

A Two-Factor Model for Commodity Prices and Futures Valuation*

Diana R. Ribeiro[†] Stewart D. Hodges[‡]

August 25, 2004

Abstract

This paper develops a reduced form two-factor model for commodity spot prices and futures valuation. This model extends the Gibson and Schwartz (1990)-Schwartz (1997) two-factor model by adding two new features. First the Ornstein-Uhlenbeck process for the convenience yield is replaced by a Cox-Ingersoll-Ross (CIR) process. This ensures that our model is arbitrage-free. Second, spot price volatility is proportional to the square root of the convenience yield level. We empirically test both models using weekly crude oil futures data from 17th of March 1999 to the 24th of December 2003. In both cases, we estimate the models parameters using the Kalman filter.

*We would like to thank Elizabeth Whalley and Mathias Muck for their helpful comments. Earlier versions of this paper were presented at the Finance Seminar Series at the Warwick Business School, the European Financial Management Association 2004 meeting and the Bachelier Finance Society Third World Congress. Comments from the seminar and conference participants were much appreciated. The first author would like to thank the financial support from the Fundacao para a Ciencia e Tecnologia, Portugal.

[†]Doctoral Researcher at the Warwick Business School, University of Warwick, Coventry, CV4 7AL, United Kingdom, phone +44 (0)24 76524465, fax: +44 (0)24 7652 7167 and email: D.Ribeiro@wbs.ac.uk

[‡]Director of the Financial Options Research Center (FORC), Warwick Business School, University of Warwick, Coventry CV4 7AL, United Kingdom, phone: +44 (0)24 76523606, fax: +44 (0)24 7652 4176 and email Stewart.Hodges@wbs.ac.uk

1 Introduction

In this paper, we introduce a new reduced form model for commodity spot prices and futures valuation which builds on and extends the reduced form models in the literature.

The earliest reduced form model for commodity prices appears to be due to Brennan and Schwartz (1985). In this model the spot commodity price follows a geometric Brownian motion and the convenience yield is treated as a dividend yield. This specification is inappropriate since it does not take into account the mean reversion property of spot commodity prices and neglects the inventory-dependence property of the convenience yield. Gibson and Schwartz (1990) introduce a two-factor, constant volatility model where the spot price and the convenience yield follow a joint stochastic process with constant correlation. Specifically, the spot price follows a geometric Brownian motion and the convenience yield follows a mean reverting stochastic process of the Ornstein-Uhlenbeck (O-U) type. The convenience yield is brought into the spot price process as a dividend yield. Schwartz (1997) further explores this model empirically by adopting a more sophisticated calibration method and tests it with several commodities. Schwartz (1997) three-factor model, Miltersen and Schwartz (1998) and Hilliard and Reis (1998) add a third stochastic factor to the model to account for stochastic interest rates. Nevertheless, the inclusion of stochastic interest rates in the commodity price models does not have a significant impact in the pricing of commodity options and futures in practice. Accordingly, interest rate can be assumed deterministic.

The reduced form class of models dominates the current literature and practice on energy derivatives. These models are particularly attractive from practitioner's perspective since they provide closed form solutions to evaluate futures and some other derivatives contracts. This in turn allows for a relatively easy calibration and computational implementation of these models.

Although these multi-factor models generate a rich set of dynamics for the commodity term structure and represent prevailing tools for derivatives pricing, they also present a number of problems. First, these models do not guarantee that the convenience yield is

always positive, possibly allowing for arbitrage opportunities. More specifically, arbitrage-free arguments require that the discounted futures prices net of carrying costs cannot be greater than the contemporaneous spot prices. By not ruling out negative values for the convenience yield, this arbitrage argument may be violated. Secondly, these models present other mis-specification problems due to the fact that both the spot price and the convenience yield have constant volatility and correlation. Accordingly, they do not allow the variance of the spot and futures, and the correlation between them, to depend on the level of the price or convenience yield, as suggested by the theory of storage. The commodity spot price volatilities are strongly heteroskedastic (see Duffie and Gray (1995)) and closely reflect the supply, demand and inventory conditions in the market. In particular, price volatilities are increasing with the degree of backwardation (see Ng and Pirrong (1994)) and Litzenberger and Rabinowitz (1995)). These mis-specifications may generate severe option mispricings, as pointed out by Pirrong (1998), Clewlow and Strickland (2000) and Routledge et al. (2000).

The reduced form model presented in this paper extends the two-factor model of Gibson and Schwartz (1990) - Schwartz (1997). More specifically, we develop a two-factor model where spot prices and instantaneous convenience yield follow a joint stochastic process with constant correlation. Our model introduces two significant additions to Gibson and Schwartz model: it rules out arbitrage possibilities and it considers time varying spot and convenience yield volatilities. Namely, the spot price follows a Geometric Brownian Motion (GBM) where the convenience yield is treated as an exogenous dividend yield and the volatility is proportional to the square root of the instantaneous convenience yield level. The instantaneous convenience yield follows a Cox-Ingersoll-Ross (CIR) which precludes negative values and makes the volatility proportional to the square root of the instantaneous convenience yield level. This ensures that our model does not allow cash-and-carry arbitrage possibilities.

We obtain a closed-form solution for the prices of futures prices of the exponential affine form¹. We solve the partial differential equation (PDE) for futures prices by sup-

¹Since we assume that the interest rate is constant, the futures and forward prices are the same.

posing that the solution has a general exponential affine form. By replacing this general affine form into the initial PDE we obtain a system of two ordinary differential equations (ODEs) with boundary conditions consistent with the futures price expiry condition. We find that each of these ODEs has a unique closed form solution. These in turn provide the solution to the PDE satisfied by the futures prices². This affine relationship is tractable and offers empirical advantages. In particular, the linear relationship between the logarithm of the futures price and the underlying state variables allows the use of the Kalman filter in the estimation of the parameters of the model. Spot prices data are not easily obtained in most of the commodity markets and therefore futures prices with closest maturity are used as a proxy for the commodity price level. Additionally, the instantaneous convenience yield is not observable and must be derived from the relationship between the spot and futures prices with closest maturity. On the other hand, futures prices are widely observed and traded in diverse markets. The non-observability of the state variables remains one of the main difficulties in modelling commodity spot prices and contracts. Due to the non-observability of the state variables, the linearity of the logarithm of the futures prices in the state variables and the Markovian property of these, the Kalman filter seems to be the most appropriate technique to estimate the model parameters. This method is also applied by Schwartz (1997). The basic principle of Kalman filter is the use of temporal series of observable variables to reconstitute the value of the non-observable variables. Accordingly, by observing futures prices, we can estimate the parameter values for the spot price and convenience yield.

