Contango and Backwardation in the Crude Oil Market: 
Regime Switching Approach

June 2, 2012

Introduction

Modeling the stochastic nature of commodities prices is a crucial step for valuing financial and real contingent claims related to commodities prices. The notion of convenience yield, defined as the benefits accrued to the owner of the physical commodity due to the flexibility in handling shocks in the market, plays a central role in commodities price modeling as it derives the relationship between futures and spot prices in the commodities markets. Early models of commodities prices, such as Brennan and Schwartz (1985), include a constant convenience yield to a one-factor geometric Brownian motion (GBM) to model the movement of the spot price. Many recent models extend this model by adding more factors. Gibson and Schwartz (1990) modeled the convenience yield as a stochastic mean reverting process and found the model able to generate various kinds of futures term structures that are commonly seen in the market. Schwartz (1997) studied the implication of a three-factor model where the third factor is a stochastic interest rate. Casassus et al. (2005) studies an extension to these models and found the importance of convenience yield being a function of the spot price and the interest rate levels. Liu and Tang (2011) introduce a stochastic volatility in the convenience yield process.
Most of these models assume a mean reverting process to model the convenience yield. That is, the convenience yield process is specified to revert to a certain level, or an equilibrium level, at a certain speed. This specification is somehow restrictive. Theoretically, convenience yield is derived as a function of the inventory level and the supply and demand conditions. Accordingly, it is too restrictive to assume that there is only one state that the market should revert to all the time. Empirically, the estimated value of this equilibrium level is very unstable as shown below. This assumption may not have much impact on the short term pricing. However, given the fast speed of mean reversion in the convenience yield process resulted from the estimation of such models for some commodities, such as crude oil, as shown by Gibson and Schwartz (1990) and Schwartz and Smith (2000), this assumption may have a significant impact in the long run.

Markov switching models provide a good tool to relax this assumption. Markov switching models were introduced by Hamilton (1989) to capture nonlinearities in GNP growth rates arising from discrete jumps in the conditional mean. Regime switching models are well developed for bond pricing and the term structure of interest rates (Bansal and Zhou (2002) and Dai et al. (2007)) and electricity futures prices (Blochlinger (2008)). However, they are less explored in studying other commodities futures term structure. Much of the attention in this literature is to capture the time series properties of the observed commodities prices. For example, Fong and See (2003) modeled the conditional volatility of crude oil futures returns as a regime switching process. The model features transition probabilities that are functions of the basis, the spread between the spot and futures prices. Alizadeh et al. (2008) purposed a regime switching conditional volatility model and studied its implication on the optimal hedge ratio and the hedging performance. Chiarella et al. (2009) modeled the evolution of the gas forward curve using regime switching. Chen and
Forsyth (2010) proposed a one-factor regime-switching model for the risk adjusted natural gas spot price and studied the implications of the model on the valuation and optimal operation of natural gas storage facilities. They solved the partial differential equation governing the futures prices numerically and used a least squares approach to calibrate the model parameters. Chen (2010) proposed a regime switching model for crude oil prices in order to capture the historically observed periods of lower but more stable prices followed by periods of high and volatile prices. The study modeled the crude oil spot price as a mean reverting process that reverts to different levels and exhibits different volatilities within each regime.

In this study, regime switching framework is exploited to study the movement in crude oil futures term structures. In particular, the Brennan and Schwartz (1985) one-factor has been extended to accommodate for shifts in the convenience yield level and, in turn, in the futures term structure in a discrete time setting. Unlike to Chen and Forsyth (2010) and Chen (2010), the model of this study allows for pricing the regime switching risk as well as the market price risk. Moreover, a closed form solution for the futures prices is derived and an extension to the Kalman filter suggested by Kim (1994) is used to estimate the model parameters.

Compared to the performance of the Gibson and Schwartz (1990) two-factor model, the regime switching one-factor model of this study does a reasonable job. In particular, the model outperforms the Gibson and Schwartz (1990) model for fitting the prices of far maturities contracts. Moreover, the transitional probabilities have been found to play an important rule in shaping the futures term structure implied by the model.

The paper is organized as follows: a background of convenience yield modeling is given in section 1. In the following section, the regime switching model is specified. Section 3 is devoted to futures price formula derivation. In 4 section, an estimation
method based on the Kalman filter is propped in detail. Data description and the model estimation results for the crude oil market are given in the next two sections. The last section is devoted to concluding remarks.

1 Convenience Yield in Commodities Price Modeling

The convenience yield is defined as the stream of benefits received by holding an extra unit of the commodity in storage rather than buying the unit in the futures market. This stream of benefits comes from the fact that holding commodity in storage enables the holder to respond flexibly and efficiently to supply and demand shocks. This concept has been introduced to reduced form modeling of commodities prices. Expressing this yield as a fraction to the commodity price, i.e. convenience yield = $\delta \cdot S_t$, where $S_t$ is the commodity spot price, Brennan and Schwartz (1985) introduces constat $\delta$ to the Geometric Brownian motion to model the price stochastic movement, that is:

$$dS_t = \left(\mu - \delta\right)dt + \sigma_s dz_{s,t},$$

where $\mu$ is the total rate of return of holding one unit of the commodity$^1$ and $\sigma_s$ is the volatility of the commodity price. $dz_{s,t}$ is a Brownian motion increment to account for the stochastic movement in the commodity price. This simple model implies only one shape of the futures term structure depending on $\delta^2$. In reality, the futures term structure is seen in different shapes. Gibson and Schwartz (1990) shows that the assumption of constat convenience yield is very restrictive. Accordingly, driven by the numerical properties of the convenience yield implied by the futures prices, they

---

1 The return on holding one unit of the commodity, $\mu$, comes from two sources: the rate of change in the commodity price (the capital gain) and the convenience yield. Thus, $\mu - \delta$ corresponds to the rate of change in the commodity price (the capital gain)

2 Brennan and Schwartz (1985) shows that the futures price for delivery in $\tau$ periods is given by $F_t(\tau) = S_t e^{(r-\delta)\tau}$. Thus, the term structure has positive (negative) slope if $r > \delta$ ($r < \delta$)
introduced a mean reverting stochastic process for the convenience yield movement as follows:

\[
\frac{dS_t}{S_t} = (\mu - \delta_t)dt + \sigma_s dz_{s,t}, \quad (2a)
\]

\[
d\delta_t = \kappa(\theta - \delta_t)dt + \sigma_\delta dz_{\delta,t}, \quad (2b)
\]

where \(\delta_t\) reverts to a long-run (or an equilibrium) value of \(\theta\) at a speed of \(\kappa\) with volatility of \(\sigma_\delta\). In this setting, the near end of the futures term structure can take any shape depending on how far is \(\delta_t\) from its long-run level, \(\theta\), and on the speed of the reversion, \(\kappa\). However, the far end of the futures term structure converges to one shape depending on the value of \(\theta\) compared to the risk free rate. Many of recent models can be seen as an extension to Gibson and Schwartz (1990) model. For example, Casassus et al. (2005) allows the convenience yield to depend on the spot and the interest rate. Liu and Tang (2011) introduces a stochastic volatility in the convenience yield process to account for the heteroscedasticity observed in the implied convenience yield.

The assumption that \(\delta_t\) reverts to a certain level in the long-run is somehow restrictive. From the theoretical side, the convenience yield is seen as a function of the level of the commodity inventory in the economy\(^3\) which is in turn a function of the supply and demand conditions. Moreover, macroeconomic conditions which run through different cycles of booms and busts are likely to have impact on the commodities markets especially for crucial commodities such as crude oil. Given that, it is unlikely that there is only one equilibrium state the commodity market should revert to. From the empirical side, estimating the Gibson and Schwartz (1990) model, the model in equation 2, using crude oil futures in different periods of time produces very different values of \(\theta\). For example, estimating the model using weekly WTI crude

\(^3\)Refer to Chapter 2 for detailed explanation about the theory of storage and the notion of convenience yield.
Figure 1. The Implied Forward Curve of Gibson and Schwartz (1990) Model

The two curves show the crude oil futures term structure implied from the estimated parameters in Table 2. Initial value of the spot price is $80 and the initial value of the convenience yield is equal to $\theta$.

(a) Parameters are estimated using WTI futures from Jan. 1992 to Jan. 2000 ($\hat{\theta} = 0.0392$)

(b) Parameters are estimated using WTI futures from Jan. 2000 to Aug. 2011 ($\hat{\theta} = 0.1149$)

Oil futures price from 01/01/1992 to 01/01/2000 and from 01/01/2000 to 6/10/2011 yields the value of $\theta$ equal to 0.0392 and 0.1149 respectively\(^4\). These two values imply very different shapes of the futures term structure as shown in figure 1.

