Basis Variation and the Role of Inventories: Evidence from the Crude Oil Market

Bassam Fattouh[1]

Oxford Institute for Energy Studies

January 2009

WPM 38

[1] Dr Bassam Fattouh is the Director of the Oil and the Middle East Programme at the Oxford Institute for Energy Studies and a University Research Lecturer at the Department of International Development, Oxford University. The contents of this paper are the author's sole responsibility.

ISBN

978-1-901795-84-4

CONTENTS

ABSTRACT

This paper investigates the dynamics of the spread between the futures price and the spot price (the basis) in the context of the crude oil market and explores to what extent these are affected by the dynamics of crude oil stocks and OPEC behaviour. The spread between futures prices at various maturities is modeled as a Markov Regime Switching (MRS) process. The estimation method allows us to identify one regime characterised by relatively low volatility and in which the mean basis is positive (contango) while the second regime is characterised by high volatility and in which the mean spread is negative (backwardation). Our results show that the basis exhibits different dynamics within these two regimes and a non-linear relationship between changes in crude oil stocks and the basis where the sensitivity of the basis to changes in crude oil stocks is higher when the market is characterised by low stocks. We also test whether crude oil stocks affect the transition probability of moving from one state to another. Interestingly, we find that an increase in the level of stocks decreases the persistence of staying in the contango regime. This result can be explained in terms of the oil market structure. A rapid accumulation of inventories and rising crude oil stocks to levels which OPEC considers undesirable may induce the Organization to engage in output cuts to trim inventories and change the shape of the forward curve.

Key Words: Convenience yield, Basis, Futures prices, Markov Regime Switching, Oil.
JEL Classifications: G13, C22

1. Introduction

Understanding the variation in the spread between the futures price and the spot price (known as the basis) is important for efficient hedging and for explaining the dynamics of commodity spot prices. Classical studies based on the theory of storage explain the variation in the basis in terms of changes in the fundamentals of supply and demand and/or storage technology of the underlying commodity (Kaldor, 1939; Working, 1948; Brennan, 1958; and Telser, 1958). Other studies explain the variation in the basis in terms of time-varying risk premiums which are influenced by preferences and beliefs of participants in the futures markets (Bailey and Chan, 1993). While the basis is relatively stable when compared to the variability of spot or futures prices, it may exhibit large variability for some commodities and may follow different dynamics depending on the behaviour of stocks of the underlying commodity.

This paper investigates the dynamic behaviour of the basis in the context of the crude oil market and explores to what extent these are affected by the dynamics of crude oil stocks and OPEC behaviour. The focus on crude oil stocks is warranted since changes in their levels are probably the most important source of information influencing short-term movements of crude oil prices and the shape of the oil forward curve. For participants in the futures market, data on crude oil stocks reveal important information about oil market balances in the short run to the medium term. For OPEC, rapid accumulation of crude oil stocks can be a major source of concern as high level of inventories may induce a steep fall in the oil price and may push the organization to take 'pre-emptive' action.[2] Thus, the fact that different market participants monitor very closely the Energy Information Administration's weekly[3] releases of crude oil stock data should come as no surprise, especially that such releases pertain to the largest oil consumer in the world.

In this paper, we model the spread between the futures prices at various maturities as a Markov Regime Switching (MRS) process. Our estimation method allows us to identify two distinctive regimes: one regime characterised by a relatively low volatility and in which the mean basis (measured by the difference between the second month and the nearby month contract) is positive (i.e. contango) while the other regime is characterised by high volatility and in which the mean spread is negative (i.e. backwardation).[4] We show that these two distinct regimes reflect periods of relatively high and low levels of commercial crude oil stocks. Our results indicate that the basis exhibits different dynamics within these two regimes. They also indicate the existence of a non-linear relationship between changes in crude oil stocks and the basis where we find that the sensitivity of the basis to changes in crude oil stocks is higher when the market is characterised by low stocks.

---

[2] The Saudi Oil Minister Ali al-Naimi, the most powerful player within OPEC, argues that "a stock build always concerns us" and that "whenever the stock level is high the price is low and vice versa". He then raises the question: "do you wait until the build-up in inventory and have a precipitous price fall or you take a pre-emptive, proactive course of action?"

[3] Data for OECD are available and published by the International Energy Agency (IEA) but on a monthly basis. Data on crude oil stocks outside the OECD are not readily available on a timely basis.

[4] There are many definitions of backwardation. Strong backwardation refers to a situation in which the spot price is above the futures price. Weak backwardation refers to a situation in which the spot price adjusted for the time value of money is above the futures price. Finally, Keynes introduces normal backwardation which refers to a situation in which the spot price is above the expected spot price. In this paper, we focus only on strong backwardation.

In the final step, we test whether crude oil stocks affect the transition probability of moving from one state to another. Interestingly, we find that an increase in the level of stocks decreases the persistence of staying in the contango regime. This result is inconsistent with storage theory where the rise in the level of inventories is expected to lower the benefit from holding stocks, increasing the futures price and lowering the basis. In other words, high levels of inventories should increase the persistence in staying in contango, or alternatively decrease the probability of moving from contango to backwardation. One way to explain this result is in terms of the oil market structure and particularly the role of OPEC. A rapid accumulation of inventories and rising crude oil stocks to levels which OPEC considers undesirable may induce the Organization to engage in output cuts to trim inventories. Thus, by inducing a reaction by OPEC, high levels of crude oil stocks increase the probability of shifting from the high stock to the low stock regime. This is consistent with some recent anecdotal evidence about OPEC behaviour.

This paper extends a very wide empirical literature that examines the implications of the basis on various aspects including the correlation between spot and futures returns, spot return volatility, forward return volatility, and hedge ratios. For instance, Fama and French (1988) find that when the interest-adjusted basis is positive (an indication that inventories are high), the variances in spot price changes are not statistically different from variances in the three month-forward prices. On the other hand, when the interest-adjusted basis is negative (an indication that inventories are low), the variances of spot prices are lower and statistically different from forward price variances. Ng and Pirrong (1994) provide a comprehensive empirical study of the spot-forward return dynamics of metal prices and their relationship to the basis. They find that spot return volatility and futures return volatility vary directly with the futures–spot price spread and that the volatility of the forward returns relative to the spot return volatility declines as the spread increases. Lin and Duan (2007) provide similar results in the context of the crude oil market.