We apply the Kalman filter method to estimate the parameters of our model using light crude oil futures data for the period from 17th of March 1999 to 24th of December 2003. Additionally, we also apply the Kalman filter to the Gibson and Schwartz (1990) - Schwartz (1997) two factor model using the same data set for estimation purposes and to compare the results. We also compare the futures volatility structure implied by the data and by both our model and Schwartz model.

Very recently and through independent work, Nielsen and Schwartz (2004) present a

²A similar solution method has been applied in the interest rate models literature such as Hull and White (1990), Brown and Schaefer (1994) and Duffie and Kan (1993, 1996).

two-factor model which also assumes stochastic volatility of the Cox-Ingersoll-Ross type. Additionally, they also compare their model to the Gibson and Schwartz (1990) - Schwartz (1997) two-factor model and conclude that the latter prices futures contracts almost as well. However, both the present model and the Nielsen and Schwartz (2004) model are special cases of the more general theory of affine term structures for forward and futures described by Bjork and Landen (2002). There are three important differences between our work and that of Nielsen and Schwartz (2004). First, Nielsen and Schwartz consider a general volatility function of the CIR type while in the present model the convenience yield follows a familiar square root process exactly as used in Cox et al. (1985). Although the Nielsen and Schwartz model appears to be more general, it does not preclude negative values for the convenience yield. In fact, our model closely resembles the Heston (1993) stochastic volatility model for standard asset prices and therefore can be interpreted as the storable commodity version of this model³. The second difference lies in the nature of the data used to test the models empirically, which follows Gibson and Schwartz (1990) - Schwartz (1997). Nielsen and Schwartz (2004) use copper futures data to calibrate their model while we use light crude oil futures data. Finally, we establish an empirical comparison between the futures volatility structure implied by the data and by both our model and Gibson and Schwartz (1990) - Schwartz (1997) two factor model. This empirical comparison is not considered by Nielsen and Schwartz (2004).

The remaining paper is organized as follows. Section 2 develops the two-factor model and derives the corresponding partial differential equation for futures valuation. Section 3 describes the empirical work, including the state-space formulation of the model, the data used and the empirical results. Section 4 concludes.

2 Valuation Model

In this section we present the commodity price model and derive the corresponding formulas for pricing futures contracts. This model has two stochastic factors. The first

³Because we are not dealing with options pricing in this paper but with futures pricing, we are able to get an analytical solution without the use of characteristic functions.

factor is the spot price, which follows as a GBM with a time-varying volatility, which is proportional to the square root of the instantaneous convenience yield level. The second factor is the convenience yield, which follows a CIR stochastic process as described by Cox et al. (1985). This process precludes negative convenience yields and implies that the absolute variance of the convenience yield increases when the convenience yield itself increases. We assume that both stochastic processes have constant correlation. The direct proportionality of the spot price and the convenience yield volatilities to the square root of the instantaneous convenience yield reflect the effect of supply, demand and inventory market conditions on the spot price and the convenience yield volatilities. As Duffie and Gray (1995) point out, the commodity spot price volatilities are strongly heteroskedastic and closely reflect the supply, demand and inventory conditions in the market. In particular, price volatilities are increasing with the degree of backwardation (see Ng and Pirrong (1994) and Litzenberger and Rabinowitz (1995)). That is, the stronger the backwardation is the higher the convenience yield is. High convenience yield levels signal low inventory and the possibility of a stockout. Therefore, the spot price volatility is positively related with the value of the convenience yield. Similarly, when the market is in contango, the spot prices volatility should be low. The market conditions and the commodity spot prices also affect similarly the volatility of the convenience yield itself.

The model we present is very tractable since it allows a closed form solution to futures prices. Specifically, we obtain a linear relation between the logarithm of futures prices and the underlying factors. This property is crucial to the empirical work that follows.

We assume that the spot price and the instantaneous convenience yield follow the joint stochastic process:

$$dp = (\mu(\cdot) - \delta)pdt + \sigma_1\sqrt{\delta}pdB_1 \quad (1)$$

$$d\delta = (\alpha(m - \delta))dt + \sigma_2\sqrt{\delta}dB_2 \quad (2)$$

where:

- $\mu(\cdot)$ is the total expected return on the spot commodity price;
- σ_1 represents the constant of proportionality between the total spot price volatility and the square root of the instantaneous convenience yield;
- σ_2 represents the constant of proportionality between the total instantaneous convenience yield volatility and the square root of the instantaneous convenience yield;
- α is the instantaneous convenience yield's speed of mean reversion;
- m is the convenience yield long-run mean, that is, the level to which δ reverts as t goes to infinity;
- B_1 and B_2 are standard Wiener processes and are correlated with $dB_1dB_2 = \rho dt$, ρ being constant.

The probability density of the convenience yield at time t conditional on its value at current time t is a non-central chi-square (see Cox et al. (1985)). The conditional moments of δ at time t under the objective measure are given by:

$$E[\delta_t|\delta_{t-dt}] = m(1 - e^{-\alpha dt}) + \delta_{t-dt}e^{-\alpha dt}, \quad (3)$$

$$Var[\delta_t|\delta_{t-dt}] = m \left(\frac{\sigma_2^2}{2\alpha} \right) (1 - e^{-\alpha dt})^2 + \delta_{t-dt} \left(\frac{\sigma_2^2}{\alpha} \right) (e^{-\alpha dt} - e^{-2\alpha dt}). \quad (4)$$

By defining $x = \ln p$ and applying Ito's Lemma, the process for the log price is given by:

$$dx = \left(\mu(\cdot) - \left(1 + \frac{1}{2}\sigma_1^2 \right) \delta \right) dt + \sigma_1\sqrt{\delta}dB_1. \quad (5)$$

Under the risk neutral measure, the stochastic processes that drive the the state variables becomes:

$$dp = (r + c - \delta)pdt + \sigma_1\sqrt{\delta}pdB_1^*, \quad (6)$$

$$d\delta = (\alpha(m - \delta) - \lambda_\delta)dt + \sigma_2\sqrt{\delta}dB_2^*, \quad (7)$$

where:

- r is the risk-free (constant) interest rate;
- c is the (constant) marginal cost of storage, which is a proportion of the spot price;
- λ is the (constant) market price of risk for the convenience yield;
- $\sigma_1, \sigma_2, \alpha$ and m are as before;
- B_1^* and B_2^* are standard Wiener processes under the risk-neutral measure and are correlated with $dB_1^*dB_2^* = \rho dt$, ρ as before.