Markov switching models provide a good venue to take to relax this assumption. In the next section, a regime switching model based on Brennan and Schwartz (1985) one-factor model is specified.

2 Regime Switching Model Specification

Let $P_t$ be the spot price of the commodity at time $t$ and let $x_t$ be the logarithm of the spot price, i.e. $x_t = \log(P_t)$. Assuming that there are a number of regimes the commodity market could run through, the dynamic of $x_t$ in each regime under the

\(^4\)The full estimation results can be seen in Table 3.
objective measure is given by:

\[ \Delta x_t = x_{t+1} - x_t = (\mu_{s_t} - \delta_{s_t}) \Delta t + \sigma_{s_t} \sqrt{\Delta t} \epsilon_{t+1}, \quad \epsilon_{t+1} \sim N(0, 1). \]  

(3)

\( \mu_{s_t} \) is the total expected return from holding one unit of the commodity, \( \delta_{s_t} \) is the convenience yield accrued by holding one unit of the commodity in storage and \( \sigma_{s_t} \) is the volatility of the commodity price change. The values of all three parameters are functions of which regime the market is in, which is indicated by the subscript \( s_t \) where \( s_t \) is the process that determines which regime the market is in at time \( t \). This specification can be seen as an extension to the continuous time Brennan and Schwartz (1985) model but with regime dependent parameters. Thus, we call this model the Brennan and Schwartz regime switching (B&S-RS) model.

As in Hamilton (1994), \( s_t \) is modeled as an S-state discrete time Markov chain process which is assumed to be independent of \( x_t \). The evolution of \( s_t \) is governed by the transitional probability matrix which specifies the probability of switching from one regime to another. In this study we are interested in the case where \( S = 2 \), i.e. there are only two regimes in the market. For a two-state Markov chain, the transitional matrix under the objective measure \( \mathbb{P} \) is then given by:

\[
\Pi^{\mathbb{P}}_{t,t+1} = \begin{bmatrix}
\pi^{\mathbb{P}}_{(1,1)} & 1 - \pi^{\mathbb{P}}_{(1,1)} \\
1 - \pi^{\mathbb{P}}_{(2,2)} & \pi^{\mathbb{P}}_{(2,2)}
\end{bmatrix},
\]  

(4)

where \( \pi^{\mathbb{P}}_{(i,i)} \) is the \( \mathbb{P} \) measure probability to stay in regime \( i \) at \( t + 1 \) given the market is in regime \( i \) at \( t \) where \( i = 1, 2 \).

The B&S-RS model of this paper is going to be compared with the Gibson and Schwartz (1990) two-factor model (G&S). The discrete time version of G&S model
can be written as follows:

\[ \Delta x_t = (\mu - \delta_t) \Delta t + \sigma_s \epsilon_{1,t} \]  
\[ \Delta \delta_t = \kappa (\theta - \delta_t) \Delta t + \sigma_\delta \epsilon_{2,t}, \]  

and

\[ \begin{bmatrix} \epsilon_{1,t} \\ \epsilon_{2,t} \end{bmatrix} \sim \mathcal{N} \left( \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & \rho_{x\delta} \\ \rho_{x\delta} & 1 \end{bmatrix} \Delta t \right). \]  

\[ x_t \] is again the log of the spot price which has a total expected return of \( \mu \) with volatility of \( \sigma_x \). \( \delta_t \) is the convenience yield and is modeled as a mean reverting process. It reverts to a long-run level of \( \theta \) at a speed of \( \kappa \) with volatility of \( \sigma_\delta \). \( \rho_{x\delta} \) is the instantaneous correlation between the shocks of the two processes.

For later development, the above two models, B&S-RS and G&S, are written in the following form:

\[ X_{t+1} = \alpha^{(s_t)} + \beta^{(s_t)} X_t + \Sigma^{(s_t)} \epsilon_{t+1}, \]  

where for the B&S-RS model: \( X_{t+1} = x_{t+1} \), \( \alpha^{(s_t)} = (\mu_{s_t} - \delta_{s_t}) \Delta t \), \( \beta^{(s_t)} = 1 \), \( \Sigma^{(s_t)} = \sigma_{s_t} \sqrt{\Delta t} \) and \( \epsilon_{t+1} \sim \mathcal{N} (0, 1) \);

and for the G&S model:

\[ \begin{bmatrix} x_t \\ \delta_t \end{bmatrix}, \alpha^{(s_t)} = \begin{bmatrix} \mu \\ \kappa \cdot \theta \end{bmatrix} \Delta t, \beta^{(s_t)} = I_{2 \times 2} + \begin{bmatrix} 0 & -1 \\ 0 & -\kappa \end{bmatrix} \Delta t, \]  
\[ \Sigma^{(s_t)} = \begin{bmatrix} \sigma_x \sqrt{1 - \rho_{x\delta}^2} & \rho_{x\delta} \sigma_x \\ 0 & \sigma_\delta \end{bmatrix} \sqrt{\Delta t}, \text{ and } \epsilon_t \sim \mathcal{N} \left( \begin{bmatrix} 0 & 0 \end{bmatrix}^T, I_{2 \times 2} \right). \]

If the market does not allow arbitrage opportunities, then according to the fundamental theory of asset pricing (see (Björk, 2003, Theorem 3.8)), there exists a positive stochastic discount factor, denoted by \( M_{t,t+1} \), underlying the time-t valua-
tion of the payoff of any contingent claim paid at date $t + 1$. That is, if $G_{t+1}$ is the payoff of the contingent claim at $t + 1$, then:

$$\text{Price}_t(G_{t+1}) = E(M_{t,t+1}G_{t+1}).$$

Following Dai et al. (2007), to allow for pricing the risk of the regime shift, $M_{t,t+1}$ is parameterized as follows:

$$M_{t,t+1} \equiv M(X_t, s_t; X_{t+1}, s_{t+1}) = \exp (-r_{t,t+1} - \gamma_{t,t+1} - \frac{1}{2} \Lambda_t \Lambda_t^\top - \Lambda_t \epsilon_{t+1}),$$

(9)

where $\gamma_{t,t+1} \equiv \gamma(X_t, s_t; s_{t+1})$ is the market price of risk associated with regime shift from regime $s_t$ at time $t$ to regime $s_{t+1}$ at time $t + 1$ and it can be a function of the state variable $X_t$. $\Lambda_t \equiv \Lambda(X_t, s_t)$ is the market price of risk associated with the stochastic movement of $X_t$ and it is also regime and state dependent. $r_{t,t+1}$ is the risk free rate at time $t$ for one period which is assumed to be deterministic, i.e $r_{t,t+1} = r \cdot \Delta t$.

The existence of a stochastic discount factor under the absence of arbitrage implies an equivalent martingale measure, $Q$ measure, under which the price of any contingent claim would be the expectation of the discounted payoff. That is, there exists a measure $Q$ such that:

$$\text{Price}(G_{t+1}) = E_t^Q(e^{-r \cdot \Delta t}G_{t+1}),$$

(10)

where $E_t^Q[\cdot]$ denotes the conditional expectation under the $Q$ measure.

Given equation (10) and the specification of the stochastic discount factor in equation (9), the equivalent $Q$ measure is then defined by (see the derivation in Appendix A):

$$\frac{Q(dX_{t+1}, s_{t+1} = k|X_t, s_t = j)}{P(dX_{t+1}, s_{t+1} = k|X_t, s_t = j)} = e^{-\gamma_{t,t+1} - \frac{1}{2} \Lambda_t^\top \Lambda_t - \Lambda_t \epsilon_{t+1}},$$

(11)
where \( \gamma^{(j,k)}_{t,t+1} \equiv \gamma(X_t, s_t = j, s_{t+1} = k) \) and \( \Lambda^{(j)}_t \equiv \Lambda(X_t, s_t = j) \).

Assuming \( \gamma^{(j,k)}_{t,t+1} \) to be constant, i.e. \( \gamma^{(j,k)}_{t,t+1} = \gamma^{(j,k)} \), then the regime switching probabilities under \( Q \) are given by:

\[
\pi^Q_{j,k} = \mathbb{E}_t[1_{s_{t+1} = k}|s_t = j] = \pi^P_{j,k} \cdot e^{\gamma^{(j,k)}},
\]

where \( 1_{s_{t+1} = k} \) is an indicator function that equals to 1 if the subscript is true and zero otherwise.