Rather than looking at the implications of the basis on the spot and forward prices, other studies have examined the variability of the basis over time and across commodities using storage costs, inventories, and seasonal supply and demand variables. Fama and French (1987) test some of the implications of storage theory but only indirectly due to data limitations that precluded the collection of data on inventories on all the commodities in their study. They postulate that seasonality in production or demand for some commodities should generate seasonality in the basis. Fama and French (1988) find that for commodities such as metals which are not subject to seasonality in supply or demand and where storage cost is low relative to its value, the standard deviation of the basis is much lower than other commodities with high storage costs and whose production and consumption are subject to seasonality. The authors also test directly for seasonal effects using seasonal dummies and find such effects only for agricultural commodities and for animal products but find no such effects for metals. Bailey and Chan (1993) model the basis in terms of macroeconomic risks that are common to all asset markets. They show that factors such as the stock index dividend yield and the corporate bond quality spread can explain a large portion of the variation in basis. They claim that such variables generate time varying risk premiums in commodity markets.

Other papers explain the variation in the basis in terms of the stocks of the underlying commodity (Telser, 1958; Geman and Nguyen, 2005; Sorensen, 2002; Carter and Giha, 2007). For instance, Sorenson (2002) uses a Kalman Filter approach to estimate a time series of convenience yields associated with futures prices and then regresses the derived convenience yields on a measure of relative inventory. In support of storage theory, he finds a negative relationship between inventories and the convenience yield. Carter and Giha (2007) revisit the original data that Working (1948) used to derive his inventory-spread curve and which formed the empirical basis for most of the subsequent theoretical literature. Controlling for the problem of potential spatial aggregation, they find that wheat stocks are carried under backwardation in a single location lending support to Working's original empirical work. The closest paper to this one is Heaney (2000) where he applies a Markov Regime Switching to an adjusted basis for three types of metals (zinc, copper and lead). Heaney (2000) finds two regimes which he labels the value in storage and value in consumption. His results indicate that inventories play an important role in explaining the basis in both of these regimes though the sensitivity of the basis to inventories varies across the two regimes.

This paper is divided in six parts. Section 2 provides a brief literature review. Section 3 describes the empirical method while Section 4 describes the dataset. Section 5 reports the empirical results. The last section concludes.

2. The Relationship between Basis and Inventories: A Literature Review

One way to explain the basis is in terms of the existence of a risk premium which arises in the process of transferring risk from hedgers to speculators. Specifically, the basis can be written as the sum of the following two main components

$$F_t^T - S_t = E_t(S^T - S_t) + E_t(\pi_t^T) \qquad (1)$$

where F_t^T is the futures price at time t for a futures contract maturing at time T, S_t is the spot price at time t, S^T is the realisation of the spot price at maturity, E_t is the expectation operator and π_t^T is the realised risk premium. The first term $E_t(S^T - S_t)$ is simply the expected change in the spot price. The second term $E_t(\pi_t^T)$ is the ex-ante risk premium. The risk premium can be positive or negative (and hence the basis can take negative or positive values) depending on investors' beliefs, endowments, and preferences (Bailey and Chan, 1993). [5]

An alternative theory, the theory of storage, explains the basis without resorting to the concept of risk premium. [6] The theory of storage comes in more than one version. One version explains the difference between the futures price and the spot price of a commodity in terms of interest foregone in purchasing and storing the commodity,

[5] The expected premium is defined as the bias of the futures price as a forecast of the future spot price (Fama and French, 1987).
[6] Fama and French argue that theories based on risk premium and theories based on storage are "alternative but not competing views of the basis". In fact, the first term in equation (1) reflects the storage cost, the time value of money and the expected convenience yield (Bailey and Chan, 1993).

storage costs, and the convenience yield (for recent studies see Miranda and Glauber (1993) and Miranda and Rui (1996)). The latter is defined as a yield or benefit that "accrues to an owner of physical asset but not to an owner of a contract for future delivery of the commodity". The yield may not necessarily be a pecuniary return. It can be an implicit and indirectly measurable return that holders place on their ability to use inventories to meet contractual obligations, to minimise transformation costs, to prevent interruption of supplies to long-standing customers, and to respond to demand shocks. The convenience yield affects the basis through arbitrage. When the convenience yield goes up, the attractiveness of holding futures contracts relative to physical stocks goes down as holding the futures contracts does not accrue any benefits. This will lower the futures price and increase the spot price until the following familiar relationship is attained

$$F_t^T = S_t e^{(r+c-y)(T-t)} \qquad (2)$$

where F_t^T is the futures price at time t for a futures contract requiring delivery at time T, S_t is the spot price at time t, r is the continuously compounded interest rate prevailing at which funds can be borrowed, c is marginal storage cost of the physical commodity per unit for the period from the purchase of the commodity until the delivery time, and y is the marginal convenience yield. Thus, in equilibrium, backwardation implies the existence of convenience yield.

Studies based on the storage model relate the convenience yield directly to the level of inventories. Generally, the theory of storage suggests that marginal convenience yield falls with inventory but at a decreasing rate (Telser, 1956; Brennan, 1956; Fama and French, 1988). At low levels of inventory, the marginal convenience yield is larger than carrying costs and the futures–spot price spread is negative. As the level of inventories goes up, the marginal convenience yield falls towards zero and the futures–spot price spread becomes positive and converges towards the cost-of-carry. Pindyck (1994) suggests a convex relationship between the convenience yield and stock levels with the marginal convenience yield rising rapidly as inventories approach zero and remaining close to zero over a wide range of moderate to high stocks. Some models consider a non-linear relationship with the marginal convenience yield rising at low level of inventories and then declining in a non-linear manner to zero. At sufficiently high level of inventories, the marginal storage becomes increasingly expensive as storage facilities reach full capacity levels and the marginal benefit from adding stocks becomes zero (Larson, 1994).[7]

[7] The convenience yield has been used extensively in modelling commodity prices. Gabillon (1991), Gibson and Schwartz (1990), Schwartz (1997) consider a two-factor model where the commodity spot price is modelled as a geometric Brownian motion (GBM) and convenience yield as a stochastic mean reverting process which enters the drift of the spot price. The two-factor model has been extended to include a third stochastic variable for the interest rate to generate three-factor models (Hilliard and Reis, 1998; Miltersen and Schwartz, 1998). In these models, the convenience yield is treated as an exogneous variable. However, some recent studies have introduced inventories into the two-factor models with some success. For instance, Ribeiro and Hodges (2004) propose a model in which commodity prices switch between two distinct stochastic processes depending on level of inventories. When inventories are high, the spot price follows GBM with a drift equal to the cost of carrying inventories. When inventories are low, the spot price follows an Ornstein-Uhlenbeck process. The authors show that this specification can explain better some of the properties of the forward curve.

