$ 108.0	$ 109.0	$ 110.0	$ 111.0
Aug-14	$ 103.0	$ 104.0	$ 105.0	$ 106.0	$ 107.0	$ 108.0	$ 109.0	$ 110.0	$ 111.0	$ 112.0	$ 113.0
Sep-14	$ 106.0	$ 107.0	$ 108.0	$ 109.0	$ 110.0	$ 111.0	$ 112.0	$ 113.0	$ 114.0	$ 115.0	$ 116.0
Oct-14	$ 100.0	$ 101.0	$ 102.0	$ 103.0	$ 104.0	$ 105.0	$ 106.0	$ 107.0	$ 108.0	$ 109.0	$ 110.0
Nov-14	$ 96.0	$ 97.0	$ 98.0	$ 99.0	$ 100.0	$ 101.0	$ 102.0	$ 103.0	$ 104.0	$ 105.0	$ 106.0
Dec-14	$ 94.0	$ 95.0	$ 96.0	$ 97.0	$ 98.0	$ 99.0	$ 100.0	$ 101.0	$ 102.0	$ 103.0	$ 104.0
Jan-15	$ 82.0	$ 83.0	$ 84.0	$ 85.0	$ 86.0	$ 87.0	$ 88.0	$ 89.0	$ 90.0	$ 91.0	$ 92.0
Feb-15	$ 74.0	$ 75.0	$ 76.0	$ 77.0	$ 78.0	$ 79.0	$ 80.0	$ 81.0	$ 82.0	$ 83.0	$ 84.0
Mar-15	$ 55.0	$ 56.0	$ 57.0	$ 58.0	$ 59.0	$ 60.0	$ 61.0	$ 62.0	$ 63.0	$ 64.0	$ 65.0
Apr-15	$ 46.0	$ 47.0	$ 48.0	$ 49.0	$ 50.0	$ 51.0	$ 52.0	$ 53.0	$ 54.0	$ 55.0	$ 56.0

The Table 6.4 show summary statistics of the variance swap strikes, and realized variances for all 21 crude oil futures contracts. Our results indicate that the mean of variance swap is smaller than the mean of realized variance, $V(t,T) > K(t,T)$.

Table 6.4: Summary Statistics of Variance Swap and the Realized Variance

	K	V	Pay-Off	Excess Return
Mean	0.06	0.09	0.03	0.03
Minimum	0.02	0.01	(0.03)	(0.27)
Maximum	0.26	0.39	0.24	0.57
Std. dev.	0.07	0.12	0.07	0.26

In this case the pay-off, $[(V(t,T)-K(t,T)]$, is positive with excess return, $\log[V(t,T)/K(t,T)]$, close to 3% if we go long variance swap.

At first, this seems contrary to the results obtained by Trolle and Schwartz (2009), however, looking at results grouped by each contract does not necessarily reject their findings. The Table 6.5 show variance swap, realized variance, pay-off and excess return for each individual contract. Our results are interesting and shed more light into risk premium in crude oil futures contract. There is usually excess return of shorting the variance swap but it seems this strategy did not pay when crude oil market was in turmoil.

Table 6.5: Variance Swaps Pay-Offs and Excess Returns of Crude Oil Futures

Contract	K	V	Pay-Off	Excess Return
Aug-13	0.051	0.040	(0.011)	(0.104)
Sep-13	0.044	0.034	(0.009)	(0.102)
Oct-13	0.041	0.040	(0.002)	(0.019)
Nov-13	0.056	0.031	(0.025)	(0.259)
Dec-13	0.042	0.031	(0.011)	(0.132)
Jan-14	0.036	0.027	(0.009)	(0.120)
Feb-14	0.035	0.028	(0.007)	(0.095)
Mar-14	0.022	0.029	0.006	0.105
Apr-14	0.039	0.027	(0.012)	(0.158)
May-14	0.031	0.025	(0.006)	(0.095)
Jun-14	0.030	0.017	(0.014)	(0.264)
Jul-14	0.029	0.016	(0.013)	(0.265)
Aug-14	0.022	0.014	(0.007)	(0.181)
Sep-14	0.022	0.020	(0.003)	(0.057)
Oct-14	0.022	0.038	0.016	0.232
Nov-14	0.023	0.064	0.040	0.434
Dec-14	0.035	0.076	0.041	0.333
Jan-15	0.066	0.205	0.139	0.494
Feb-15	0.089	0.332	0.243	0.571
Mar-15	0.240	0.391	0.151	0.213
Apr-15	0.259	0.330	0.071	0.105

As Table 6.5 indicates, the average excess return of shorting crude oil futures variance swap was well over 10% for August 2013 through September 2014 futures contracts. However, it's average excess return dropped significantly for October 2014 through April 2015 by around 40%.