Moreover, assuming a constant market price of risk within each regime, i.e. \( \Lambda^{(s_t)}_t = (\Sigma^{(s_t)})^{-1} \Lambda^{(s_t)} \), where \( \Lambda^{(s_t)} \) is vector of constant within each regime, then it is shown in Appendix B that the behavior of \( X_t \) in the \( Q \) measure is given by:

\[
X_{t+1} = \hat{\alpha}^{(s_t)} + \beta^{(s_t)} X_t + \Sigma^{(s_t)} \epsilon_{t+1},
\]

where: \( \hat{\alpha}^{(s_t)} = \alpha^{(s_t)} - \Lambda^{(s_t)} \).

For B&S-RS model, \( \Lambda^{(s_t)} = \lambda^{(i)} \Delta t, i = 1, 2 \), while for G&S model, since it is regime-independent, \( \Lambda^{(s_t)} = [\lambda_x \Delta t \ \lambda_\delta \Delta t]^\top \).

If the commodity is a traded asset, then absence of arbitrage implies that the total expected return of holding a unit of the commodity should be equal to the risk-free rate. This is due the fact that one can design a portfolio of the commodity and the derivatives and choose the weights to eliminate the risk. Approximating the return on the commodity by the difference in the logarithm\(^5\), i.e. \( \Delta x_t \), this dynamic hedging implies:

\[
E^Q[\Delta x_t|J_t, s_t = j] \approx (r - \delta^{(j)}) \Delta t,
\]

\(^5\)The approximating is reasonable for small time step. In this paper, weekly data is used, i.e. \( \Delta t = 0.0192 \). Thus the error is negligible.
for B&S-RS one factor model, and

\[ E^Q [\Delta x_t | J_t] \approx (r - \delta_t) \Delta t, \]  

(15)

for the G&S two factor model.

Thus, for the B&S-RS one factor model:

\[ \hat{\alpha}^{(s_t)} = (r - \delta^{(s_t)}) \Delta t \]

and, for the G&S two factor model:

\[ \hat{\alpha}^{(s_t)} = \begin{bmatrix} r \\ \kappa \cdot \theta - \lambda \delta \end{bmatrix} \Delta t \]

. 

3 Futures Pricing

Denote \( F_{t,n} \equiv F_n(X_t, s_t) \) to be the futures price at time \( t \) of a unit of the commodity delivered in \( n \) periods. A futures contract entered at time \( t \) has a payoff at time \( t + 1 \) of \( F_{t+1,n-1} - F_{t,n} \). Since, there is no payment at the inception at time \( t \), this payoff must have a price of zero, that is:

\[ 0 = e^{-r \Delta t} E^Q_t [F_{t+1,n-1} - F_{t,n}], \]  

(16)

which implies:

\[ F_{t,n} = E^Q_t [F_{t+1,n-1}], \]  

(17)

Appendix C shows that the futures price is an exponential affine function of the state variable within each regime. Specifically:
\[ F_n(X_t, s_t = j) = e^{A_n(j) + B_n X_t}, \]  

where for B&S-RS model:

\[ A_n(j) = \log \left( \sum_{k=1}^{S} \pi_{jk} \cdot e^{A_{n-1}(k)} \right) + B_{n-1} \hat{\alpha}(j) + \frac{1}{2} B_{n-1} \Sigma(j) \Sigma(j)^\top B_{n-1}^\top \]  

\[ B_n = B_{n-1} \beta, \]

with \( A_0(i) = 0 \) for \( i = 1, 2 \) and \( B_0 = 1 \).

For the G&S model, \( S = 1 \) which implies:

\[ A_n = A_{n-1} + B_{n-1} \hat{\alpha} + \frac{1}{2} B_{n-1} \Sigma \Sigma^\top B_{n-1}^\top \]

\[ B_n = B_{n-1} \beta, \]

with \( A_0 = 0 \) and \( B_0 = [1 \ 0] \).

### 4 Estimation Methodology

As the focus of this study is on the crude oil markets, the B&S-RS one factor model can be estimated using the time series of the crude oil spot price or, if not available, the first contract of futures prices as a proxy. However, the parameters estimated this way would not correspond to the \( Q \) measure which is the relevant measure for pricing contingent claims. Moreover, using such methods, the convenience yield, \( \delta_i \) where \( i = 1, 2 \), cannot be identified.

In bonds pricing literature, where regime switching models have been studied extensively, several estimation methods have been proposed. Bansal and Zhou (2002) used efficient method of moments (EMM). Dai et al. (2007) relied on maximum likelihood method (MLE) which involves inverting equation (18) to extract the states vector, \( X_t \), which has Gaussian conditional likelihood. However, this requires one to
chose a number of futures contracts or to design a number of portfolios of futures contracts that is the same as the number of factors to be extracted. The contracts choice or the portfolio weights are chosen arbitrarily. Duffee and Stanton (2004) compared the performance of the three methods in estimating affine term structure models: MLE, EMM and methods based on Kalman filter. They found that MLE is a good method for simple term structure models and the performance of EMM (a commonly used method for estimating complicated models) is poor even in the simple term structure models. According to Duffee and Stanton (2004), Kalman filtering procedure is found to be a tractable and reasonably accurate estimation technique that they recommend in settings where maximum likelihood is impractical.

Thus, in this paper we used an extension to the Kalman filter for estimating the parameters of the B&S-RS model proposed in Kim (1994). Blochlinger (2008) used this procedure to estimate electricity price models in regime switching framework.

The Kalman filter is a recursive procedure for computing the optimal estimator of the state vector at time $t$, based on the information available up to time $t$, and it enables the estimate of the state vector to be continuously updated as new information becomes available. When the disturbances and the initial state vector are normally distributed, the Kalman filter enables the likelihood function to be calculated, which allows for the estimation of any unknown parameters of the model and provides the basis for statistical testing and model specification. For a detailed discussion of state space models and the Kalman filter see Chapter 3 in Harvey (1989).

The first step in using Kalman filtrating procedure is to cast the model in the state space form. To do this, one needs to specify the transition equation that governs the dynamic of the state variables and the measurement equation that relates the observable variables to the state variables.
The transition equation is represented by equation (8), which is:

\[ X_{t+1} = \alpha^{(s_t)} + \beta^{(s_t)} X_t + \Sigma^{(s_t)} \epsilon_{t+1} \]

where for B&S-RS model,

\[
\alpha^{(s_t)} = \begin{cases} 
(r + \lambda_1 - \delta_1) \Delta t & \text{if } s_t = 1 \\
(r + \lambda_2 - \delta_2) \Delta t & \text{if } s_t = 2 
\end{cases}, \\
\Sigma^{(s_t)} = \begin{cases} 
\sigma_1 \Delta t & \text{if } s_t = 1 \\
\sigma_2 \Delta t & \text{if } s_t = 2 
\end{cases},
\]

and \( \beta^{(s_t)} = 1 \) in both regimes, and for the G&S model, the matrices \( \alpha^{(s_t)} \), \( \beta^{(s_t)} \) and \( \Sigma^{(s_t)} \) are same as defined in section 2. At each time, a vector of (log) future prices of the commodity for different maturities is observed. Assuming that these prices are observed with measurement error (these errors may be caused by bid-ask spreads, the non-simultaneity of the observations, etc. see Schwartz (1997)), the measurement equation will then be:

\[
\begin{bmatrix}
  f_t(n_1) \\
f_t(n_2) \\
  \vdots
\end{bmatrix} =
\begin{bmatrix}
  A_{n_1}^{(s_t)} \\
  A_{n_2}^{(s_t)} \\
  \vdots
\end{bmatrix} +
\begin{bmatrix}
  B_{n_1} \\
  B_{n_2} \\
  \vdots
\end{bmatrix} X_t +
\begin{bmatrix}
  e_{1,t} \\
  e_{2,t} \\
  \vdots
\end{bmatrix}
\]

\[ Y_t = A^{(s_t)} + BX_t + e_t \quad e_t \sim N(0, Q_{s_t}), \quad (24) \]

where \( e_t \) represent the measurement error in the futures prices. It is assumed that the measurement errors are not correlated and have regime independent volatilities. That is, \( Q_{s_t} = Q \) where the off diagonal elements of \( Q \) and are zeros and the diagonal elements, denoted by \( \nu_{i}^2 \), are to be estimated.