Using the insights of option theory, more recent models describe the convenience yield as a financial call option held by storage agents (Heinkel et al, 1990; Milonas and Thomadakis, 1997). The call option can have value due to a variety of reasons and that the value of the call option may be high enough to justify holding inventories when costs-of-carry are high and/or prices are expected to fall. Some authors argue that the option would be valuable whenever the probability of stock-out is non-zero (Milonas and Thomadakis, 1997). Others emphasise that the call option will have value whenever demand shocks create the probability that agents can sell their stocks at higher price during the storage period (Heinkel, Howe and Hughes, 1990; Larson, 1992). If the current price of the commodity exceeds the sum of the purchasing price and the costs of carry, then the option is 'in the money' and the value of holding inventories is positive. In line with the theory of storage, studies that model the convenience as a call option assume the value of the option to be dependent on stock levels with higher stocks reducing the value of the call option. For instance, in Larson's model, when inventories are low, the range of possible price outcomes becomes more skewed towards higher prices. By carrying inventories during this period, agents possess a more valuable option because they can take advantage of higher prices if they materialise.[8]

The concept of convenience yield has been criticised as not being derived from optimising conditions but rather introduced heuristically into the storage model (Williams and Wright, 1991; Deaton and Laroque, 1996; Brennan, Williams and Wright, 1997). More recent studies propose a competitive rational expectation model that generates an embedded timing option with implications on spot/futures pricing (Routledge, Seppi and Spatt, 2000). The value of the timing option arises due to the possibility of stock-out. When prices rise faster than the costs of carrying the commodity, then it is optimal to store a commodity for future consumption. In this case, the commodity is priced like a financial asset and in equilibrium the difference between the spot price and futures price would only reflect the cost of carrying the commodity. Conversely, if spot prices are expected to fall or spot prices are expected to rise less than storage costs, then it is optimal to sell inventories. However, discretionary stocks can only be reduced to zero: Traders may wish to sell more inventories but they are physically constrained. This non-negativity constraint which implies a non-zero probability of current or future stock-out creates a valuable timing option.[9] In terms of pricing, when a stock-out occurs, the commodity price would be linked only to the good's immediate use and thus commodities should be priced like consumption goods. In a high demand state and in the presence of stock-out, the relationship between spot prices and futures prices is broken and the spot price can rise relative to the forward price resulting in backwardation (Routledge, Seppi and Spatt, 2000).

Although the stock-out theory does not rely on an explicit convenience yield, it has similar implications regarding the role of inventories (Ng and Pirrong, 1994).[10] When

[8] Models that incorporate options do not focus only on inventories but also introduce other variables such as volatility that may affect the value of the call option.

[9] In Routledge, Seppi and Spatt (2000), the value of the option varies over time and is affected by inventory levels and exogenous transitory supply-demand shocks.

[10] The timing option refers to the ability of the holder to consume the commodity in high demand states and buy it back at a lower expected price in the future. This benefit could be expressed in terms of

inventories are high, the current or the future probability of stock-out is low and thus we are in the ‘value in-storage’ state. As inventories decline, the probability of stock-out increases and the probability of entering the value in consumption state becomes higher (Heaney, 2000).

In short, the literature suggests the following two general observations. First, regardless of the theory adopted, the behaviour of inventories is a key factor in explaining the variation in the basis over time. Second, the basis is likely to follow different dynamics depending on whether stocks are low or high. In what follows, we test these implications in the context of the crude oil market and using a Markov Regime Switching model.

3. Empirical Method

In the Markov Switching Model, the regression coefficients and the variance of the error term are assumed to be state-dependent. Let y_t be a vector of dependent variable and X_t a matrix that includes a constant, lags of y_t and other set of exogenous variables. Assuming only two states of the world, we can write the following two conditional distribution functions:

$$\begin{aligned} &f(y_t \mid \theta_1, X_t) \ if \ s_t = 1 \\ &f(y_t \mid \theta_2, X_t) \ if \ s_t = 2 \end{aligned} \qquad (3)$$

Where θ_1 and θ_2 contain the parameters of the model to be estimated and s_t defines whether we are in regime 1 or 2. The state variable s_t which determines the conditional distribution evolves over time as a discrete time, discrete space Markov chain process[11] with the following transition probabilities

$$P[s_t = 1 \mid s_{t-1} = 1] = p_{11} \qquad P[s_t = 2 \mid s_{t-1} = 1] = 1 - p_{11} = p_{12} \qquad (4)$$

where p_{11} is the probability of remaining in regime 1 at time t given that y_t is in regime 1 at $t-1$; p_{22} is the probability of remaining in regime 2 at time t given that y_t is in regime 2 at $t-1$. The transitional probabilities from switching from one regime to another are given by p_{12} and p_{21}. Based on transition probabilities, it is possible to compute the unconditional probability of being in any of the regimes. For instance, the unconditional probability of being in regime 1 is given by

$$\bar{p} = \frac{p_{12}}{1 - p_{11} + p_{12}}$$

In the basic model, the transition probabilities are assumed to be fixed i.e. transitional probability is not allowed to vary over time. This is quite a restrictive assumption and it is desirable to endogenise transition probabilities and allow them to vary over time.

---

convenience yield but unlike the classical storage model, the convenience yield is an “out of the model rather than an input” (see Routledge, Seppi and Spatt, 2000 for details).

[11] The Markov property implies that the probability of being in a certain regime at time t depends only on the state at time $t-1$. States at period $t-2$, $t-3$ and so on don’t affect the transition probability.

Following Diebold et al (1999) and Filardo (1994), we allow the probability of switching from one regime to another to be a function of z_{t-1} where z_{t-1} is a conditioning vector containing economic information that is expected to affect the state transition probabilities. Specifically, we write (5) as:

$$P[s_t = 1 \mid s_{t-1} = 1] = p_{11}(z_{t-1}) \qquad P[s_t = 2 \mid s_{t-1} = 1] = 1 - p_{11}(z_{t-1})$$
$$P[s_t = 2 \mid s_{t-1} = 2] = p_{22}(z_{t-1}) \qquad P[s_t = 1 \mid s_{t-1} = 2] = 1 - p_{22}(z_{t-1})$$

The transition probabilities of moving from one regime to another are modelled as logistic functions of z_{t-1} such that (see Diebold, Lee and Weinbach (1994) for details):

$$\begin{aligned} P[s_t = 2 \mid s_{t-1} = 2, z_{t-1}; \beta_1] &= 1 - \frac{\exp(z'_{t-1}\beta_1)}{1+\exp(z'_{t-1}\beta_1)} \\ P[s_t = 2 \mid s_{t-1} = 1, z_{t-1}; \beta_2] &= 1 - \frac{\exp(z'_{t-1}\beta_2)}{1+\exp(z'_{t-1}\beta_2)} \end{aligned} \qquad (5)$$

where β_1 and β_2 are the vector of coefficients measuring the impact of the conditioning vector on the probability of moving from one regime to another. The variances, the parameters of each state, the transitional probabilities and their determinants are jointly estimated using the maximum likelihood method (see Filardo, 1994 and Diebold et al, 1999 for details).