The expected growth of the commodity price in a risk-neutral world is $\mu - \lambda_p\sigma_1$, where λ_p is the market price of risk of commodity price. Since the commodity behaves like a traded security that provides a dividend rate equal to δ , the expected growth rate of the commodity price under the risk-neutral measure is also given by $r + c - \delta$. Therefore $r + c - \delta = \mu(\cdot) - \lambda_p\sigma_1\sqrt{\delta}$. Accordingly, the drift of the spot price process, $\mu(\cdot)$ in the real measure is replaced by $(r + c - \delta)$ under the risk neutral measure. Equation (6) is an extension of a standard process for the commodity process allowing for a stochastic convenience yield and a time varying volatility. This volatility is proportional to the square root of the time-varying stochastic convenience yield. On the other hand, since the convenience yield is non-traded, the convenience yield risk cannot be hedged and will have a market price of risk, λ_δ associated with it. Therefore, the drift of the convenience yield under the risk neutral measure becomes $[\alpha(m - \delta) - \lambda_\delta]$. Although λ_δ is a function of $\alpha(m - \delta)$ and $\sigma_2\sqrt{\delta}$, we assume that the convenience yield market price of risk is constant. This assumption is standard and is also followed by Schwartz (1997). The process for the log price then becomes:

$$dx = \left(r + c - \left(1 + \frac{1}{2}\sigma_1^2 \right) \delta \right) dt + \sigma_1 \sqrt{\delta} dB_1^*. \quad (8)$$

By assuming that the instantaneous convenience yield follows a CIR process we ensure that our model is arbitrage-free because it precludes negative values. This assumption ensures that the discounted futures prices net of carrying costs cannot be greater than the discounted contemporaneous spot prices, which is derived by arbitrage-free arguments. Considering that $\tau = T - t$ represents time to maturity the arbitrage-free condition can be written as:

$$F(\tau) \leq p_t \exp\{(r + c)(\tau)\}, \quad (9)$$

where

- $F(\tau)$ is the forward price at time t , for delivery of a commodity at time $T > t$;
- p_t is the spot price of the commodity at time t ;
- c is the (constant) proportion of the spot price which defines the marginal cost of storage;
- r is risk-free (constant) interest rate.

If the instantaneous convenience yield is always non-negative the arbitrage-free condition in equation (9) is satisfied⁴.

The futures prices must satisfy the partial differential equation (PDE):

$$\frac{1}{2}\sigma_1^2 \delta p^2 F_{pp} + \frac{1}{2}\sigma_2^2 \delta F_{\delta\delta} + \rho\sigma_1\sigma_2 \delta p F_{p\delta} + ((r + c) - \delta)p F_p + \quad (10)$$

$$(\alpha(m - \delta) - \lambda) F_\delta - F_\tau = 0 \quad (11)$$

⁴Note that Gibson and Schwartz (1990) - Schwartz (1997) does ensure non-negative convenience yield given that the stochastic convenience yield in his two-factor model follows an Ornstein-Uhlenbeck. This may generate arbitrage possibilities in his model.

subject to the boundary condition $F(p, \delta, 0) = p$. This PDE suggests an exponential affine form solution:

$$F(p, \delta, \tau) = pe^{A(\tau) - B(\tau)\delta}, \quad (12)$$

with

$$A(0) = 0; \quad B(0) = 0. \quad (13)$$

Equivalently, the logarithm of the futures prices is given by:

$$\ln F(p, \delta, \tau) = \ln p + A(\tau) - B(\tau)\delta. \quad (14)$$

The futures prices as given by equation (12) satisfy the PDE (10) and the boundary condition when

$$\frac{1}{2}\sigma_2^2 B^2 + (\alpha - \rho\sigma_1\sigma_2)B - 1 + B_\tau = 0, \quad (15)$$

and

$$(r + c) + (\lambda - \alpha m)B - A_\tau = 0, \quad (16)$$

with initial conditions

$$A(0) = 0; \quad B(0) = 0. \quad (17)$$

It follows that if (15) and (16) are solved subject to the boundary conditions in (17), equation (12) provides the price of a futures contract maturing at time T . In appendix we provide the derivation of the solution to the ODEs above, which is given by:

$$B(\tau) = \frac{2(1 - e^{-k_1\tau})}{k_1 + k_2 + (k_1 - k_2)e^{-k_1\tau}} \quad (18)$$

and

$$A(\tau) = (r + c)\tau + (\lambda - \alpha m) \int_t^T B(q) dq, \quad (19)$$

where:

$$\int_t^T B(q) dq = \frac{2}{k_1(k_1 + k_2)} \ln \left[\frac{(k_1 + k_2)e^{k_1\tau} + k_1 - k_2}{2k_1} \right] + \quad (20)$$

$$\frac{2}{k_1(k_1 - k_2)} \ln \left[\frac{k_1 + k_2 + (k_1 - k_2)e^{-k_1\tau}}{2k_1} \right], \quad (21)$$

with:

$$k_1 = \sqrt{k_2^2 + 2\sigma_2^2} \quad (22)$$

$$k_2 = (\alpha - \rho\sigma_1\sigma_2) \quad (23)$$

The solution to equation (10) with initial boundary condition $F(p, \delta, 0) = p$ is given by (12) with $A(\tau)$ and $B(\tau)$ given by (18) and (19).

3 Empirical Estimation of the Joint Stochastic Process

In this section we estimate and empirically test both our model and the Gibson and Schwartz (1990) - Schwartz (1997) two-factor model. Data for most of the spot commodity prices are extremely difficult to obtain price for most of the commodities. On the other hand, we are able to observe daily several futures prices at different maturities. This non-observability and the linear relationship between futures prices and the state variables in the model clearly suggest that the Kalman filter is the most appropriate technique to estimate the model's parameters. The principle of Kalman filter is to use a time series of observable variables and to infer the value of the non-observable variables. This technique

is suitable whenever there is a linear dependency of the observable variables upon the state variables and when the latter are Markovian processes. Kalman filter is a technique which has become increasingly popular in Finance and has been applied to both Gaussian and CIR type interest rate models and in commodity futures valuation in (e.g. Schwartz (1997), Schwartz and Smith (2000) and Manoliu and Tompaidis (2002)). Affine models are particularly suited for estimating using Kalman filter because of their linear structure. In the context of interest rate models Gaussian examples can be found in Babbs and Nowman (1999) and Lund (1997), who estimate a two-factor generalized Vasicek model. In the CIR case, there are examples due to Ball and Tourous (1996), Duan and Simonato (1999) and Lund (1997). The Gibson and Schwartz (1990) - Schwartz (1997) two-factor model belongs to the Gaussian class while our model fits in the CIR class.