The Kim (1994) filter extends the Kalman filter to accommodate state space
models with regime switching. To ease the explanation of the algorithm, let $J_t \equiv (Y_t, Y_{t-1}, Y_{t-2}, \ldots, Y_1)$ and define the following:

\[ X_{t|t-1}^{(i,j)} = E[X_t|J_{t-1}, s_t = j, s_{t-1} = i] \]
\[ P_{t|t-1}^{(i,j)} = E[(X_t - X_{t|t-1})(X_t - X_{t|t-1})^\top|J_{t-1}, s_t = j, s_{t-1} = i] \]
\[ X_{t|t}^{(j)} = E[X_t|J_t, s_t = j] \]
\[ P_{t|t}^{(j)} = E[(X_t - X_{t|t})(X_t - X_{t|t})^\top|J_{t-1}, s_t = j] \]

That is, $X_{t|t-1}^{(i,j)}$ is the forecast of $X_t$ based on information up to time $t-1$ conditional on $s_t$ being in the regime $j$ and on $s_{t-1}$ being on regime $i$ and $P_{t|t-1}^{(i,j)}$ is the associated mean square error. On the other hand, $X_{t|t}^{(j)}$ is the inference about $X_t$ based on information up to time $t$, given that $s_t$ is in regime $j$ and $P_{t|t}^{(j)}$ is the associated mean square error.

Given these definitions, the algorithm goal is to start with $X_{t|t-1}^{i}$ and $P_{t|t-1}^{i}$ from the previous step to produce $X_{t|t}^{i}$ and $P_{t|t}^{i}$ of the current step using the above model and the current observation of time $t$. Specifically, it goes by:

\[ X_{t|t-1}^{(i,j)} = \alpha_i X_{t|t-1}^{i} \]  \quad (25)
\[ P_{t|t-1}^{(i,j)} = \alpha_i P_{t|t-1}^{i} + \Sigma_i \Sigma_i^\top \]  \quad (26)
\[ \eta_{t|t-1}^{(i,j)} = Y_t - (A_j + B_j X_{t|t-1}^{(i,j)}) \]  \quad (27)
\[ H_{t}^{(i,j)} = B_j P_{t|t-1}^{(i,j)} B_j^\top + Q^j \]  \quad (28)
\[ K_{t}^{(i,j)} = P_{t|t-1}^{(i,j)} B_j^\top [H_{t}^{(i,j)}]^{-1} \]  \quad (29)
\[ X_{t|t}^{(i,j)} = X_{t|t-1}^{(i,j)} + K_{t}^{(i,j)} \eta_{t|t-1}^{(i,j)} \]  \quad (30)
\[ P_{t|t}^{(i,j)} = (I - K_{t}^{(i,j)}) P_{t|t-1}^{(i,j)}. \]  \quad (31)

These step constitutes the Kalman filtering procedure and given the normality
assumption of the pricing errors, the likelihood of observing $Y_t$ conditional on $J_{t-1}$ and on $s_t = j$ and $s_{t-1} = i$ can be evaluated as follows:

$$f(Y_t|s_{t-1} = i, s_t = j, J_t) = (2\pi)^{-N/2}|H_t^{(i,j)}|^{-1/2}\exp(-\frac{1}{2}H_t^{(i,j)\top}H_t^{(i,j)}),$$

(32)

where $N$ is the size of $X_t$. For Gibson and Schwartz (1990) model, where the model is regime independent (i.e. $S = 1$), the above likelihood reduced to $f(Y_t|J_{t-1})$.

However, if the number of the regimes is $S > 1$ (in our case $S = 2$), then the results of the above filtration procedure is $S^2$ forecasts, $X_{t|t}^{(i,j)}$, and $S^2$ associated forecast errors, $P_{t|t}^{(i,j)}$. Thus, each iteration would require $S$-fold of cases to consider. Kim (1994) suggested the following approximation, in each iteration, to collapse the $S^2$ forecasts and their associated forecast errors to only $S$ cases:

$$X_{t|t}^{(j)} = \sum_{i=1}^{S} \frac{Pr[s_{t-1} = i, s_t = j|J_t] \cdot X_{t|t}^{(i,j)}}{Pr[s_t = j|J_t]}$$

(33)

$$P_{t|t}^{(j)} = \sum_{i=1}^{S} \frac{Pr[s_{t-1} = i, s_t = j|J_t] \cdot \left(P_{t|t}^{(i,j)} + (X_{t|t}^{(j)} - X_{t|t}^{(i,j)})(X_{t|t}^{(j)} - X_{t|t}^{(i,j)})\top\right)}{Pr[s_t = j|J_t]}.$$  

(34)

Kim (1994) gives the details behind this approximation of this procedure. The outputs $X_{t|t}^{(j)}$ and $P_{t|t}^{(j)}$ are then used as inputs to the Kalman filtration procedure in the next step.

To achieve this recursive procedure, one needs to calculate the probabilities terms appearing in equations (33) and (34). Kim (1994) suggested to use the Hamilton (1994) procedure to obtain these probabilities recursively. This procedure is explained in detail in Appendix D.

As shown in Appendix D, as a by product of the Hamilton (1994) filtration procedure, the conditional likelihood of each iteration, $f(Y_t; \psi|J_{t-1})$ is obtained, where $\psi$ is the set of parameters to be estimated.
Having the likelihood of each observation, the parameters of the two models can then be estimated by maximizing the likelihood of the sample, that is:

\[
\hat{\psi} = \arg \max_{\psi} \sum_{t=1}^{T} \log(f(Y_t; \psi|J_{t-1})).
\]

5 Data Description

To estimate the parameters of the two models, weekly data of West Texas Intermediate (WTI) crude oil futures are used. WTI crude oil futures contracts for more than four years maturities are traded in New York Mercantile Exchange (NYMEX). WTI futures contracts are very liquid and are among the most traded commodity futures worldwide. Data from January 1992 to the August of 2011 has been obtained from Datastream and the Energy Information Administration (US Department of Energy). For the risk free rate, the average of the 3 months U.S. treasury bill is used.

To construct continuous series of futures prices, following the literature, futures prices are sorted each week according to the contract horizon with "first month" contract being the contract with the earliest delivery date, the "second month" contract being the contract with the next earliest delivery date, etc. Each contract will switch to the next one just before it expires.\(^6\)

The performance of our specification of the B&S-RS one-factor model is compared with the G&S two-factor model. The estimation and performance analysis are done for the whole sample period and for two subperiods, namely: from January 1992 to January 2000 and from January 2000 to August 2011.

Table 1 reports descriptive statistics for the weekly returns of the spot, 6th, 12th, 17th and 20th months contracts: unconditionally and conditional on the slope of

---

\(^6\)For WTI, trading in the current delivery month ceases on the third business day prior to the twenty-fifth calendar day of the month preceding the delivery month. More details can be seen in http://www.cmegroup.com.
Right scale is for WTI oil spot price, left scale is for its log return and shaded areas are for the periods of contango markets.
Table 1. Descriptive Statistics

Descriptive statistics for the crude oil (log) returns for the whole sample and for the two sub-samples. Contango and backwardation is defined by the sign of the difference between $F_6$ and $F_1$. Positive sign indicates contango market and negative sign indicates backwardation market.