4. Data and Descriptive Statistics

We use weekly data from 5 January 1990 to 26 December 2008 on the basis and US crude oil stocks. Data on spot oil prices do exist but as noted by Pindyck (1994), spot prices do not necessarily reflect actual transactions on the day and thus cannot be matched with prices on futures contracts. Instead, we use futures price of NYMEX Light Sweet crude oil on maturing contracts (the first nearby or the first-month contract) as proxy for the spot price. The basis ($BASIS_{21}$) is calculated as the log difference in the settlement prices between the second month and the first month futures contract.[12]

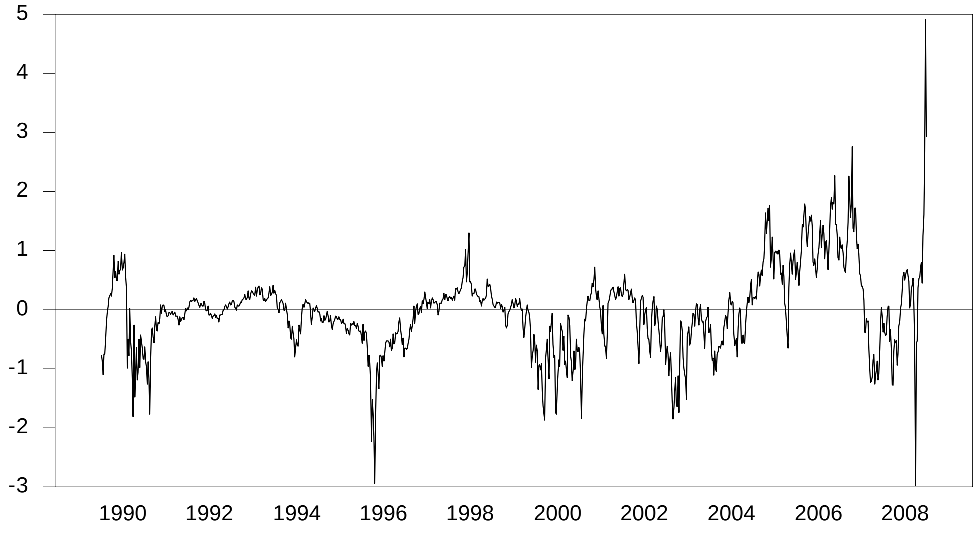
Crude oil stocks are total crude oil stocks in the USA excluding the Special Petroleum Reserve (SPR). The series (STOCKS) is obtained from the Energy Information Administration (EIA) website. It is important to stress that the data refer only to what are known as primary stocks. Primary stocks encompass crude oil stocks in refining and storage facilities of the industry such as crude oil in export and import terminals, in distribution terminals, in refinery columns, and in specific large storage facilities. The data do not cover secondary stocks which usually refer to stocks outside the main terminals or tertiary stocks which are owned by consumers. For crude oil, the distinction between the various types of stocks is less important than for petroleum

[12] For robustness, we also consider the log difference between the the third month and the first month futures contract ($basis_{31}$) and the log difference between the fourth month and the first month futures contract ($basis_{41}$). The results are very similar and are not reported here but are available from the author upon request.

products as primary stocks constitute the bulk of crude oil stocks. This is in contrast to petroleum products where secondary stocks (such as stocks with wholesale distributors, marketers, small distribution centres) and territory stocks (such as fuel oil in power stations, utilities, gasoline in trucks, and industrial and commercial stocks) constitute a large fraction of total petroleum products stocks.

The figure below shows the weekly basis (not in logs) over the period 1990 to 2008. It reveals two interesting observations. The first observation concerns the high volatility of the basis, especially towards the end of the sample. The second observation concerns the frequency in which the basis is in contango. In an influential article, Litzenberger and Rabinowitz (1995) noted that the nine month futures crude oil prices are below the crude oil spot price (i.e. the market is backwardated) almost 80% of the time. However, examining more recent data and focusing on the very front part of the futures curve (i.e. the difference in prices between the second month and the first month futures contract), we obtain a very different picture. As can be seen from this figure, since 1997, the oil market has witnessed many switches from backwardation to contango and in two occasions, the crude oil market entered into two sessions of prolonged contango which lasted for a number of weeks. The first occurred around the first quarter of 1997 and lasted until the third quarter of 1999. A more recent episode occurred towards the last quarter of 2004 and lasted until the mid of 2007. This raises a number of questions: can the behaviour of inventories explain such switches and long episodes of contango? Second, given OPEC’s central role in the market and recent claims that the Organization has shifted to targeting inventories, to what extent does OPEC influence the shape of the basis through its impact on inventories?

Figure 1: Weekly Basis (Front Month minus Second Month)



Source: Energy Information Administration Website

Table 1 below reports some basic statistics for the basis and crude oil inventories. The mean of the basis is positive but with a very high standard deviation. The skewness of the basis is positive suggesting a heavier right tail while excess kurtosis is very high suggesting fat-tailed distribution. The Jarque-Bera test rejects the null hypothesis that the basis is normally distributed. Regarding stocks, the mean of level of stocks stood at 319,250 thousand barrels with a high standard deviation of 23,639 thousand barrels.

The minimum and maximum range of crude oil stocks varied from around 264,000 thousand barrels to 392,000 thousand barrels over the sample period. There is no evidence for skewness or kurtosis and the Jarque-Bera test cannot reject the null hypothesis of normality.

Table 1: Summary Statistics

	Sample Mean	**Standard Error**	**Min**	**Max**	**Skewness**	**Kurtosis (excess)**	**J-B**
$BASIS_{21}$	-0.241	2.17	+11.67	-12.84	-0.38	3.57	553.15***
STOCKS	319,250	23,639	263,666	391,907	0.10	-0.20	3.37

Table 2 below reports the correlation matrix between the basis based on different maturities, the logarithm of stocks (LSTOCKS) and the change in logarithm of stocks (ΔLSTOCKS). As expected, the correlation between the spreads at various maturities is close to 1 and highly significant. The correlation between the basis and logarithm of stocks is negative and highly significant. On the other hand, correlation between the (ΔLSTOCKS) and the various spreads is close to zero and marginally significant.