The Figure 6.1 depicts the variance swap and realized variance along with average crude oil prices between inception and exit dates for each futures contract.

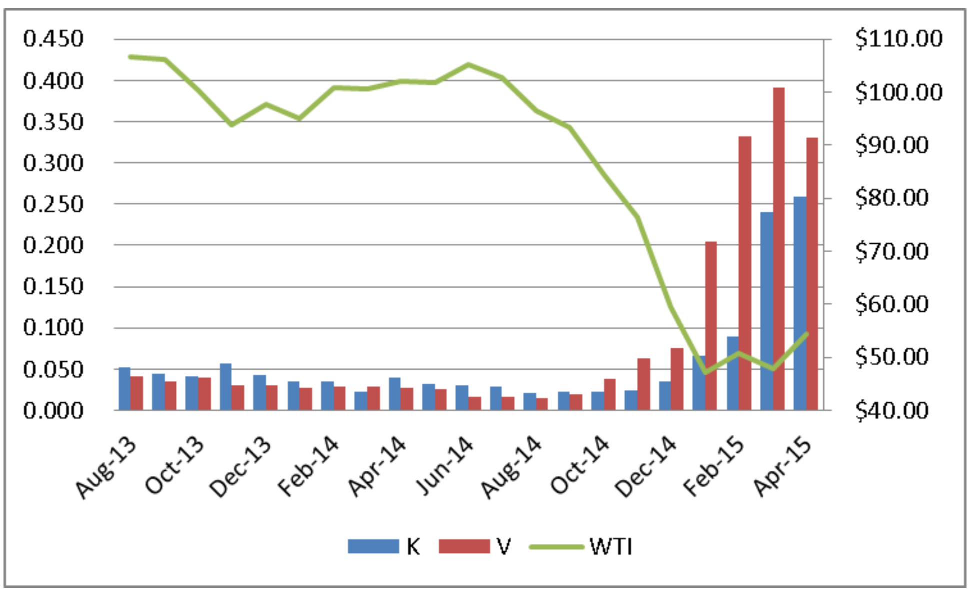


Figure 6.1: Variance Swap and Realized Variance vs. Average Crude Oil Prices

It seems variance swap was consistently higher than realized variances through August 2014. However, as soon as crude oil prices started to go down significantly as of late July 2014, the realized variances significantly exceeded variance swaps. This is a very interesting observation.

This is probably the reason why negative variance premium, $[(V(t,T)-K(t,T)]$, exists most of the time. Market participants tend to pay some premium to protect themselves from unexpected events in crude oil market. As shown in Figure 6.2, the so-called strategy of shorting variance risk premium pays off in "normal" situations but the strategy gives all back "at-once" as soon as a very unlikely outcome takes place.