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Mean</th>
<th>Std. Dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>From January 1992 to the August of 2011</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unconditional</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F1</td>
<td>0.00199</td>
<td>0.05307</td>
<td>0.00044</td>
<td>0.05999</td>
<td>0.00318</td>
<td>0.04866</td>
</tr>
<tr>
<td>F6</td>
<td>0.00188</td>
<td>0.03866</td>
<td>-0.00007</td>
<td>0.04413</td>
<td>0.00378</td>
<td>0.03421</td>
</tr>
<tr>
<td>F12</td>
<td>0.00184</td>
<td>0.03256</td>
<td>-0.00022</td>
<td>0.03794</td>
<td>0.00393</td>
<td>0.02752</td>
</tr>
<tr>
<td>F17</td>
<td>0.00182</td>
<td>0.02991</td>
<td>-0.00021</td>
<td>0.03513</td>
<td>0.00394</td>
<td>0.02471</td>
</tr>
<tr>
<td>F20</td>
<td>0.00181</td>
<td>0.02868</td>
<td>-0.00018</td>
<td>0.03387</td>
<td>0.00392</td>
<td>0.02336</td>
</tr>
<tr>
<td>Contango (49.4%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F1</td>
<td></td>
<td></td>
<td>0.00044</td>
<td>0.05999</td>
<td>0.00318</td>
<td>0.04866</td>
</tr>
<tr>
<td>F6</td>
<td></td>
<td></td>
<td>-0.00007</td>
<td>0.04413</td>
<td>0.00378</td>
<td>0.03421</td>
</tr>
<tr>
<td>F12</td>
<td></td>
<td></td>
<td>-0.00022</td>
<td>0.03794</td>
<td>0.00393</td>
<td>0.02752</td>
</tr>
<tr>
<td>F17</td>
<td></td>
<td></td>
<td>-0.00021</td>
<td>0.03513</td>
<td>0.00394</td>
<td>0.02471</td>
</tr>
<tr>
<td>F20</td>
<td></td>
<td></td>
<td>-0.00018</td>
<td>0.03387</td>
<td>0.00392</td>
<td>0.02336</td>
</tr>
<tr>
<td>Backwardation (50.6%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F1</td>
<td></td>
<td></td>
<td>0.00031</td>
<td>0.05333</td>
<td>0.00363</td>
<td>0.04456</td>
</tr>
<tr>
<td>F6</td>
<td></td>
<td></td>
<td>-0.00145</td>
<td>0.03190</td>
<td>0.00343</td>
<td>0.02850</td>
</tr>
<tr>
<td>F12</td>
<td></td>
<td></td>
<td>-0.00211</td>
<td>0.02298</td>
<td>0.00302</td>
<td>0.02268</td>
</tr>
<tr>
<td>F17</td>
<td></td>
<td></td>
<td>-0.00240</td>
<td>0.01989</td>
<td>0.00272</td>
<td>0.02014</td>
</tr>
<tr>
<td>F20</td>
<td></td>
<td></td>
<td>-0.00239</td>
<td>0.01863</td>
<td>0.00251</td>
<td>0.01885</td>
</tr>
<tr>
<td><strong>From January 1992 To January 2000</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unconditional</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F1</td>
<td>0.00220</td>
<td>0.04686</td>
<td>-0.00031</td>
<td>0.05333</td>
<td>0.00363</td>
<td>0.04456</td>
</tr>
<tr>
<td>F6</td>
<td>0.00131</td>
<td>0.02930</td>
<td>-0.00145</td>
<td>0.03190</td>
<td>0.00343</td>
<td>0.02850</td>
</tr>
<tr>
<td>F12</td>
<td>0.00076</td>
<td>0.02244</td>
<td>-0.00211</td>
<td>0.02298</td>
<td>0.00302</td>
<td>0.02268</td>
</tr>
<tr>
<td>F17</td>
<td>0.00044</td>
<td>0.01978</td>
<td>-0.00240</td>
<td>0.01989</td>
<td>0.00272</td>
<td>0.02014</td>
</tr>
<tr>
<td>F20</td>
<td>0.00030</td>
<td>0.01855</td>
<td>-0.00239</td>
<td>0.01863</td>
<td>0.00251</td>
<td>0.01885</td>
</tr>
<tr>
<td>Contango (41.3%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F1</td>
<td></td>
<td></td>
<td>-0.00031</td>
<td>0.05333</td>
<td>0.00363</td>
<td>0.04456</td>
</tr>
<tr>
<td>F6</td>
<td></td>
<td></td>
<td>-0.00145</td>
<td>0.03190</td>
<td>0.00343</td>
<td>0.02850</td>
</tr>
<tr>
<td>F12</td>
<td></td>
<td></td>
<td>-0.00211</td>
<td>0.02298</td>
<td>0.00302</td>
<td>0.02268</td>
</tr>
<tr>
<td>F17</td>
<td></td>
<td></td>
<td>-0.00240</td>
<td>0.01989</td>
<td>0.00272</td>
<td>0.02014</td>
</tr>
<tr>
<td>F20</td>
<td></td>
<td></td>
<td>-0.00239</td>
<td>0.01863</td>
<td>0.00251</td>
<td>0.01885</td>
</tr>
<tr>
<td>Backwardation (58.7%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F1</td>
<td></td>
<td></td>
<td>0.00031</td>
<td>0.05333</td>
<td>0.00363</td>
<td>0.04456</td>
</tr>
<tr>
<td>F6</td>
<td></td>
<td></td>
<td>-0.00145</td>
<td>0.03190</td>
<td>0.00343</td>
<td>0.02850</td>
</tr>
<tr>
<td>F12</td>
<td></td>
<td></td>
<td>-0.00211</td>
<td>0.02298</td>
<td>0.00302</td>
<td>0.02268</td>
</tr>
<tr>
<td>F17</td>
<td></td>
<td></td>
<td>-0.00240</td>
<td>0.01989</td>
<td>0.00272</td>
<td>0.02014</td>
</tr>
<tr>
<td>F20</td>
<td></td>
<td></td>
<td>-0.00239</td>
<td>0.01863</td>
<td>0.00251</td>
<td>0.01885</td>
</tr>
<tr>
<td><strong>From January 2000 to August 2011</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unconditional</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F1</td>
<td>0.00184</td>
<td>0.05555</td>
<td>0.00068</td>
<td>0.06212</td>
<td>0.00281</td>
<td>0.05070</td>
</tr>
<tr>
<td>F6</td>
<td>0.00211</td>
<td>0.04205</td>
<td>0.00039</td>
<td>0.04751</td>
<td>0.00390</td>
<td>0.03682</td>
</tr>
<tr>
<td>F12</td>
<td>0.00230</td>
<td>0.03605</td>
<td>0.00041</td>
<td>0.04172</td>
<td>0.00438</td>
<td>0.02970</td>
</tr>
<tr>
<td>F17</td>
<td>0.00242</td>
<td>0.03332</td>
<td>0.00051</td>
<td>0.03885</td>
<td>0.00458</td>
<td>0.02675</td>
</tr>
<tr>
<td>F20</td>
<td>0.00246</td>
<td>0.03204</td>
<td>0.00055</td>
<td>0.03755</td>
<td>0.00465</td>
<td>0.02534</td>
</tr>
</tbody>
</table>
Figure 3. The Implied Term Structure of Futures

For both models, the spot price is set equal to 80$. For G&S model, the initial value of convenience yield is set to the delirium level, that is $\delta_0 = \theta$

(a) The Whole Sample
(b) The First Sub-sample
(c) The Second Sub-Sample

the futures term structure being positive or negative. The table shows that crude oil market visits backwardation regime and contango regime half of the time in the whole sample period and in the two sub-samples. Moreover, periods of backwardation generate higher returns. The market has higher volatility when being in contango than the case when it is being in backwardation. It is also clear that volatility declines with maturity; an observation known in futures prices literature as Samuelson’s effect.
Table 2. Kim Filter Estimation Results of the B&S-RS Model

<table>
<thead>
<tr>
<th>From</th>
<th>From</th>
<th>From</th>
</tr>
</thead>
<tbody>
<tr>
<td>01/01/1992</td>
<td>01/01/1992</td>
<td>01/01/2000</td>
</tr>
<tr>
<td>To</td>
<td>To</td>
<td>To</td>
</tr>
<tr>
<td>6/10/2011</td>
<td>01/01/2000</td>
<td>6/10/2011</td>
</tr>
<tr>
<td>(r = 0.0315)</td>
<td>(r = .0455)</td>
<td>(r = 0.0205)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>F1, F3, F7 &amp; F12</th>
<th>F1, F3, F7 &amp; F12</th>
<th>F1, F3, F7 &amp; F12</th>
</tr>
</thead>
<tbody>
<tr>
<td>value</td>
<td>value</td>
<td>value</td>
<td></td>
</tr>
<tr>
<td>(\delta_1)</td>
<td>-0.0916</td>
<td>-0.1321</td>
<td>-0.1155</td>
</tr>
<tr>
<td>(\sigma_1)</td>
<td>0.2783</td>
<td>0.2629</td>
<td>0.3137</td>
</tr>
<tr>
<td>(\lambda_1)</td>
<td>-0.1290</td>
<td>-0.3115</td>
<td>-0.1210</td>
</tr>
<tr>
<td>(\delta_2)</td>
<td>0.2425</td>
<td>0.1777</td>
<td>0.1983</td>
</tr>
<tr>
<td>(\sigma_2)</td>
<td>0.2223</td>
<td>0.1659</td>
<td>0.2514</td>
</tr>
<tr>
<td>(\lambda_2)</td>
<td>0.3723</td>
<td>0.2023</td>
<td>0.3406</td>
</tr>
<tr>
<td>(\pi_{Q1})</td>
<td>0.9850</td>
<td>0.9846</td>
<td>0.9812</td>
</tr>
<tr>
<td>(\pi_{Q2})</td>
<td>0.9905</td>
<td>0.9815</td>
<td>0.9937</td>
</tr>
<tr>
<td>(\pi_{P1})</td>
<td>0.9889</td>
<td>0.9921</td>
<td>0.9879</td>
</tr>
<tr>
<td>(\pi_{P2})</td>
<td>0.9920</td>
<td>0.9922</td>
<td>0.9965</td>
</tr>
<tr>
<td>(v_1)</td>
<td>0.0512</td>
<td>0.0558</td>
<td>0.0501</td>
</tr>
<tr>
<td>(v_2)</td>
<td>0.0317</td>
<td>0.0342</td>
<td>0.0279</td>
</tr>
<tr>
<td>(v_3)</td>
<td>0.0001</td>
<td>0.0002</td>
<td>0.0001</td>
</tr>
<tr>
<td>(v_4)</td>
<td>0.0220</td>
<td>0.0268</td>
<td>0.0219</td>
</tr>
<tr>
<td>LL</td>
<td>6914</td>
<td>2605</td>
<td>4084</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>F1, F4, F9 &amp; F13</th>
<th>F1, F4, F9 &amp; F13</th>
<th>F1, F4, F9 &amp; F13</th>
</tr>
</thead>
<tbody>
<tr>
<td>value</td>
<td>value</td>
<td>value</td>
<td></td>
</tr>
<tr>
<td>(\delta_1)</td>
<td>-0.0842</td>
<td>-0.1194</td>
<td>-0.1089</td>
</tr>
<tr>
<td>(\sigma_1)</td>
<td>0.2685</td>
<td>0.2556</td>
<td>0.3047</td>
</tr>
<tr>
<td>(\lambda_1)</td>
<td>-0.1205</td>
<td>-0.2933</td>
<td>-0.1122</td>
</tr>
<tr>
<td>(\delta_2)</td>
<td>0.2310</td>
<td>0.1707</td>
<td>0.1900</td>
</tr>
<tr>
<td>(\sigma_2)</td>
<td>0.2093</td>
<td>0.1574</td>
<td>0.2393</td>
</tr>
<tr>
<td>(\pi_{Q1})</td>
<td>0.9864</td>
<td>0.9863</td>
<td>0.9863</td>
</tr>
<tr>
<td>(\pi_{Q2})</td>
<td>0.9927</td>
<td>0.9926</td>
<td>0.9926</td>
</tr>
<tr>
<td>(\pi_{P1})</td>
<td>0.9889</td>
<td>0.9921</td>
<td>0.9878</td>
</tr>
<tr>
<td>(v_1)</td>
<td>0.0541</td>
<td>0.0468</td>
<td>0.0419</td>
</tr>
<tr>
<td>(v_2)</td>
<td>0.0267</td>
<td>0.0268</td>
<td>0.0219</td>
</tr>
<tr>
<td>(v_3)</td>
<td>0.0001</td>
<td>0.0002</td>
<td>0.0001</td>
</tr>
<tr>
<td>(v_4)</td>
<td>0.0220</td>
<td>0.0268</td>
<td>0.0319</td>
</tr>
<tr>
<td>LL</td>
<td>6924</td>
<td>2607</td>
<td>4089</td>
</tr>
</tbody>
</table>
6 Estimation Results