Table 2: Correlation Table

	$BASIS_{21}$	**$BASIS_{31}$**	**$BASIS_{41}$**	**LSTOCKS**	**INTRATE**
$BASIS_{21}$	1				
$BASIS_{31}$	0.986 (0.000)	1			
$BASIS_{41}$	0.964 (0.000)	0.994 (0.000)	1		
LSTOCKS	0.445 (0.000)	0.468 (0.000)	0.483 (0.000)	1	
ΔLSTOCKS	0.042 (0.177)	0.050 (0.110)	0.056 (0.077)	0.082 (0.009)	1

Significance Levels in parentheses; Number of observations = 990

5. Empirical Analysis and Results

5.1 Unit Root Tests

As an initial step in our econometric analysis, we test for non-stationarity of the basis and the logarithm of stocks. The augmented Dickey-Fuller (ADF), Philips-Perron (PP), DF-GLS test (Ng and Perron, 2001), and the Zivot-Andrews unit root test indicate that we can reject the null hypothesis of unit root for the BASIS. The results for LSTOCKS are not clear cut where the ADF and DF-GLS suggest that we can't reject the null of unit root, while the Philips-Perron and the Zivot-Andrews Unit Root test indicate that we can reject the null of unit root but only at the 10% level. The Kwiatowski, Phillips, Schmidt & Shin (1992) Unit Root Test (KPSS), which tests for the null of stationarity around a level, indicate that we can reject the null of stationarity for LSTOCKS but not for the basis. Based on the various unit root tests, we conclude that the basis is stationary while the logarithm of stocks is integrated of order 1.

Table 3: Unit Root Tests

	ADF	DFGLS	PP	Z-V	KPSS
$BASIS_{21}$	-5.080	-2.657	-6.012	-5.785	0.699
LSTOCKS	-2.721	-1.670	-2.912	-4.710	6.359
ΔLSTOCKS	-12.944	-12.061	-31.317	-20.018	0.024

Number of lags was selected using the BIC criterion.
Critical values for ADF: 1% = -3.440; 5% = -2.865; 10% = -2.569.
Critical values for DFGLS: 1%: -2.58; 5% = -1.95; 10% = -1.62.
Critical values for PP: 1% = -3.440; 5%= -2.865; 10%= -2.569.
Critical values for Zivot-Andrews Unit Root Test: 1% = -5.34; 5% = -4.80
Critical values for KPSS: 10%= 0.347; 5%= 0.574; 1% = 0.739.

5.2 Empirical Results

We estimate the MRS model of equation (3) with X_t containing a constant term and three lags of the basis.[13] [14] The results are reported in Table 4. They reveal some interesting findings. First, the mean basis in state 1 is 0.033/(1-0.929+0.135-0.147) = 0.55 while the mean basis in state 2 is -0.162/(1-0.794-0.00-0.095) = -1.73. Thus, state 1 corresponds to weeks where the mean basis is positive (i.e. contango) while state 2 corresponds to weeks where the mean basis is negative (i.e. backwardation). Second, the results suggest that the variance in the error terms in state 2 is almost four times higher than that in state 1. In other words, the basis exhibits greater volatility when the market is in backwardation. This is expected as backwardation is associated with low and declining stocks. Third, the transition probabilities estimates indicate differences across the different states where it is more likely for the basis to get out from backwardation into contango than the other way around. Fourth, the estimated transition probabilities indicate strong persistence within each of the regimes. Specifically, the expected duration of the contango and backwardation are approximately 32 weeks and 14 weeks respectively.[15] Finally, the estimated coefficients across the two states are statistically different implying that dynamics of the basis are different between contango and backwardation.

[13] The number of lags has been chosen to ensure there is no residual correlation.
[14] We also included lags of US 3-month treasury-bill rate in the regression but the coefficients were found to be insignificant in both regimes and hence the interest rate was dropped from the equation.
[15] The expected duration of regime j is given by: $E(D=j)=\sum_{i=1}^{\infty} i p_{jj}^{i-1}(1-p_{jj})=\frac{1}{1-p_{jj}}$.

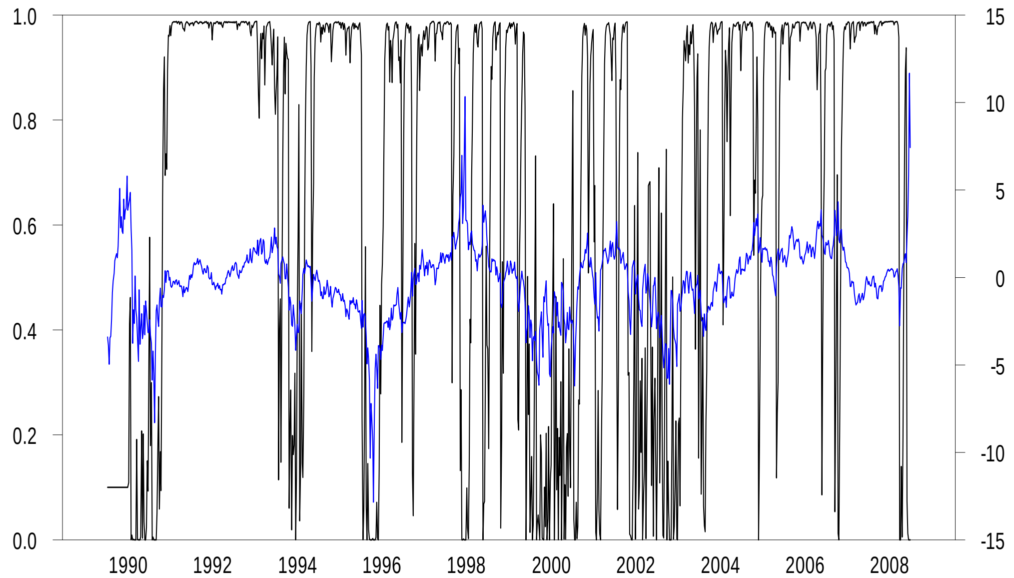
Table 4: Markov-Regime Switching Results

Regime 1		*Regime 2*	
σ_1	0.407 (0.041)	**σ_2**	1.445 (0.145)
P_{12}	0.031 (0.014)	**P_{21}**	0.069 (0.023)
Intercept	-0.033 (0.017)	**Intercept**	0.162 (0.096)
y_{t-1}	0.929 (0.038)	**y_{t-1}**	0.811 (0.033)
y_{t-2}	-0.135 (0.054)	**y_{t-2}**	-0.000 (0.025)
y_{t-3}	0.147 (0.042)	**y_{t-3}**	0.095 (0.047)

Notes:
Robust standard error in parentheses
Coefficient σ_1 and σ_2 denotes the variance of the error term in regime 1 and regime 2 respectively.
P_{12} is probability of moving from regime 1 (low volatility regime) to regime 2 (high volatility regime).
P_{21} is probability of meeting from regime 2 (high volatility regime) to regime 1 (low volatility regime).

Figure 2 shows the filtered probability of being in regime 1 and $BASIS_{21}$. As can be seen from this figure, the probability that the basis is in the low volatility regime varies frequently over the weeks but spends most of its time close to one or zero. The probability that basis is in contango seems to match the broad trends in the actual variation of the basis. Specifically, the model captures adequately the main episodes when the market entered in long periods of contango. It also captures the latest switch into contango towards the end of the sample.