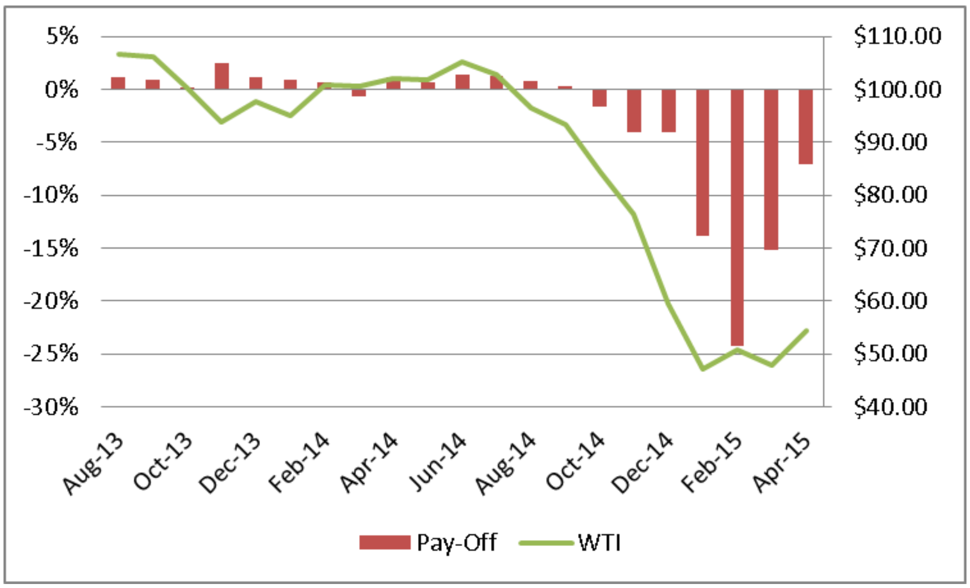


Figure 6.2: Pay-off of Shorting $[(V(t,T)-K(t,T)]$**.($1) vs. Average Crude Oil Prices**

Chapter Seven: Conclusion

7.1 Overview

In Chapter 2 we review the literature regarding quantitative analysis of crude oil prices. In general literature related to crude oil prices is partly focused on analyses of crude oil prices from statistical point of views. Some of the literature deals with developing risk metrics for crude oil prices, and finally there are some works on pricing instruments trading in the crude oil markets.

In Chapter 3 we focus on come up with fair value of crude oil futures contracts in any given day after having spot prices. We use a set of one and two factor models to model crude oil prices. We carefully construct our data sets by changing reference of data from calendar date to time to expiry of each future contract. In this way we do not jump from one contract to the next one by ignoring how and when contracts role. We use daily spot and forward data and calibrate parameters using an optimization approach known as Particle-Swamp Optimization (PSO), and argue why my approach is superior to others in the literature. We provide a practical method which could help producers, market makers and other market participants to take spot prices traded in the market as given, and price long dated crude forward contracts on a real time basis.

In Chapter 4 we focus on pricing crude oil options using Merton's Jump Diffusion Model (MJDM), Normal Inverse Gaussian Model (NIGM), and Variance Gamma Model (VGM), which belong to family of Levy processes. We present characteristic functions of these processes, and then we use current market option prices to calibrate parameters of these models. We used Fast Fourier Transform (FRFT) algorithm to calibrate parameters of NIGM and VGM, and used PSO to calibrate MJDM. Our results are satisfactory of options not very far from ATM strikes.

We allocate Chapter 5 to building a bridge between risk-neutrality and structure of the crude oil markets. This is an improvement to what we have done in Chapter 3. We try to provide a framework for producers and other market participants to use physical information and feed them into futures pricing methods. We first present a set of structural models to show how the dynamics of crude market works. We use the framework to calibrate crude oil prices, and then we use the framework of our two-factor model presented in Chapter 3 to build a relationship between level of inventories and other key physical variables in the crude oil market and valuation methods. We provide out-of-sample results and compare them with actual data as well.

In Chapter 6 we investigated variance risk premia in crude oil prices using information obtained from crude oil option prices. We provide detailed steps on constructing the required data sets, and designing a realistic experiment. Our results indicate that there is a negative risk premium in crude oil prices but that does not necessarily provide trading opportunity for market participants because excess return of shorting the variance swap show huge losses when crude oil market is in turmoil.

7.2 Our Contributions

There are very few researches in the literature applying mathematical finance to crude oil futures markets. We have come to believe that the main reason could be complexities of the data sets. There are lots of studies on behavior of crude oil spot prices and crude oil risks but not much on futures curve in crude oil prices.

- We used spot and futures data sets and imposed risk neutrality on our two factor models to calibrate our parameters. We believe this is the first time this is done for crude oil markets. Our models help to solve a problem that market participants have to deal with in a given day; what is the fair value of a future contract when spot market settles? Our models in Chapter 3 provide a practical solution for this question.