Table 2 shows the results of the Kim filter estimation and Figure 3 shows the implied term structure of the futures prices for the whole sample and the two sub-samples. For comparison purposes, the futures term structures implied by the Gibson and Schwartz (1990) two-factor model for the three periods are included.

Regarding the estimated parameters of B&S-RS model, all the parameters are significant except those for the market price of risk. Regime one is characterized by negative convenience ($\delta_1$) and higher volatility ($\sigma_1$). On the other hand, regime two is characterized by positive convenience ($\delta_2$) and lower volatility ($\sigma_2$). It is clear from the figure that the first regime corresponds to a positive slope of the futures term structure while the second regime corresponds to a negative slope of the futures term structure. The volatility estimates of both regimes are high reflecting the higher variability of crude oil markets as it is the case in energy markets and commodity markets in general (see for example Deaton and Laroque (1996)). However, this result is at odds with the theory of storage prediction which asserts that higher convenience yield is associated with low level of inventory and high volatility.

The classification of observation into the two regimes coincides with the observed slope of the futures term structure as shown in figure 4. There are two methods in literature to classify the observation into its regime. The first one is to use the smoothed regime probabilities, i.e. $Pr[s_t = j|J_T]$. If $Pr[s_t = j|J_T] > .5$, then $t$ is classified into regime $j$. This method was suggested by Hamilton (1994). The other method is to compare the fitted prices of each regime with the actual prices at each time and classify the observation to the regime that has the least error. That is, if the fitted futures price at time $t$ for regime $i$ is closer to the actual price, then, $s_t$ is set to equal $i$. This method has been used by Bansal and Zhou (2002).

From figure 4, comparing 4(a) with 4(b) and 4(c), it is clear that the first regime
Figure 4. Observation Classification into Regimes

Second Sub-period

(a) The sign of $F_6 - F_1$: shaded area is for positive sign

(b) Implied Classification Using Minimum Error Method: Shaded area for regime 1

(c) Implied Classification Using Smoothed Probability Method: Shaded area for regime 1
corresponds to the market having positive slope while the second one corresponds
to the market having a negative slop. In other words, the first regime shows the
market when being in contango and the second regime shows the market when being
in backwardation.

Transitional probabilities, \( \pi_1^P \) and \( \pi_2^P \), in both regimes reflect the persistence of
each regime. Note that the reported values are the probabilities to switch between
the two regimes in one week. The values are very high for both regimes because the
a period of one week is too short to allow for switching. To make a clearer picture,
the corresponding annual transitional probabilities are calculated as follows:

\[
\begin{bmatrix}
0.9905 & 1 - 0.9905 \\
1 - 0.9889 & 0.9889
\end{bmatrix}^{(52)} = \begin{bmatrix}
0.6951 & 0.3049 \\
0.3563 & 0.6437
\end{bmatrix}.
\]

That is, the probability to stay in the same regime after one year is 0.6951 for regime
1 and 0.6437 for regime 2. The result shows that both regimes are highly persistence.
Moreover, the market stays slightly longer in the first regime where it has lower
volatility and positive convenience yield. The recent persistence in regime one at the
end of the financial crisis and afterward, as shown in Figure 4, contributes to the fact
that \( \pi_1^P \) is slightly higher than \( \pi_2^P \) although the market visits regime two more than
regime one in the estimation period, as shown in Table 1.

The transitional probabilities in the pricing measure, measure \( Q \), is given by \( \pi_1^Q \)
and \( \pi_2^Q \). These estimates reflects how the market reacts to the risk of the regime
shift. Again, the annual transitional probabilities are calculated as follows:

\[
\begin{bmatrix}
0.9850 & 1 - 0.9850 \\
1 - 0.9920 & 0.9920
\end{bmatrix}^{(52)} = \begin{bmatrix}
0.5423 & 0.4577 \\
0.2441 & 0.7559
\end{bmatrix}.
\]

These number shows that the risk of switching from regime one to regime two in one
year is higher in \( Q \) than in the actual measure \( P \) (0.4577 > 0.3049). On the other
hand, the risk of switching from regime two to regime one in one year is lower in
than in the actual measure $\mathbb{P}$ ($0.2441 < 0.3563$). If the market is currently in regime one, where futures prices are higher than the spot price, risk-averse market participants set the switching risk (the probability that futures prices will drop lower than the spot price) to be higher than actual risk. On the other hand, if the market is currently in regime two, where futures prices are lower than the spot price, they set the switching risk (the probability that futures prices will jump up above the spot price) to be lower than the actual risk. In both cases, they achieve that by setting futures prices lower than what is expected to be if they are risk neutral in term of regime shift. Figure 5 explains this fact by depicting the implied futures term structure in each regime using the transitional probabilities of both measures. It shows that risk averse market participants, to account for the risk of regime shift, set
the futures prices lower than what spot price is expected to be at expiry. Moreover, the reduction increases as the maturity of the futures contract increases.

Table 2 also shows that the fist regime has a negative market price of risk, \( \lambda_1 = -0.129 \), while the second regime has a high and a positive market price of risk, \( \lambda_2 = 0.37 \). That is, in the first regime, market participants trade futures contracts at higher prices than expected at maturity as a reward of bearing the price risk. On the other hand, in the second regime, they trade futures contracts at lower prices than expected at maturity. This is also confirmed by the observation from Table 1 where the returns are negative in contango (regime one according to the estimates of our model) and positive in backwardation (regime two according to the estimates of our model). However, the standard errors of the market price of risk (MPR) parameters, \( \lambda_1 \) and \( \lambda_2 \), are relatively high compared with the other parameters. The reason for that is, as explained in Schwartz and Smith (2000), the MPR parameters cannot be directly identified from futures prices.

Table 2 shows that the the same pattern of results appears also in the two sub-samples reflecting the stability of the parameters.