Figure 2: Filtered Probability of Being in Regime 1 (Table 4)



In the above analysis, inventories did not play a role in the identification of the two regimes. To explore the interaction between inventories and the basis, we check whether the two states reflect periods of low stocks versus period of high stocks. Table 5 below compares the level of stocks when the market is in state 1 and when it is in state 2. As can be seen from this table, the level of stocks is higher in regime 1 than the level of stocks in regime 2. Based on the *t-test* and the non-parametric Mann-Whitney test, we find that the difference across the two regimes is statistically significant. Interestingly, the standard deviation in regime 2 is much higher than that of regime 1. Based on this analysis, it is possible to stipulate the MRS splits into two states: the high inventory state (regime 1) and the low inventory state (regime 2). In the high inventory state, market participants receive little benefit from holding stocks and thus the convenience yield is likely to be zero and the basis will only reflect the cost of holding the inventory. Thus, the mean basis for regime 1 reflects the cost of holding inventories (both financing and storage costs). In the second regime, when inventories are low, the cost of holding inventories is more than offset by the existence of convenience yield and hence the negative mean basis.

Table 5: Level of Stocks across the Two States

	Level of Stocks (Thousands barrels)	***Standard Deviation***
Regime 1	320626.6	19714.74
Regime 2	311899.8	27231.25
t-test	-5.582***	
Mann-Whitney Test	*5.822****	

5.3 Basis and the Role of Inventories

Much of the empirical evidence assumes that the sensitivity of the basis to the change in inventories is the same across contango and backwardated markets. The MRS model allows us to test whether the impact of inventories differ across the two regimes. To do so, we re-estimate our original model with lagged values of changes in crude oil stocks. Models based on the convenience yield suggest that increases in crude oil inventories would lower the convenience yield and hence lower the basis. Specifically, the convenience yield is modelled in terms of lagged changes in the level of stocks

$$cy = \alpha + \sum_{k=1}^{k} \Delta \ln(stocks_{\text{t-k}})$$

where $\Delta \ln(stocks)$ is the change in the natural log of the level of stocks. Thus, equation (3) takes the following form:

$$BASIS = \alpha + \sum_{t=1}^{k} BASIS + \sum_{t=1}^{k} \Delta \ln(stocks_{\text{t-k}}) \qquad (6)$$

As before, we include lags of basis in the estimation. The results are reported in Table 6 below. The MRS regime splits the model into regimes (contango and backwardation). The estimates change very little suggesting that our model is robust to the inclusion of changes in log stocks. As expected, increases in lagged stocks of crude oil lead to increases in the spread. Interestingly, we find that the sensitivity of the basis to change in stocks differs across the two states where the coefficients on lagged stocks are different. This suggests the existence of non-linearity in the relationship between change in inventories and the basis. Specifically, small changes in crude oil stocks have a greater impact on the basis when the market is in backwardation (see Heaney (2000) for similar results for metals). This is expected as in the low stock regime a decrease in inventories would raise the convenience yield and depress the futures price resulting in large movements in the basis. In the high stock regime, changes in inventories will affect the spread insomuch as it affects the cost of the carrying which is usually bounded.

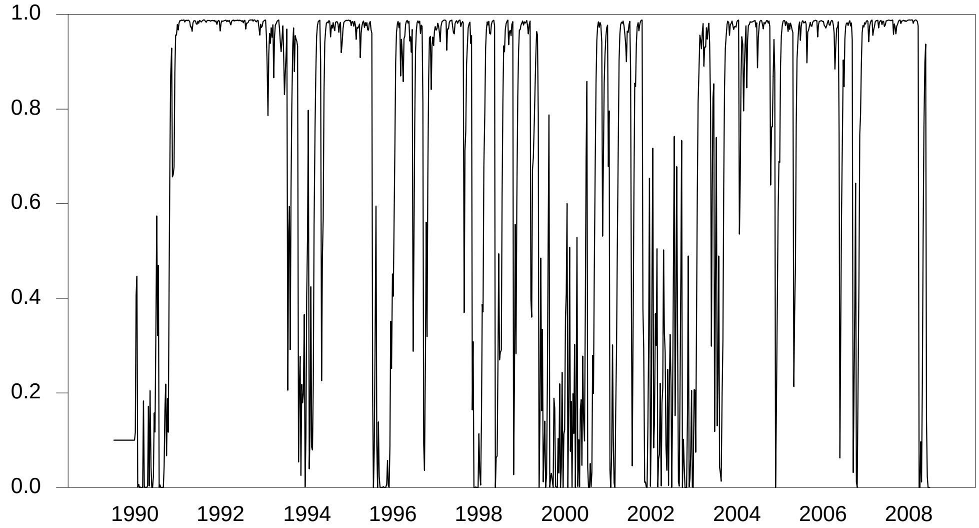
Table 6: Markov-Regime Switching Results

Regime 1		***Regime 2***	
$\boldsymbol{\sigma_1}$	0.405 (0.031)	$\boldsymbol{\sigma_2}$	1.435 (0.138)
$\mathbf{P_{12}}$	0.029 (0.011)	$\mathbf{P_{21}}$	0.059 (0.020)
Intercept	0.031 (0.017)	**Intercept**	-0.140 (0.103)
$\mathbf{y_{t-1}}$	0.917 (0.043)	$\mathbf{y_{t-1}}$	0.804 (0.048)
$\mathbf{y_{t-2}}$	-0.127 (0.039)	$\mathbf{y_{t-2}}$	-0.017 (0.061)
$\mathbf{y_{t-3}}$	0.146 (0.061)	$\mathbf{y_{t-3}}$	0.118 (0.051)
Δln(Stocks$_{t-1}$)	4.542 (1.341)	**Δln(Stocks$_{t-1}$)**	14.149 (6.965)
Δln(Stocks$_{t-2}$)	2.742 (1.559)	**Δln(Stocks$_{t-2}$)**	0.389 (7.471)
Δln(Stocks$_{t-3}$)	0.605 (1.253)	**Δln(Stocks$_{t-3}$)**	8.223 (7.578)

Notes:
Robust standard error in parentheses
Coefficient σ_1 and σ_2 denotes the variance of the error term in regime 1 and regime 2 respectively.
P_{12} is probability of moving from regime 1 (low volatility regime) to regime 2 (high volatility regime).
P_{21} is probability of meeting from regime 2 (high volatility regime) to regime 1 (low volatility regime).

Figure 3 shows the filtered probability of being in regime 1. The inclusion of changes in log stocks has little impact on the identification of the two regimes. In fact, the correlation between the filtered probability of model (1) and filtered probability of model (2) is quite high at 0.99.