The state form is applied to a multivariate time series of observable variables, which in our case are a futures prices time series at several different maturities. These observed variables are related to the state vector which consist of the state variables, which in our model are the spot price and the instantaneous convenience yield via the *measurement equation*. The measurement equation is then given by equation (14) by adding serially and cross sectionally uncorrelated disturbances with mean zero and variance to take into account for the irregularities of the observations. In the Kalman filter, the non-observable state variables are generated by first-order Markov processes which correspond to the discrete time form of equations (1) and (2). The latter are arranged in a vector, which forms the *transition equation*. See Harvey (1989) for a detailed description of this method. We calibrate Gibson and Schwartz (1990) - Schwartz (1997) two-factor⁵ model using exactly the same methodology as described in the latter article. To calibrate our model we follow the same steps but we need to take into account an important difference between the two empirical models. The state-space form of the Gibson and Schwartz

⁵According to Gibson and Schwartz (1990) - Schwartz (1997) two-factor model, the commodity spot price and the convenience yield follow a joint stochastic process with constant correlation given by:

$$\begin{aligned} dp &= (\mu(\cdot) - \delta)pdt + \sigma_1 pdB_1, \\ d\delta &= (\alpha(m - \delta))dt + \sigma_2 dB_2, \end{aligned}$$

where $dB_1dB_2 = \rho dt$.

model is Gaussian while the state-space form of our model is non-Gaussian, given that we do not have constant volatility.

For a Gaussian state-space model, the Kalman filter provides an optimal solution to prediction, updating and evaluating the likelihood function. The Kalman filter recursion is a set of equations which allows an estimator to be updated once a new observation becomes available. The Kalman filter first forms an optimal predictor of the unobserved state variable vector given its previously estimated value. This prediction is obtained using the distribution of unobserved state variables, conditional on the previous estimated values. These estimates for the unobserved state variables are then updated using the information provided by the observed variables. Prediction errors, obtained as a by-product of the Kalman filter, can then be used to evaluate the likelihood function.

When the state-space model is non-Gaussian, the Kalman filter can still be applied and the resulting filter is quasi optimal. This filter is then used to obtain a quasi-likelihood function and the estimates obtained is linearly optimal. This approximation is needed because of the non-Gaussian nature of the problem, which can be compared to linearizing a non-linear function in the typical Kalman filtering applications. Duan and Simonato (1999) and Geyer and Pichler (1998) apply Kalman filter to estimate and test exponential-affine term structure models for both the Gaussian and non-Gaussian cases. For a detailed discussion see Duan and Simonato (1999) and Harvey (1989).

It is also important to mention that the CIR process also differs from standard Kalman filter application because of the non-negative constraint on the convenience yield. Following Duan and Simonato (1999) and Geyer and Pichler (1998) we modify the standard Kalman filter by simply replacing any negative element of the convenience yield estimate with zero.

3.1 State Space Formulation

From the valuation formula given by equations (14), (18) and (19), the measurement equation can be written as:

$$Y_t = d_t + Z_t[x_t, \delta_t]' + \varepsilon_t, \quad t = 1, \dots, N \quad (24)$$

where:

- $Y_t = [\ln F(\tau_i)]$, for $i = 1, \dots, n$ is a $n \times 1$ vector of observations where $F(\tau_i)$ is the observed futures price at time t for maturity τ_i . At each time t we observe n futures prices which correspond to n different maturities;
- $d_t = [A(\tau_i)]$ for $i = 1, \dots, n$ is a $n \times 1$ where $A(\cdot)$ is given by equation (19).
- $Z_t = [1, -B(\tau_i)]$, for $i = 1, \dots, n$ is a $n \times 2$ matrix where $B(\cdot)$ is calculated according to equation (18);
- ε_t is a $n \times 1$ is $n \times 1$ vector of serially uncorrelated disturbances with $E[\varepsilon_t] = 0$, $Var[\varepsilon_t] = H_t$. This vector is introduced to account for possible errors in the data. The covariance matrix H_t is taken to be diagonal for computational simplicity;

The transition equation is given by:

$$[x_t, \delta_t]' = c_t + Q_t[x_t, \delta_t]' + R_t\eta_t, \quad t = 1, \dots, NT \quad (25)$$

where:

•

$$c_t = \begin{bmatrix} \mu\Delta t \\ m(1 - e^{-\alpha\Delta t}) \end{bmatrix} \quad (26)$$

•

$$Q_t = \begin{bmatrix} 1 & -(1 + \frac{1}{2}\sigma_1^2)\Delta t \\ 0 & e^{-\alpha\Delta t} \end{bmatrix} \quad (27)$$

- R_t is a 2×2 identity matrix;
- η_t is a 2×1 vector of serially uncorrelated disturbances with $E[\eta_t] = 0$ and $Var[\eta_t] = V_t$.

The covariance matrix of η_t is given by:

$$V_t = \begin{bmatrix} \sigma_1^2 \Delta t \delta_{t-dt} & \rho \sigma_1 \sqrt{\Delta t} \sqrt{\delta_{t-dt}} \sqrt{Var[\delta_t | \delta_{t-1}]} \\ \rho \sigma_1 \sqrt{\Delta t} \sqrt{\delta_{t-dt}} \sqrt{Var[\delta_t | \delta_{t-1}]} & Var[\delta_t | \delta_{t-1}] \end{bmatrix} \quad (28)$$

where:

$$Var[\delta_t | \delta_{t-1}] = m \left(\frac{\sigma_2^2}{2\alpha} \right) (1 - e^{-\alpha \Delta t})^2 + \delta_{t-1} \left(\frac{\sigma_2^2}{\alpha} \right) (e^{-\alpha \Delta t} - e^{-2\alpha \Delta t}) \quad (29)$$

The observation and state equation matrices Z_t , d_t , H_t , Q_t , c_t and V_t depend on the unknown parameters of the model. One of the main purposes of the Kalman filter implementation is to find estimates for these parameters. This can be done by maximizing the quasi likelihood function with respect to the unknown parameters through an optimization procedure.