6.1 Model Comparison

The Kalman filter estimation results for the G&S model two-factor model are shown in Table 3. The G&S model implies that there is only one equilibrium long run slope of the futures term structure that is dictated by the parameter \( \theta \). That is, the possibility that the futures term structure shift its long run slope in the future is ignored. If the speed to return to this long-run slope, which is dictated by the parameter \( \kappa \) is fast (which is the case for estimated \( \kappa \) here and is also shown by Gibson and Schwartz (1990)), this would have significant impact on long term forecasting, investment and risk management decisions where the futures term structure is used as a risk-adjusted
Table 3. Kalman Filter Estimation Results of the G&S Two-factor Model

<table>
<thead>
<tr>
<th>From</th>
<th>From</th>
<th>From</th>
</tr>
</thead>
<tbody>
<tr>
<td>01/01/1992</td>
<td>01/01/1992</td>
<td>01/01/2000</td>
</tr>
<tr>
<td>To</td>
<td>To</td>
<td>To</td>
</tr>
<tr>
<td>6/10/2011</td>
<td>01/01/2000</td>
<td>6/10/2011</td>
</tr>
</tbody>
</table>

F1, F3, F7 & F12 F1, F3, F7 & F12 F1, F3, F7 & F12

<table>
<thead>
<tr>
<th>value</th>
<th>SE</th>
<th>value</th>
<th>SE</th>
<th>value</th>
<th>SE</th>
</tr>
</thead>
<tbody>
<tr>
<td>θ</td>
<td>0.0827</td>
<td>0.0023</td>
<td>0.0392</td>
<td>0.0032</td>
<td>0.1149</td>
</tr>
<tr>
<td>κ</td>
<td>1.3994</td>
<td>0.0219</td>
<td>1.6254</td>
<td>0.0348</td>
<td>1.3637</td>
</tr>
<tr>
<td>σₓ</td>
<td>0.3718</td>
<td>0.0052</td>
<td>0.3186</td>
<td>0.0088</td>
<td>0.3873</td>
</tr>
<tr>
<td>σδ</td>
<td>0.3892</td>
<td>0.0080</td>
<td>0.4327</td>
<td>0.0150</td>
<td>0.4063</td>
</tr>
<tr>
<td>ρₓδ</td>
<td>0.8721</td>
<td>0.0085</td>
<td>0.9516</td>
<td>0.0080</td>
<td>0.7567</td>
</tr>
<tr>
<td>λₓ</td>
<td>0.1360</td>
<td>0.0764</td>
<td>0.0789</td>
<td>0.0902</td>
<td>0.1002</td>
</tr>
<tr>
<td>λδ</td>
<td>0.0166</td>
<td>0.0839</td>
<td>0.1642</td>
<td>0.1502</td>
<td>-0.1029</td>
</tr>
<tr>
<td>v</td>
<td>0.0107</td>
<td>0.0001</td>
<td>0.0108</td>
<td>0.0001</td>
<td>0.0096</td>
</tr>
<tr>
<td>LL</td>
<td>9802.6</td>
<td>3727.2</td>
<td>56454</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

forecast for the future spot prices. Figure 3 shows how the equilibrium slope of the G&S model is different across the two subperiods which reflect the instability of θ, the parameter that dictates the equilibrium slope of the futures term structure.

In term of comparing the performance of B&S-RS model with the G&S model in fitting the futures term structure, figure 6 shows that although the G&S model outperforms B&S-RS model for short term maturities, B&S-RS model outperforms the G&S model for long-term maturities. Figures 7 and 8 show the same pattern for the two sub-periods. This result is an implication of the assumption of the G&S model that the convenience yield reverts to only one equilibrium level which implies that the futures term structure should revert to one slope all the time.

6.2 The Impact of the Transitional Probabilities

Within the framework of our regime switching model, the transitional probabilities π₁₁ and π₂₂ have an important role in shaping the term structure of the futures prices.
Figure 6. Model Performance Compared to Gibson and Schwartz (1990) Model

(a) Pricing Error for log(F2)

(b) Pricing Error for log(F15)
Figure 7. Model Performance Compared to Gibson and Schwartz (1990) Model

First Sub-period

(a) Pricing Error for log(F2)

(b) Pricing Error for log(F15)
Figure 8. Model Performance Compared to Gibson and Schwartz (1990) Model

Second Sub-period

(a) Pricing Error for log(F2)

(b) Pricing Error for log(F15)
Figure 9. The Impact of The Transitional Probabilities

For easing the display of the graphs, $\pi_{11}^Q = p$ and $\pi_{22}^Q = q$

(a) Regime One: $p$ changes and $q$ is at estimated value (b) Regime One: $q$ changes and $p$ is at estimated value

(c) Regime Two: $p$ changes and $q$ is at estimated value (d) Regime Two: $q$ changes and $p$ is at estimated value

Figure 9 shows how the term structure changes with different values of the transitional probabilities and for each regime.

For easy presentation, denote $p = \pi_{11}^Q$ and $q = \pi_{22}^Q$. Consider the market to be currently in regime one where the slope of the forward curve is positive. Figure 9(a) shows that as $p$ decrease relative to $q$, the far end of the curve bends down. This is because the probability to be in regime 2, which is characterized by having negative slope, is now higher. Figure 9(b) shows that as $q$ decreases, the far end of the curve
bends up because the probability to be in regime 1, which is characterized by having positive slope, is now higher. However, the effect of changing $q$ is much smaller than the effect of changing $p$ because the market is now in regime one. The same pattern is also seen when the market is currently in regime 2 where the forward curve has a negative slope. Figure 9(d) shows that as $q$ decrease relative to $p$, the far end of the curve bends up because the probability to be in regime 1, which has positive slope, is now higher. Figure 9(c) shows that as $p$ increases, the far end of the curve bends up because the probability to be in regime 1. However, the effect of changing $p$ is much smaller than the effect of changing $q$ because the market is currently in regime two.

The estimates of the transitional probabilities reflects the persistence of the regimes during the estimation period, which is not necessary true for future periods. Thus, a way to forecast the transitional probability might be needed to improve the regime switching model prediction and hence the decision making based on it. An appealing feature of B&S-RS model of this study is that the regimes that the market runs through correspond to the slope of futures term structure. A large amount of research has been conducted to explain the factors behind the changes in the shape of the futures term structure, especially what factors make the market being in contango or backwardation. Examples of such factors are: inventory level and volatility of the market (see Williams and Wright (1991) and Deaton and Laroque (1996)); hedging pressure, the net hedging position of short and long sides of the market, (see Roon et al. (2000), Cantekin Dinceler and Titman (2003) and Gorton et al. (2007)). These factors can be exploited to estimate and forecast the probability of switching from a regime to another, and then use these estimates as inputs to the pricing formula instead of relying on the historical estimates. For example, Brooks et al. (2011) run a logistic regression to estimate the effect of the inventory on the sign of the slope of the futures term structure.
7 Concluding Remarks

In this paper, we exploit regime switching models to study the movement in the crude oil futures term structures. In particular, we extend Brennan and Schwartz (1985) one-factor model to accommodate for shifts in in the futures term structure between contango and backwardation and the reverse in discrete time setting. We allow for the regime shift risk as well as the market price risk. Moreover, we derive a closed from solution for the futures prices and used an extension to the Kalman filter suggested by Kim (1994) to estimate the model parameters. Compared to the performance of the Gibson and Schwartz (1990) two-factor, the regime switching one-factor model of this study did a reasonable job. In particular, the model outperforms Gibson and Schwartz (1990) model for fitting the prices of longer maturities contracts. Moreover, we found that the transitional probabilities played an important rule in shaping the futures term structure implied by the model.

As a future extension to the model, one might add more state factors other than the spot price movement. In this case, one needs to be careful in choosing the added factors as an important assumption in deriving the futures formula is that the state variables are independent from the Markov chain process governing the regime switching. Another direction is to test the improvement of the model when the transitional probabilities are estimated using the variables that the literature finds have explanation power in determining the shape of the futures term structure such as inventory level and the net hedging position.