Figure 3: Filtered Probability of Being in Regime 1 (Table 6)



5.4 Time Varying Transition Probabilities

Rather than examining the impact of inventories on the basis, we test whether crude oil stocks affect the probability of moving from one regime to another. In other words, rather than treating the transition probabilities as fixed we endogenise the transitional probabilities by making them dependent on behaviour of crude oil stocks. To implement this, the transition probabilities of moving from one regime to another are modelled as a logistic function of lagged changes in the level of stocks. The model's parameters are then estimated jointly using the maximum likelihood method. The results reported below indicate that lagged change in the level of stocks does not have a significant impact on transitional probabilities.

Table 7: Estimates of Time Varying Transition Probabilities

Regime 1		***Regime 2***	
σ_1	0.414 (0.037)	σ_2	1.451 (0.152)
Intercept	0.032 (0.015)	Intercept	-0.114 (0.097)
y_{t-1}	0.930 (0.035)	y_{t-1}	0.810 (0.042)
y_{t-2}	-0.121 (0.052)	y_{t-2}	0.008 (0.059)
y_{t-3}	0.131 (0.043)	y_{t-3}	0.100 (0.053)
	Time Varying Transitional Probability		
$\Delta LSTOCKS_{t-1}$		$\Delta LSTOCKS_{t-1}$	
Intercept	3.49 (0.44)	Intercept	2.86 (0.52)
Slope Coefficient[1]	-4.04 (30.04)	Slope Coefficient[2]	-50.23 (44.94)

Notes:

Robust standard errors in parentheses. Slope coefficient measures the effect of $\Delta\ln(Stocks_{t-1})$ on the probability of moving from regime 1 (low volatility regime) to regime 2 (the high volatility regime). Slope coefficient measures the effect of $\Delta\ln(Stocks_{t-1})$ on the probability from moving from regime 2 (high volatility regime) to regime 1 (the low volatility regime).

We next examine the impact of the level of crude oil stocks on transition probability. After all, for OPEC, it is the level of inventories rather than the change in the level of inventories that can induce the Organization to react by adjusting output. To test for this hypothesis, we model the transition probabilities of moving from one regime to another as logistic function of lagged level of crude oil stocks. The results are reported in Table 7 . As before, the Markov Regime Switching estimation method is able to separate between the two regimes. The estimated coefficients are very similar to those obtained in Table 1 and need no further commenting. In the second part of the table, we report the impact of the level of crude oil stocks on the probability of switching from one regime to another. As expected, an increase in the level of inventories increases the probability of moving from backwardation to contango; however the estimated coefficient is not significant at the conventional levels. Similarly, we find that an increase in crude oil stocks increases the probability of moving from contango to backwardation with a significant estimated coefficient at the 1% level. This result is at odds with storage theory. According to this theory, a rise in inventories should increase the probability of staying in the contango regime.

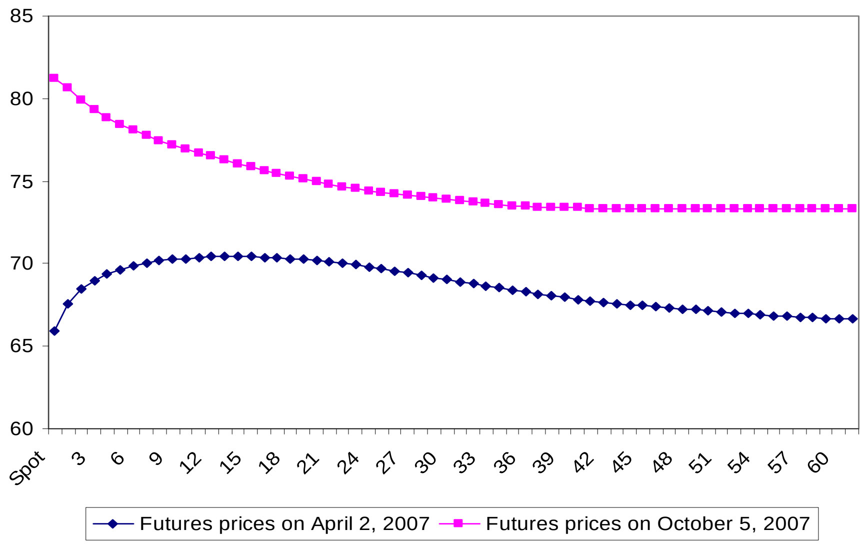
Table 8: Estimates of Time Varying Transition Probabilities

Regime 1		***Regime 2***	
σ_1	0.398 (0.036)	σ_2	1.449 (0.148)
Intercept	0.036 (0.019)	Intercept	-0.114 (0.081)
y_{t-1}	0.933 (0.042)	y_{t-1}	0.809 (0.048)
y_{t-2}	-0.139 (0.058)	y_{t-2}	0.014 (0.061)
y_{t-3}	0.144 (0.043)	y_{t-3}	0.097 (0.054)
	Time Varying Transitional Probability		
LSTOCKS$_{t-1}$		*LSTOCKS*$_{t-1}$	
Intercept	-107.03 (36.692)	Intercept	-32.649 (38.556)
Slope Coefficient[1]	8.695 (2.911)	Slope Coefficient[2]	2.761 (3.047)

Notes: Robust standard errors in parentheses. Slope coefficient measures the effect of LSTOCKS$_{t-1}$ on the probability of moving from regime 1 (low volatility regime) to regime 2 (the high volatility regime). Slope coefficient measures the effect of LSTOCKS$_{t-1}$ on the probability from moving from regime 2 (high volatility regime) to regime 1 (the low volatility regime).

One possible way to explain this result is in terms of OPEC behaviour. High levels of stocks may increase the incentive for OPEC to engage in output cuts if the Organization feels that high levels of stocks can induce a sharp downturn in oil prices. OPEC cuts would have the effect of lowering inventories and raising the price at the front end of the futures curve increasing the probability of the basis moving back into backwardation. There is some anecdotal evidence to support this explanation. During the period 2004–2006, inventories were accumulating at a very fast rate. By the end of 2006, crude oil inventories in the USA stood at 321 million barrels, 25 million barrels above the five-year average. Since 2004, OPEC officials have been conveying their strong concerns about the high build-up of inventories in the USA and other OECD countries. Their main concern was that the release of large stocks of crude oil can flood the market with the effect of driving oil prices downwards to levels they consider unacceptable. The sharp fall in oil prices from above $77 in July 14, 2006 to around $50.4 in January 18, 2007 prompted OPEC's decision to cut output in mid 2007. This had the effect of changing the shape of the curve from contango to backwardation as shown in Figure 4.