For notational simplicity, consider θ the vector of unknown parameters and $Y_t = \{y_t, y_{t-\Delta t}, \dots, y_t, y_{t_0}\}$ the information vector at time t , which are not independent. We assume that the distribution of Y_t conditional on $Y_{t-\Delta t}$ under the objective measure is normal with mean $\hat{Y}_{t|t-\Delta t} = E[Y_t | Y_{t-\Delta t}]$ and covariance matrix F_t . The vector of prediction errors is given by $v_t = Y_t - \hat{Y}_{t|t-\Delta t}$. The logarithm of the *quasi-likelihood function* is given by:

$$\log L(Y; \theta) = -\frac{1}{2} \frac{n(t_{final} - t_0)}{\Delta t} \log 2\pi - \frac{1}{2} \sum_t \log |F_t| - \frac{1}{2} \sum_t v_t F_t^{-1} v_t \quad (30)$$

Since both F_t and v_t depend upon θ , θ_t is chosen to maximize the quasi-likelihood function⁶. This estimation procedure is recursive and it is calculated at each time t as part of the Kalman filter.

⁶To maximize the quasi-likelihood function we used the Matlab routine "maxlik.m". This function is part of the Econometrics Toolbox by James P. LeSage and can be downloaded for free at: www.spatial-econometrics.com.

To calibrate the Gibson and Schwartz (1990) - Schwartz (1997) two-factor model we use the state-space formulation as described in the later article.

3.2 Empirical Results

The data set used in this study consists of weekly observations of New York Mercantile Exchange (NYMEX) light crude oil futures which covers the period from 17th of March 1999 to 24th of December 2003 (243 observations)⁷. At each observation we consider 7 contracts ($n = 7$) corresponding to 7 different maturities. Naturally, the time to maturity changes as we evolve in time and to force the time to maturity to stay within a narrow range we roll over the contracts during the period of observations. Table 1 describes the data used. We denote by F0 the contract closest maturity, F1 the second contract closest to maturity and so on. We assume that the interest rate, r , is equal to 0.04 and the marginal storage cost of storage is equal to 0.20.

Table 2 reports the estimation results for both our model and Gibson and Schwartz (1990) - Schwartz (1997) two-factor model. The values obtained for the parameters are comparable for both cases. The most noticeable difference lies in the value of the long-run mean for the convenience yield, m . However, this difference is approximately 0.2 which is consistent with the storage cost value of 0.2 that we assume in our model.

The speed of mean reversion in the convenience yield equation, α , and the coefficient of correlation between the spot price and convenience yield, ρ , are high and significant for both cases. The total expected return on the spot commodity, μ , and the market price of convenience yield, λ , are also positive and high. In particular, it is worth mentioning the high value of average convenience yield. This indicates that during this period the market is predominantly in backwardation. Additionally, both spot price and convenience yield volatilities are also high. This behavior in the crude oil market can be explained by the world events which took place after September 2001 and, in particular, the recent Golf

⁷The data was retrieved from the Internet on the 31st of October 2003 and on the 17th of February 2004 from *Futures Guide* TM, <http://www.futuresguide.com/index.php>. The original data set consists of daily observations. Weekly data was obtained by using every Wednesday (to avoid weekend effects) observation.

war. These events lead to increasingly uncertainty in the world markets in general and in particular in the oil supply. This uncertainty naturally rises the value of having crude oil in storage, which implies a high convenience and a market in strong backwardation.

As mentioned before, we assume a diagonal covariance structure for the measurement errors. These are denoted by v_0, v_1, v_2, v_3, v_4 and v_5 and v_6 and correspond to each of the futures contracts used, namely $F0, F1, F2, F3, F4, F5$ and $F6$ respectively. These values are also displayed in Table II. The magnitude of this errors is the same for both our model and Gibson and Schwartz (1990) - Schwartz (1997) two-factor model.

Table 1: Light Crude Oil Futures weekly data from 17th of March 1999 to 24th of December 2003.

Futures Contract	Mean Maturity (Std)	Mean Price (Std)	Volatility of Returns
F0	0.044 (0.024)	27.058 (4.711)	0.377
F1	0.127 (0.025)	26.724 (4.435)	0.346
F2	0.348 (0.024)	25.712 (3.884)	0.273
F3	0.598 (0.024)	24.735 (3.514)	0.231
F4	0.931 (0.024)	23.759 (3.173)	0.191
F5	1.181 (0.024)	23.189 (2.974)	0.176
F6	1.931 (0.024)	21.963 (2.599)	0.158

Table 2: Estimation results and standard errors (in parentheses) for both our model and the Gibson and Schwartz (1990) - Schwartz (1997) two-factor model using all the futures contracts F0, F1, F2, F3, F4, F5 and F6 from 17th of March 1999 to 24th of December 2003.

Parameters	Schwartz (1997)	
	Our model	two-factor model
m	0.526 (0.011)	0.356 (0.011)
λ	1.617 (0.008)	1.869 (0.002)
α	6.301 (0.019)	6.273 (0.002)
σ_1	0.434 (0.005)	0.441 (0.001)
σ_2	0.725 (0.004)	0.720 (0.000)
ρ	0.899 (0.003)	0.800 (0.001)
μ	0.514 (0.006)	0.500 (0.000)
v_0	0.051 (0.001)	0.050 (0.002)
v_1	0.048 (0.001)	0.052 (0.003)
v_2	0.041 (0.001)	0.036 (0.001)
v_3	0.034 (0.001)	0.028 (0.001)
v_4	0.033 (0.001)	0.026 (0.001)
v_5	0.041 (0.002)	0.031 (0.002)
v_6	0.056 (0.003)	0.040 (0.002)
Log-likelihood function	3718	3597

Figure 1: This figure illustrates the weekly evolution of the state variables implied by this model for the whole sample period. The spot price is measured in U.S. Dollar per barrel. The convenience yield is multiplied by 10 to facilitate the comparison between the state variables.