References


A The Definition of the $\mathbb{Q}$ Measure

\[
\text{Price}_t(\mathcal{G}_{t+1}|X_{t}, s_t = j) = \mathbb{E}^\mathbb{P}(\mathcal{M}_{t,t+1}\mathcal{G}_{t+1}|X_{t}, s_t = j)
\]

\[
= e^{-r_t} \sum_{k=1}^{S} \frac{M_{t+1}^{j,k}}{e^{-r_t}} \mathbb{P}(dX_{t+1}, s_{t+1} = k|X_{t}, s_t = j)
\]

\[
= e^{-r_t} \sum_{k=1}^{S} \int e^{-\gamma_{t+1}} e^{-\frac{1}{2} \Lambda_t^\top \Lambda_t - \Lambda_t^\top \epsilon_{t+1}} \mathcal{G}_{t+1} \mathbb{P}(dX_{t+1}, s_{t+1} = k|X_{t}, s_t = j)
\]

\[
= e^{-r_t} \sum_{k=1}^{S} \int \mathcal{G}_{t+1} \mathbb{Q}(dX_{t+1}, s_{t+1} = k|X_{t}, s_t = j)
\]

\[
= e^{-r_t} \mathbb{E}^\mathbb{Q}[\mathcal{G}_{t+1}]
\]

B $X_t$ in the $\mathbb{Q}$ Measure

Denote $\mathbb{E}^\mathbb{P}_{t,j}[\cdot] \equiv \mathbb{E}^\mathbb{P}[\cdot|X_{t}, s_t = j]$. Before deriving the dynamic of $X_t$ in measure $\mathbb{Q}$, observe that the no arbitrage price of a zero coupon bond that pays $\$1$ at $t+1$ is:

\[
e^{-r_t} = \mathbb{E}^\mathbb{P}_{t,j}[1 \cdot \mathcal{M}_{t,t+1}] = \mathbb{E}^\mathbb{P}_{t,j}
\left[e^{-r_t - \gamma_{t+1}} - \frac{1}{2} \Lambda_t^\top \Lambda_t - \Lambda_t^\top \epsilon_{t+1}\right],
\]

which implies:

\[
1 = \mathbb{E}^\mathbb{P}_{t,j}[e^{-\gamma_{t+1}}] = \sum_{k=1}^{S} \pi_{j,k} e^{-\gamma(k,k)}
\]
The conditional moment generation function (MGF) of $X_{t+1}$ given $s_t = j$ is

$$
E^Q \left[ e^{u^T X_{t+1}} | X_t, s_t = j \right] = e^{\gamma^T s_t} \mathbb{E}^P \left[ \mathcal{M}_{t,t+1} e^{u^T X_{t+1}} | X_t, s_t = j \right] \\
= E^P_{t,j} \left[ e^{-\gamma_{t,t+1} - \frac{1}{2} \Lambda^T_t \Lambda_t - \Lambda^T_t \epsilon_{t+1}} , e^{u^T X_{t+1}} \right] \\
= E^P_{t,j} \left[ e^{-\gamma_{t,t+1} - \frac{1}{2} \Lambda^T_t \Lambda_t - \Lambda^T_t \epsilon_{t+1}} , e^{u^T (\alpha(s_t)+\beta(s_t) X_t) + u^T \Sigma(s_t) \epsilon_{t+1}} \right] \\
= E^P_{t,j} \left[ e^{-\gamma_{t,t+1}} \cdot E^P_{t,j} \left[ e^{-\frac{1}{2} \Lambda^T_t \Lambda_t - \Lambda^T_t \epsilon_{t+1}} , e^{u^T (\alpha(s_t)+\beta(s_t) X_t) + u^T \Sigma(s_t) \epsilon_{t+1}} \right] \right] \\
= E^P_{t,j} \left[ e^{-\frac{1}{2} \Lambda^T_t \Lambda_t + u^T (\alpha(s_t)+\beta(s_t) X_t) + \frac{1}{2} \left(u^T \Sigma(s_t) - \Lambda^T_t \right) \left(u^T \Sigma(s_t) - \Lambda^T_t \right)^T} \right] \\
= e^{u^T (\alpha(s_t) + \beta(s_t) X_t) - \frac{1}{2} u^T \Sigma(s_t) \Sigma(s_t)^T u} \\
= e^{u^T (\alpha(s_t) + \beta(s_t) X_t) + \frac{1}{2} u^T \Sigma(s_t) \Sigma(s_t)^T u} \\
= e^{u^T (\hat{\alpha} + \beta(s_t) X_t) + \frac{1}{2} u^T \Sigma(s_t) \Sigma(s_t)^T u},
$$

which implies that the behavior of $X_{t+1}$ is as follows:

$$
X_{t+1} = \hat{\alpha} + \beta(s_t) X_t + \Sigma(s_t) \epsilon_{t+1}
$$
C Futures Price Formula Derivation

\[ F_{t,n} = E^Q[F_{t+1,n-1}|X_t, s_t = j] \]
\[ F_n(X_t, s_t) = E^Q[F_{n-1}(X_{t+1}, s_{t+1})|X_t, s_t = j] \]
\[ e^{A_n^{(j)} + B_n^{(j)}X_t} = E^Q\left[e^{A_{n-1}^{(st+1)} + B_{n-1}^{(st+1)}X_{t+1}}\big|X_t, s_t = j\right] \]
\[ = \sum_{k=1}^{S} \pi_{jk} E^Q\left[e^{A_{n-1}^{(k)} + B_{n-1}^{(k)}X_{t+1}}\right] \]
\[ = \sum_{k=1}^{S} \pi_{jk} E^Q\left[e^{A_{n-1}^{(k)} + B_{n-1}^{(k)}(\alpha^{(j)} + \beta^{(j)}X_t + \Sigma^{(j)}\epsilon_{t+1})}\right] \]
\[ = \sum_{k=1}^{S} \pi_{jk} e^{A_{n-1}^{(k)} + B_{n-1}^{(k)}(\alpha^{(j)} + \beta^{(j)}X_t)} E^Q\left[e^{B_{n-1}^{(k)}\Sigma^{(j)}\epsilon_{t+1}}\right] \]
\[ = \sum_{k=1}^{S} \pi_{jk} e^{A_{n-1}^{(k)} + B_{n-1}^{(k)}(\alpha^{(j)} + \beta^{(j)}X_t)} + \frac{1}{2}B_{n-1}^{(k)}\Sigma^{(j)}\Sigma^{(j)\top} B_{n-1}^{(k)\top} \]

\[ A_n^{(j)} + B_n^{(j)}X_t = \log \left(\sum_{k=1}^{S} \pi_{jk} e^{A_{n-1}^{(k)} + B_{n-1}^{(k)}(\alpha^{(j)} + \beta^{(j)}X_t)} + \frac{1}{2}B_{n-1}^{(k)}\Sigma^{(j)}\Sigma^{(j)\top} B_{n-1}^{(k)\top}\right) \]

since \( \beta^{(1)} = \beta^{(1)} \) for both B&S-RS and G&S models and setting \( B(n)^{(1)} = B(n)^{(2)} = B(n) \), one gets:

\[ A_n^{(j)} + B_nX_t = B_{n-1}\beta X_t + \log \left(\sum_{k=1}^{S} \pi_{jk} e^{A_{n-1}^{(k)} + B_{n-1}^{(k)}(\alpha^{(j)} + \beta^{(j)}X_t)} + \frac{1}{2}B_{n-1}^{(k)}\Sigma^{(j)}\Sigma^{(j)\top} B_{n-1}^{(k)\top}\right) \]

Wich implies:

\[ A_n^{(j)} = \log \left(\sum_{k=1}^{S} \pi_{jk} e^{A_{n-1}^{(k)}}\right) + B_{n-1}\alpha^{(j)} + \frac{1}{2}B_{n-1}\Sigma^{(j)}\Sigma^{(j)\top} B_{n-1}^{\top} \]
\[ B_n = B_{n-1}\beta. \]
### D Hamilton (1994) Filtration Procedure

Step 1 Calculate $Pr[s_{t-1} = i, s_t = j|J_{t-1}]$ as following:

$$Pr[s_{t-1} = i, s_t = j|J_{t-1}] = Pr[s_t = j|s_{t-1}] \cdot \sum_{h=1}^{S} Pr[s_{t-2} = h, s_{t-1} = i|J_{t-1}]$$

$$= \pi^{(i,j)}_{t-1,t} \cdot \sum_{h=1}^{S} Pr[s_{t-2} = h, s_{t-1} = i|J_{t-1}]$$

where $\pi^{(i,j)}_{t-1,t}$ is the transitional probability and can be taken from the transitional matrix.

Step 2 Calculate the joint density function of $y_t$ and $(s_{t-1}, s_t)$ as following:

$$f(y_t, s_{t-1} = i, s_t = j|J_t) = f(y_t|s_{t-1} = i, s_t = j, J_t) \cdot Pr[s_{t-1} = i, s_t = j|J_{t-1}]$$

where

$$f(y_t|s_{t-1} = i, s_t = j, J_t) = (2\pi)^{-N/2}|H_t^{(i,j)}|^{-1/2}\exp(-\frac{1}{2} \eta^{(i,j)}_{t-1} H_t^{(i,j)} \eta^{(i,j)}_{t-1})$$

Step 3 Calculate:

$$Pr[s_{t-1} = i, s_t = j|J_t] = \frac{f(y_t, s_{t-1} = i, s_t = j|J_{t-1})}{f(y_t|J_{t-1})} \quad (35)$$

where

$$f(y_t|J_{t-1}) = \sum_{j=1}^{S} \sum_{i=1}^{S} f(y_t, s_{t-1} = i, s_t = j|J_{t-1})$$