Figure 4: A Shift from Partial Contango to Full Backwaradation



Source: IMF, World Economic Outlook: Globalization and Inequality, October 2007, Figure 1.9

6. Conclusion

One of the very interesting features in the recent behaviour of crude oil prices has been the increase in the variability of the basis and the more frequent switches between backwardation and contango at the front end of the forward curve. This paper uses a Markov Regime Switching model to explore the dynamic behaviour of the basis in the crude oil market. This approach allows us to identify two regimes with the basis following different dynamics in each of the two regimes. We find that inventories play a major role in driving these dynamics providing support for the insights from storage theory though the impact of inventories varies depending on whether we are in a high stock or a low stock regime. On the other hand, given the structure of the oil market and the importance of OPEC in this market, some of the switches in the forward curve and the variability of the basis can be explained by OPEC behaviour. The paper suggests that further research is needed to explore the interplay between OPEC and the market to gain a clearer picture of the role of OPEC in influencing the shape of the forward curve.

References

Bailey, W. and K. C. Chan (1993). Macroeconomic Influences and the Variability of the Commodity Futures Basis, *Journal of Finance*, 48(2): 555–73.

Brennan, D., J.C. Williams, and B.D. Wright (1997). Convenience Yield without the Convenience: A Spatial-Temporal Interpretation of Storage under Backwardation, *Economic Journal*, 107(443):1009–22.

Brennan, M. (1958). The Supply of Storage, *American Economic Review*, 48:50–72

Carter, Colin A. and Giha, Cesar L. Revoredo (2007). The Working Curve and Commodity Storage under Backwardation, *American Journal of Agricultural Economics*, 89(4):864–72.

Deaton, A., and G. Laroque (1996). Competitive Storage and Commodity Price Dynamics, *The Journal of Political Economy*, 104(5): 896–923.

Dickey, D. A. and W. A. Fuller (1979). Distribution of Estimates for Autoregressive Time Series with a Unit Root, *Journal of American Statistical Association*, 24:427–31.

Diebold, F.X., Lee, J.-H. and Weinbach, G. (1994). Regime Switching with Time-Varying Transition Probabilities, in C. Hargreaves (ed.), Nonstationary Time Series Analysis and Cointegration. (Advanced Texts in Econometrics, C.W.J. Granger and G. Mizon, eds.), 283-302. Oxford: Oxford University Press.

Fama, E. and K. French (1987). Commodity Futures Prices: Some evidence on forecast power, premiums, and the theory of storage, *Journal of Business*, 60(1):55–73.

Fama, E. and K. French, (1988). Business Cycles and Behavior of Metals Prices, *Journal of Finance*, 43:1075–93.

Filardo, Andrew J. (1994). Business-Cycle Phases and Their Transitional Dynamics, *Journal of Business & Economic Statistics, American Statistical Association,* 12(3): 299-308.

Gabillon, J. (1991). The Term Structure of Oil Futures Prices. Working Paper M17, Oxford Institute of Energy Studies.

Geman, H. and V. Nguyen (2005). Soybean Inventory and Forward Curve Dynamics, *Management Science*, 51(7):1076–91.

Gibson, R. and E.S. Schwartz (1990). Stochastic Convenience Yield and the Pricing of Oil Contingent Claims, *Journal of Finance*, 45:959–76.

Heaney, R. (2002). Approximation for Convenience Yield in Commodity Futures Pricing, *The Journal of Futures Markets*, 22:1005–17.

Heinkel, R., Howe, M.E. and Hughes, J.S. (1990). Commodity: The convenience yield as an option profit, *The Journal of Futures Markets*, 10:519–33.

Kaldor, N. (1939) Speculation and Economic Stability, *The Review of Economic Studies*, 7(1):27.

Kwiatkowski, D., P. C. Phillips, P. Schmidt, and Y. Shin. Testing the Null Hypothesis of

Larson, D.F. (1994). Copper and the Negative Price of Storage, PRWP 1282, International Economics Department, World Bank, Washington, D.C.

Lin, T., and C. W. Duan (2007). Oil Convenience Yields Estimated under Demand/Supply Shock, *Review of Quantitative Finance and Accounting* 28:203–225.

Litzenberger, R.H. and Rabinowitz, N. (1995). Backwardation in Oil Futures Markets: theory and empirical evidence, *The Journal of Finance*, 50:1517–45.

Milonas, N. and Thomadakis, S. (1997). Convenience Yields as Call Options: An empirical analysis, *The Journal of Futures Markets*, 17:1–15.

Miltersen, K. R. and E. S. Schwartz (1998). Pricing of Options on Commodity Futures with Stochastic Term Structures of Convenience Yields and Interest Rates, *The Journal of Financial and Quantitative Analysis*, 33(1):33–59.

Miltersen, K.R. (2003). Commodity Price Modeling that matches Current Observables: A new approach, *Quantitative Finance*, 3(1):51–58.

Miranda, Mario J. and Glauber, Joseph W. (1993). Estimation of Dynamic Nonlinear Rational Expectations Models of Primary Commodity Markets with Private and Government Stockholding, *Review of Economic and Statistics*, 75(2):463–70.

Miranda, Mario J. and Rui, Xiongwen (1996). An Empirical Reassessment of the Commodity Storage Model, Mimeo, Department of Agricultural Economics, Ohio State University.

Ng, S. and F.J. Ruge-Murcia (2000). Explaining the Persistence of Commodity Prices. *Computational Economics*, 16:149–71.

Ng, V., and C. Pirrong (1994). Fundamentals and Volatility: Storage, Spreads, and the Dynamics of Metals Prices, *Journal of Business*, 67(2): 203–30.

Ng**,** S. and P. Perron **(**2001**),** Lag Length Selection and the Construction of Unit Root Tests with Good Size and Power”, *Econometrica***,** 69: 1519-1554.

Pindyck, R.S. (1994). Inventories and the Short-Run Dynamics of Commodity Prices, RAND Journal of Economics, 25:141-59.

Ribeiro, D. and Hodges, S.D. (2004). A Two-Factor Model for Commodity Prices and Futures Valuations, Working Paper, Financial Options Research Centre, Warwick Business School.

Routledge, B.R., D.J. Seppi and C.S. Spatt (2000). Equilibrium Forward Curves for Commodities., *Journal of Finance*, 55(3):1297–338.

Samuelson, P. A. (1965). Proof that Properly Anticipated Prices Fluctuate Randomly,.*Industrial Management Review* 6:41–49.

Schwartz, E. S. (1997). The Stochastic Behavior of Commodity Prices: Implications for valuation and hedging, *Journal of Finance*, 52:923–73.

Sorensen, C. (2002). Modeling Seasonality in Agricultural Commodity Futures, *Journal of Futures Markets*, 22:393–426.

Stationarity against the Alternative of a Unit Root (1992). *Journal of Econometrics*, 54:159–78.

Telser, Lester G. (1958). Futures Trading and the Storage of Cotton and Wheat, *Journal of Political Economy*, 66(3):233–55.

Williams, J. C. and B. D. Wright (1991). *Storage and Commodity Markets*. Cambridge, UK: Cambridge University Press.

Working, H. (1949). The Theory of the Price of Storage, *American Economic Review,* 39:1254–62.