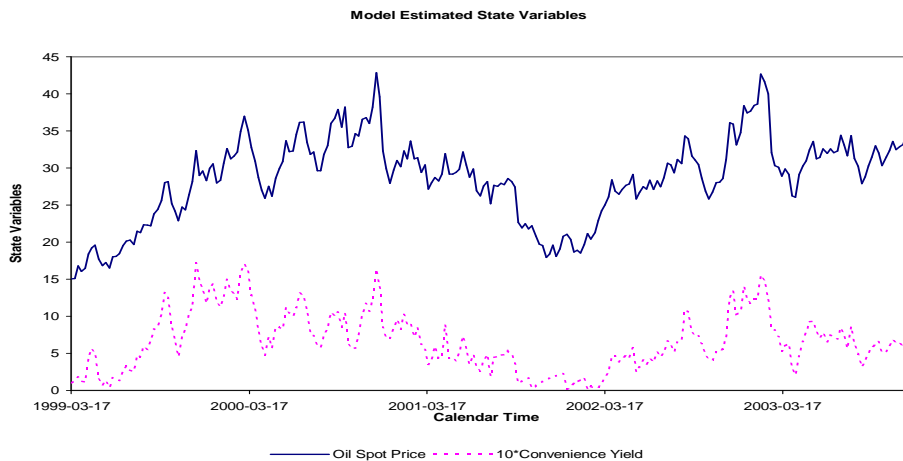
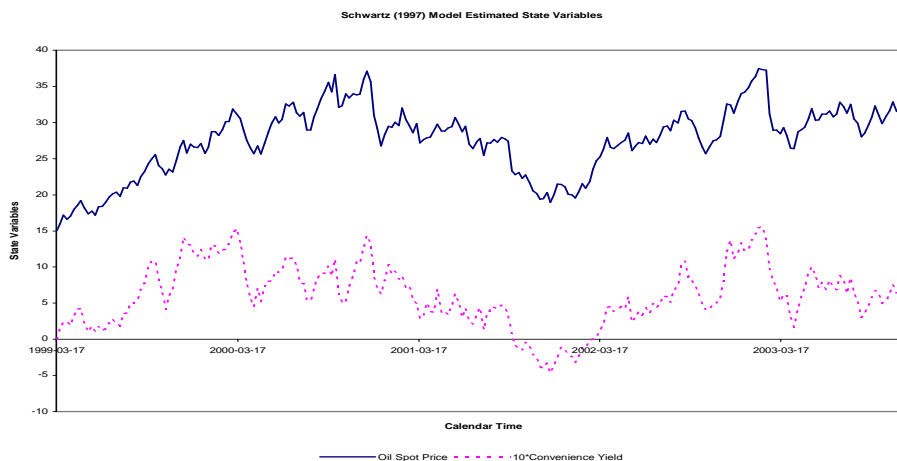


Figure 2: This figure illustrates the weekly evolution of the state variables implied by the Gibson and Schwartz (1990) - Schwartz (1997) two-factor model for the whole sample period. The spot price is measured in U.S. Dollar per barrel. The convenience yield is multiplied by 10 to facilitate the comparison between the state variables.



Figures 1 and 2 show the weekly evolution of the two state variables over the whole sample period for both this model and the Gibson and Schwartz model. In both cases, we can observe a strong correlation between the two state variables. There are two main differences between the behaviour of the state variables displayed in figures 1 and 2. The first is that the convenience yield is always non-negative for this model whereas we observe negative values for the convenience yield implied by Schwartz model. The second difference is that both state variables implied by this model present heteroskedasticity whereas the respective volatilities implied by Schwartz model are constant. Both differences are consistent with the definition of each models.

Table 3 displays the mean pricing errors (MPE) and the root mean squared errors (RMSE) for all the observations. Both error measures are small and of the same order of magnitude for both our model and Schwartz model. The values of the log-likelihood function and the pricing errors indicate that our model fits better the data but by only marginally. These results are consistent with the results presented by Nielsen and Schwartz (2004). These authors also find that both Gibson and Schwartz (1990) - Schwartz (1997) two factor model and the CIR stochastic volatility model perform almost as well. In the present paper, we also note that the performance of both models decreases

as maturity of the futures contract increases. This highlights the fact that both models become less efficient as we increase the maturity of the futures contracts.

Figures 3 and 4 illustrate examples of the evolution of the forward curve of the market prices and both models.

Table 4 and Figure 5 display the volatilities implied by the market, by the present model and by the Gibson and Schwartz (1990) - Schwartz (1997) two-factor model. For short maturities both models underestimate the market volatilities and for longer maturities, these models increasingly overestimate them. This indicates that neither model is able to fit accurately the market volatility term structure. This certainly has implications in the valuation of financial or real asset contingent on a commodity price.

Table 3: Summary statistics for both our model and Gibson and Schwartz (1990) - Schwartz (1997) two-factor model pricing errors in valuing futures contracts during the whole period 17th of March 1999 to 24th of December 2003.

Futures Contract	Our Model		Schwartz (1997) Two-factor Model	
	RMSE	MPE	RMSE	MPE
F0	0.023	0.017	0.021	-0.001
F1	0.022	-0.017	0.034	-0.020
F2	0.044	-0.036	0.034	-0.023
F3	0.027	-0.019	0.014	-0.007
F4	0.020	0.000	0.038	0.013
F5	0.033	0.015	0.034	0.024
F6	0.059	0.029	0.060	0.020
Total	0.033	-0.002	0.034	0.002

Figure 3: This figure illustrates the evolution of the forward curve for the market of futures prices and both our model and the Gibson and Schwartz (1990) - Schwartz (1997) two-factor model on the 5th of November 2002.

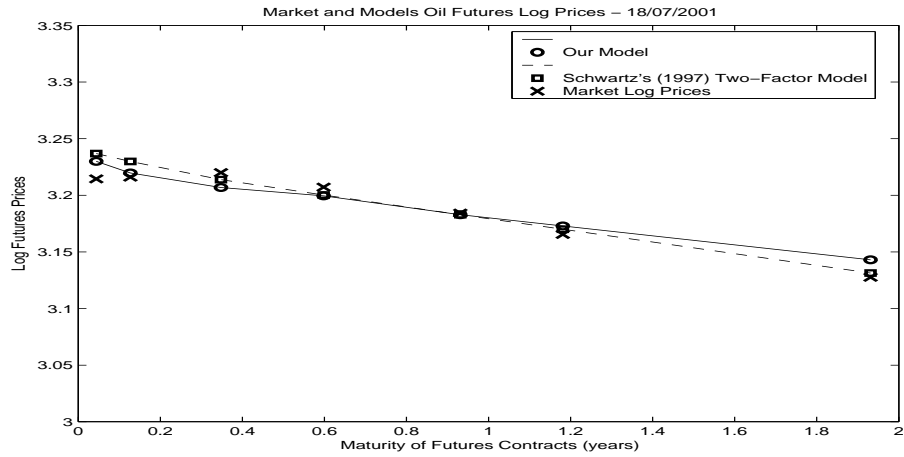


Figure 4: This figure illustrates the evolution of the forward curve for the market of futures prices and both our model and the Gibson and Schwartz (1990) - Schwartz (1997) two-factor model on the 18th of July 2001.