- We provided a practical solution to calculate fair value of crude oil future contracts by utilizing inventories and capacity data in the process. We discussed detail steps on how to implement it. We also calculated fair value of future spreads along with sensitivity analysis of fair values of futures and future spreads with respect to level of inventories and draws from storages. We provided out of sample results as well.
- We believe this is most likely the first application of NIG, JDM, and VG models in crude oil markets by calibrating their parameters using option markets and futures prices. This is of course done by others for stocks and bonds, where all forward option contracts have the same spot price. However, in the case of crude oil market each option contract on a future has a different spot price from another option contract. This adds significant complexity to calibration and implementation process.
- All past studies on Variance Risk Premia in pretty much every asset in general, and crude oil markets in particular have concluded that there is always a negative excess return in Variance Swap. This was always a puzzle in our minds, in a sense that how and why market could misprice an instrument all the time? We designed a very realistic and real-world type experiment by building forward-type curves on crude oil options from August 2013 through April 2015 with total of about 3100

option contracts on those futures contracts. Our findings are different from the other studies on crude oil Variance Swap. Our results indicate that the negative excess return is not an opportunity at all because as soon as a turmoil starts in crude oil markets that strategy results in hefty losses.

Bibliography

1. Applebaum, A., 2003. Levy Processes and Stochastic Calculus, Cambridge University Press.
2. Askari, H., and Krichene, N., 2008, Oil Price Dynamics, Energy Economics 30, p 2134–2153.
3. Bailey, D. H., and Swarztrauber, P. N. 1991, The Fractional Fourier Transform and applications, SIAM Review 33 (3), 389–404.
4. Bjerksund, P., and S. Ekern, 1995, Contingent claims evaluation of mean-reverting cash flows in shipping, in L. Trigeorgis, Ed.: Real Options in Capital Investment: Models, Strategies, and Applications, Preager.
5. Brennan, M.J., 1958. The supply of storage. American Economic Review 48, 50–72
6. Bollerslev D., Tauchen, G., and Zhou, H., Expected Stock Returns and Variance Risk Premia, The Review of Financial Studies, V 22 n11 2009.
7. Benth, F.E., and Benth, J., 2004. The normal inverse gaussian distribution and spot price modeling in energy markets. Intern. J. Theor. Appl. Finance, Pages, 7:177.
8. Brennan, M.J., 1958. The supply of storage. American Economic Review 48, 50–72
9. Cabedo, J.D., Moya, I., 2003. Estimating oil price value-at-risk using the historical simulation approach. Energy Economics 25, 239–253.
10. Carmona, R., Ludkovski, M., 2004, Spot Convenience Yield Models for the Energy Markets. Contemporary Mathematics, 351, 65-80
11. Carr, P. and D. Madan, 1998. Towards a theory of volatility trading, in Robert Jarrow, 7^{th} Editon: Risk Book on Volatility, Risk, New York.
12. Chourdakis, K., 2004. Option Pricing Using the Fractional FFT. Working Paper available on www.theponytail.net
13. Crosby, J., 2008, A multi-factor jump-diffusion model for commodities, Quantitative Finance, Vol. 8, No. 2.

14. Deville, D., 2008. On Levy Processes for Option Pricing: Numerical Methods and Calibration to Index Options. PhD thesis, Universit`a Politecnica delle Marche.
15. Doran, S., 2004., On the Market Price of Volatility Risk, PhD Thesis, University of Texas at Austin.
16. Eberlein, E., and Raible, S., 1999. Term Structure Models Driven by General Levy Processes, Mathematical Finance, Vol. 9, Issue 1, Jan 1999.
17. Ellwanger, R., 2015. The tail Risk Premium in the Crude Oil Futures Market, Job Market Paper, European University Institute.
18. Eryigit, M., 2009., Effects of Oil Price Changes on the Sector Indices of Istanbul Stock Exchange, International Research Journal of Finance and Economics, Issue 25.
19. Fan, Y., Zhange Y., Tsai, H., Wei Y., 2008., Estimating 'Value at Risk' of crude oil price and its spillover effect using the GED-GARCH approach, Energy Economics, 30, 3156-3171.
20. Fama, E.F., French, K.R., 1987. Commodity futures prices: some evidence on forecast power, premiums and the theory of storage. Journal of Business 60, 55–73.
21. Gatheral, J., 2006, The volatility surface : a practitioner's guide.
22. Geman, H., Ohana, S., 2009. Forward curves, scarcity and price volatility in oil and natural gas markets, Energy Economics 31 (2009) 576–585.
23. Gibson, R., and Schwartz, E., 1990., Stochastic convenience yield and the pricing of oil contingent claims, Journal of Finance, 959-976.
24. Giot, P., Laurent, S., 2003. Market risk in commodity markets: a VaR approach. Energy Economics 25, 435–457.
25. Gorton, G., Hayashi, F., Rouwenhorst, K., 2007. "The Fundamentals of Commodity Futures Returns," NBER,
26. Gronwald, M., 2009. Jumps in Oil Prices – Evidence and Implications, IFO Institute for Economic Research at the University of Munich, Working Paper No. 75.
27. Hamilton, James D., 1983. Oil and the Macroeconomy since World War II, Journal of Political Economy, vol. 91(2), pages 228-248, April.