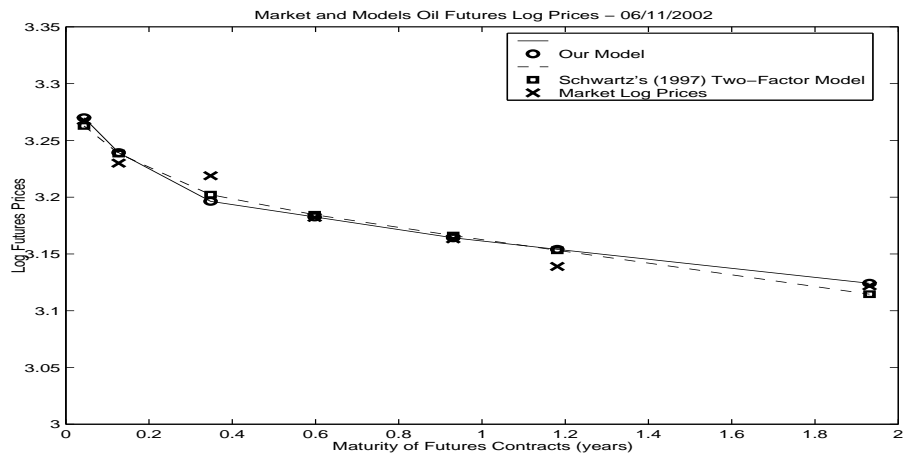
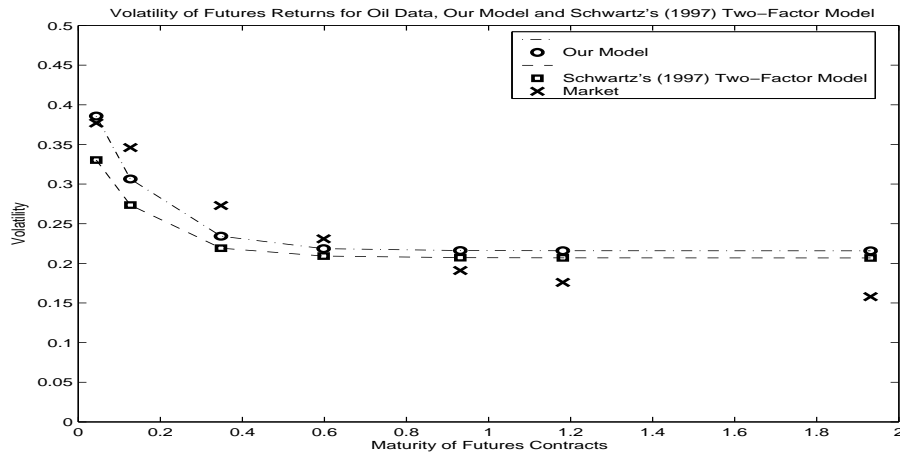


Table 4: Market, our model and the Gibson and Schwartz (1990) - Schwartz (1997) Two-Factor Model implied volatilities of annualized log-returns of futures prices.

Futures Contract	Market Volatility	Our Model Volatility	Schwartz's Model Volatility
F0	0.377	0.386	0.330
F1	0.346	0.306	0.274
F2	0.273	0.234	0.219
F3	0.231	0.219	0.209
F4	0.191	0.216	0.207
F5	0.176	0.216	0.207
F6	0.158	0.216	0.207

Figure 5: This figure illustrates the annualized volatility of futures returns implied by our model, Gibson and Schwartz (1990) - Schwartz (1997) and the market data.



4 Conclusion

In this paper we presented a two-factor model for commodity prices and the corresponding futures valuation. This model extends Gibson and Schwartz (1990) - Schwartz (1997) two factor model by adding two important features. First, the O-U process for the convenience yield is replaced by a CIR process. This allows us to maintain the mean-reverting property of the convenience yield and to additionally ensure that our model is arbitrage-free. Second, we consider that both the spot price and the convenience yield volatilities are proportional to the square root of the instantaneous convenience yield level. This specification establishes a dependency between the commodity price volatility and the inventory levels. As explained before, this property is predicted by the theory of storage and by the structural models in the literature.

We derived the analytical solution for the futures partial differential equation. By guessing a solution of the exponential affine form, we transformed the initial PDE into a system two ODEs with analytical solution. This solution also satisfies the original PDE with the appropriate boundary conditions. This method has been also applied in the interest models literature such as Hull and White (1990), Brown and Schaefer (1994) and Duffie and Kan (1993, 1996).

We implemented empirically both our model and Gibson and Schwartz (1990) - Schwartz (1997) two factor model for purposes of comparison using crude oil futures prices applying the Kalman filter. We found that the parameters estimates are analogous for both models and indicate a strong mean reversion in the convenience yield and strong backwardation. Both the 11th of September event and the recent war crisis might explain such behavior in the oil prices.

Although our model adds valuable characteristics to the existing reduced form models in the literature, it only outperforms Gibson and Schwartz (1990) - Schwartz (1997) model marginally, in terms of both the pricing errors and the value of the log-likelihood function. Both models achieve very good results when valuating short-term maturity data but their efficiency decreases when the futures maturity increases. Neither can they

reproduce the empirical volatility structure accurately. Both models underestimate short-term volatilities and over estimate longer term ones. This certainly has implications on contingent claims pricing for which the volatility is a fundamental element. Therefore, this study suggests that both our model and the Gibson and Schwartz (1990) - Schwartz (1997) models should be extended in order to reproduce accurately the volatility structure of the commodity forward curves. This is particularly relevant to evaluate long-term investments on commodities.

It is important to point out that although the present model only outperforms the Gibson and Schwartz (1990) - Schwartz (1997) model marginally, we believe that the empirical comparison is affected by the peculiarity of the data set used. Specifically, a close analysis of our data revealed that we do not observed contango for a significant period of time in this data set available to us. In fact, one of the advantages of our model in relation to the Gibson and Schwartz (1990) - Schwartz (1997) model is the exclusion of arbitrage possibilities when the forward curve is in contango. Therefore, the empirical analysis of this particular data set did not allow us to illustrate this advantage. Accordingly, one of the directions of further work is to test both models for a different commodity or for different time period. Another direction for future work is the extension of this analysis to price commodity options in order to study the implications of this model in option pricing.

A Derivation of the Solution to the Futures Partial Differential Equation

Because the solution to equation (16) depends on the solution to equation (15) we start by solving the latter.