28. Hamilton, J.D., 2001. A parametric approachto 'exible nonlinear inference. Econometrica 69, 537–573.
29. Hamilton, James D., 2003. "What is an oil shock?," Journal of Econometrics, Elsevier, vol. 113(2), pages 363-398, April.
30. Heston, S., 1993., A closed-form solutions for options with stochastic volatility, Review of Financial Studies, 6, 327–343.
31. Heston, S. and S. Nandi., 1997., A Closed Form GARCH Option Pricing Model, Federal Reserve Bank of Atlanta Working Paper 97-9, 1–34.
32. Jamshidian, F. and M. Fein, 1990, Closed-form solutions for oil futures and European options in the Gibson-Schwartz model: A note, Working paper, Merrill Lynch Capital Markets.
33. Kaldor, N., 1939. Speculation and Economic Stability. The Review of Economic Studies, 7, pp. 1–27.
34. Kaijie, C., Swishchuk, A., 2015, "Applications of weather derivatives in the energy market." The Journal of Energy Markets 8.1.
35. Kennedy, J., Eberhart, R., 1995, "Particle Swarm Optimization". Proceedings of IEEE International Conference on Neural Networks **IV**. pp. 1942–194
36. Kou, S.G., 2008. Jump-Diffusion Models for Asset Pricing in Financial Engineering, J.R. Birge and V. Linetsky (Eds.), Handbooks in OR & MS, Vol. 15.
37. Kuper, G., 2002. Measuring Oil Price Volatility. CCSO Working Paper, No. 2002/08. Groningen University. http://www.eco.rug.nl/ccso/quarterly/2002q2.html.
38. Mandelbrot, B. 1963a, The variation of certain speculative prices, Journal of Business, 36, 394-419.
39. Mandelbrot, B. 1963b, New methods in statistical economics', Journal of Political Economy, 61, 421-440.
40. Marimoutou, V., Raggad, B., Trabelsi, S., 2009. Extreme Value Theory and Value at Risk: Application to oil market. Energy Economics 31, 519–530.
41. Merton, R., 1976. Option pricing when underlying stock returns are discontinuous. Bell Journal of Financial Economics, 3, 125–144.

42. Miltersen. K., 2003., Commodity price modeling that matches current observables: A new approach, Quantitative Finance 3(1), 51-58.
43. Rydberg, T., 1997. The normal inverse Gaussian L´evy process: Simulation and approximation. Comm. Statist. Stochastic Models, 13, 887–910.
44. Schwartz, E., Gibson, R, 1990. Stochastic Convenience Yield and the Pricing of Oil Contingent Claims, The Journal of Finance, Vol. XLV, No. 3.
45. Schwartz, E., 1997. The stochastic behavior of commodity prices: implications for pricing and hedging, J. Finance, 52, 923-973.
46. Shahmoradi, A., Swishchuk, A., 2015. Pricing and Sensitivity Analysis of Crude Oil Futures, and Future Spreads, ***Submitted to*** Journal of Empirical Finance.
47. Shahmoradi, A., Swishchuk, A., 2015. Pricing Crude Oil Options Using Levy Processes, ***(Forthcoming: Accepted for publication on Sep 29, 2015)*** Journal of Energy Markets.
48. Shahmoradi, A., Swishchuk, A., 2015. Variance Risk Premium in Crude Oil Futures, ***Submitted to*** Journal of Futures.
49. Takayama, A., 1993, Analytical Methods in Economics, The University of Michigan Press.
50. Trolle, A., Schwartz., E., 2009. Varinace risk premia in energy commodities, J. of Derivatives, Spring, v. 17, No. 3, 15-32.
51. Yang, F., 2011. Investment Shocks, Capacity Constraints, and the Commodity Basis Spread, Department of Finance, Carlson School of Management, University of Minnesota.