Write $a_1 = \frac{1}{2}\sigma_2^2$ and $a_2 = \alpha - \rho\sigma_1\sigma_2$, then equation (15) becomes:

$$a_1B^2 + a_2B - 1 + B_\tau = 0, \quad \text{with } B(0) = 0. \quad (31)$$

We can write this equation in the format:

$$\frac{d}{d\tau}\Phi(\tau, B) = 0 \quad \Leftrightarrow \quad M(\tau, B) + N(\tau, B)\frac{dB}{d\tau} = 0 \quad (32)$$

if and only if there exists a function $\Phi(\tau, B)$ such that:

$$M(\tau, B) = \frac{\partial\Phi}{\partial\tau} \quad \text{and} \quad N(\tau, B) = \frac{\partial\Phi}{\partial B} \quad (33)$$

In this case we have:

$$M(\tau, B) = a_1B^2 + a_2B - 1, \quad \text{and} \quad (34)$$

$$N(\tau, B) = 1. \quad (35)$$

The equation

$$M(\tau, B) + N(\tau, B)\frac{dB}{d\tau} = 0 \quad (36)$$

is exact if and only if:

$$\frac{\partial M}{\partial B} = \frac{\partial N}{\partial \tau}. \quad (37)$$

Given that we have a non-exact equation, we need to multiply both sides of the equation by the following integrating factor:

$$\mu(B) = \exp \int Q(B)dB = \frac{a_1}{a_1B^2 + a_2B - 1} \quad (38)$$

where:

$$Q(B) = -\frac{2a_1B + a_2}{a_1B^2 + a_2B - 1}. \quad (39)$$

Equation (31) then becomes:

$$a_1 + \frac{a_1}{a_1B^2 + a_2B - 1}B_\tau = 0, \quad \text{with } B(0) = 0. \quad (40)$$

The solution to this equation is

$$B(\tau) = \frac{2(1 - e^{-k_1\tau})}{k_1 + k_2 + (k_1 - k_2)e^{-k_1\tau}}, \quad (41)$$

where

$$k_1 = \sqrt{k_2^2 + 2\sigma_2^2} \quad (42)$$

$$k_2 = (\alpha - \rho\sigma_1\sigma_2). \quad (43)$$

Accordingly, the solution to equation (16) is:

$$A(\tau) = r\tau + (\lambda - \alpha m) \int_0^\tau B(q)dq \quad (44)$$

where:

$$\int_t^T B(q)dq = \frac{2}{k_1(k_1 + k_2)} \ln \left[\frac{(k_1 + k_2)e^{k_1\tau} + k_1 - k_2}{2k_1} \right] + \frac{2}{k_1(k_1 - k_2)} \ln \left[\frac{k_1 + k_2 + (k_1 - k_2)e^{-k_1\tau}}{2k_1} \right] \quad (45)$$

where k_1 and k_2 are as before.

References

- Babbs, S. H. and Nowman, K. B. (1999). Kalman filtering of generalized vasicek term structure models. *Journal of Financial and Quantitative Analysis*, pages 115–130.
- Ball, C. A. and Tourous, W. N. (1996). Unit roots and the estimation of interest rate dynamics. *Journal of Empirical Finance*, 3:215–238.
- Bjork, T. and Landen, C. (2002). On the term structure of futures and forward prices. In Geman, Madan, Pliska, and Vorst, editors, *Mathematical Finance - Bachelier Congress 2000*. Springer-Verlag.
- Brennan, M. J. and Schwartz, E. S. (1985). Evaluating natural resource investments. *Journal of Business*, 58(2):135:157.
- Brown, R. H. and Schaefer, S. M. (1994). Interest rate volatility and the shape of the term structure. *Philosophical Transactions of the Royal Society of London A*, 347:563–576.
- Clewlow, L. and Strickland, C. (2000). *Energy Derivatives: Pricing and Risk Management*. Lacima Publications, London, England.
- Cox, J. C., Jonathan E. Ingersoll, J., and Ross, S. A. (1985). A theory of the term structure of interest rates. *Econometrica*, 53(2):385–407.
- Duan, J. C. and Simonato, J. G. (1999). Estimatinh and testing exponential-affine term structure models by kalman filter. *Review of Quantitative Finance and Accounting*, 13:111–135.
- Duffie, D. and Gray, S. (1995). *Managing Energy Price Risk*, chapter Volatility in Energy Markets. Risk Publications, London.
- Duffie, D. and Kan, R. (1993). Multi-factor term structure models. *Philosophical Transactions of the Royal Society of London, Series A*, 347:577–586.
- Duffie, D. and Kan, R. (1996). A yield factor model of interest rates. *Mathematical Finance*, 6:379–406.

- Geyer, A. L. H. and Pichler, S. (1998). A state-space approach to estimate and test multifactor cox-ingersoll-ross models of the term structure. *Journal of Financial Research*, 22:107–130.
- Gibson, R. and Schwartz, E. S. (1990). Stochastic convenience yield and the pricing of oil contingent claims. *Journal of Finance*, 45(3):959–976.
- Harvey, A. C. (1989). *Forecasting, structural time series model and the Kalman filter*. Cambridge University Press.
- Heston, S. L. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies*, 6:327–343.
- Hilliard, J. and Reis, J. (1998). Valuation of commodity futures and options under stochastic convenience yields, interest rates, and jump diffusions on the spot. *Journal of Financial and Quantitative Analysis*, 33(1):61–86.
- Hull, J. and White, A. (1990). Pricing interest-rate derivative securities. *The Review of Financial Studies*, 3:573–592.
- Litzenberger, R. H. and Rabinowitz, N. (1995). Backwardation in oil futures markets: Theory and empirical evidence. *The journal of Finance*, 50(3):1517–1545.
- Lund, J. (1997). Non linear kalman filter techniques for term structure models of the term structure of interest rates. Working Paper, University of Aarhus.
- Manoliu, M. and Tompaidis, S. (2002). Energy futures prices: Term structure models with kalman filter estimation. *Applied Mathematical Finance*, 9:21–43.
- Miltersen, K. R. and Schwartz, E. S. (1998). Pricing of options on commodity futures with stochastic term structures of convenience yields and interest rates. *Journal of Financial and Quantitative Analysis*, 33(1):33–59.
- Ng, V. K. and Pirrong, C. (1994). Fundamentals and volatility: Storage, spreads and the dynamics of metal prices. *Journal of Business*, 67(2):203–230.

- Nielsen, M. J. and Schwartz, E. S. (2004). Theory of storage and the pricing of commodity claims. *Review of Derivatives Research*, 7(1):5–24.
- Pirrong, S. C. (1998). Price dynamics and derivatives prices for continuously produced, storable commodities. Working Paper, Washington University.
- Routledge, B., Seppi, D. J., and Spatt, C. (2000). Equilibrium forward curves for commodities. *The Journal of Finance*, 55(3):1297–1338.
- Schwartz, E. S. (1997). The stochastic behaviour of commodity prices: Implications for valuation and hedging. *Journal of Finance*, 52(3):923–973.
- Schwartz, E. S. and Smith, J. E. (2000). Short-term variations and long-term dynamics in commodity prices. *Management Science*, 46(7):893–911.