Residual model for Futures prices.

Journal of Futures

Authors: Marcos Escobar, Luis Seco

Abstract

In this paper we study the modelling of the term structure of futures prices from commodities. A factor-model for the term structure of futures prices is presented aimed at providing flexibility on the drivers. This model allows a good fitting of curve movements as for example changes in slope and concavity. We study the estimation of the parameters of the model based on Kalman-filter techniques. On the other hand, the factors' processes are found based on the absence of arbitrage. We propose and price derivatives where the underlyings are the model factors.

1 Introduction.

The stochastic behaviour of spot and futures prices for a commodity is of paramount importance to analysts interested in risk management or pricing contingent claims (derivatives) on the underlying commodity. Analysts make explicit assumptions about the stochastic behaviour of the spot and futures prices when they model the underlying assets. The accuracy of these models depends on the accuracy of the assumptions. Commodity prices present a variety of features, which make them a very special case in the financial market. They show a very flexible term structure (TS) of futures prices ($F_{t,T}$), even richer than the term structure of bond prices (i.e. changes in slope and curvature in time to maturity, $T-t$ are quite common); which implies a higher level of complexity for risk management and derivative pricing purposes.

It is important to be aware of some important features of this term structure:

- The future price for a commodity, at time t, can be observed for fixed maturity times T_i (time to maturity $T_i - t$). For example, in the case of oil, T_i would correspond to 20th of month, for as long as 12 months.
- futures prices for more than nine months of time-to-maturity are rarely observed. Consequently, the prices of futures contracts with maturities greater than nine months are frequently only an extrapolation of the prices of shorter maturity contracts computed from actual trades.

The standard modelling of this vector of data $(F(t,T_1),...,F(t,T_n))$, is based on assuming them as n points of a curve on R^n for each time t. Then this curve is modeled as a function of several stochastic factors. All proposed models in the literature assume this procedure, (see next section).

Early studies in this area typically assumed that all uncertainty arose from the spot price of the commodity. (This is the model of stock price uncertainty underlying the famous Black-Scholes option pricing formula and it leads to closed-form solutions for many derivatives prices). See for example, Schwartz (1982) and Brennan and Schwartz (1985). Recognition of the importance of the variability of the spreads between spot and futures prices (term structure) led to the development of several multi-factor models of commodity prices. For example, Gibson and Schwartz (1990) introduced a two-factor model where the spot price of the commodity and the convenience yield [1] followed a joint stochastic process. Schwartz (1997) presented a three-factor model where the logarithm of the spot price, the convenience yield and interest rates followed mean reverting processes.

Based on the fact that many find the notion of convenience yield elusive, Schwartz -Smith (2000) proposed a different approach to commodity modelling. They considered a model with two factor, called short term and equilibrium term which, although equivalent to Gibson-Schwartz (1990) model, leads to analytic results that are more transparent and allows a simplification of the analysis of long-term investment. Ross (1995) and Pilopovic (1998) have also recently developed bi-factor price models.

One of the most prominent features of the term structure of commodity (such as energy, agricultural products and metals), futures prices is the sudden change in sign of its slope and convexity (see Figure 3.2), which translate into interesting movements of the futures prices curve. This behaviour is generally known as backwardation (decreasing) and contango [2] (increasing). Models in the literature do not

---

[1] the flow of services that accrues to the holder of the physical commodity, but not to the owner of a contract for future delivery.

[2] Contango: A market state where futures prices are above expected spot prices and fall as maturity approaches. Backwardation: A market state where futures prices are below expected spot prices and rise as maturity approaches.

explicitly address this problematic nor do they allow for a wider set of underlying factors to be used. This is why, the finding of a flexible set of factors capable of explaining important features (movements) of the term structure of future prices is the main focus of this paper.

This paper is organized as follows: Subsection 1.1 presents a survey of some standard model for futures prices. In section 2, a (minimum) three factors model is develop under the historical measure (P), these new factors are studied showing simple stochastic behaviour and second-moment dependency structures. The estimation of the model parameters and factors is addressed in subsection 2.2. Our model shows a flexible fitting, for up to 1 year time-to-maturity, of the term structure of futures prices (see subsection 2.3). The risk neutral (arbitrage free, Q-measure) processes for these underlying factors is found in section 3. Several examples of contingent claims (derivatives) based on these factors are provided in section 5, aimed at hedging against contango and backwardation.

The developments in these sections attempt to provide a complete picture of the two leading problems of the commodity world, which are forecasting and pricing of commodity variables.

1.1 Futures Prices Models.

Let us first provide some notation.

- $F_{t,T}$: Future price, at time t of a contract for delivery at time T.
- S_t: Spot price of the commodity at time t.
- $r(t,T)$: Interest rate at time t for $T-t$ time to maturity, bootstrapped from the zero curve.
- $\varepsilon(t,T)$: Convenience yield: The benefit from the ownership of the physical

commodity that may include the ability to profit from temporary local storages or the ability to keep a production process running, at time-t for (T-t)-time to maturity.

- The instantaneous interest rate and instantaneous convenience yield are defined as $\lim_{T\to t} r(t,T)$,$\lim_{T\to t} \varepsilon(t,T)$ respectively.

We present some of the leading models for commodities, Ross 1995, Gibson-Schwartz 1990, Schwartz 1997, Miltersen and Schwartz 1998, Urich 2000 and Schwartz-Smith 2000. The processes for the factors are given under the risk neutral Q-measure, the difference with respect to the P-measure, is the addition of a term called "market price of risk" (see appendix **??**), usually denoted by λ.

Notice, the uncertainty of futures prices is described by modelling the stochastic behaviour of some underlying factors. It is also interesting to note that in real-life practice, the number of factors have been reduced mainly to three: spot price, interest rate and convenience yield. There have recently appeared a number of models based on new meaningful factors underlying the term structure. An example is the approach of Schwartz and Smith 2000, which is based on finding the relationship between futures prices and a two factors model (called short term/equilibrium term).

Ross Model, 1995: The Ross model posits:

$$dX_t = k(\mu - \frac{\sigma^2}{2k} - X_t)dt + \sigma dW \tag{1}$$

Ornstein-Uhlenbeck (OU process). X_t denote the log of the current spot price (S_t). Then:

$$\ln(F_{t,T}) = e^{-k(T-t)}X_t + (1 - e^{-k(T-t)})\alpha + \frac{\sigma^2}{4k}(1 - e^{-2k(T-t)}) \tag{2}$$

It assumes only one source of randomness.

Gibson-Schwartz Model, 1994:

$$dX_t = (\mu - \varepsilon_t - \frac{\sigma_1^2}{2})dt + \sigma_1 dW_1 \tag{3}$$

$$d\varepsilon_t = [k(\alpha - \varepsilon_t) - \lambda]dt + \sigma_2 dW_2 \tag{4}$$

W_1 and W_2 are standard Brownian motion with $dW_1 dW_2 = \rho dt$, ε_t denotes the instantaneous convenience yield.

$$\ln(F_{t,T}) = X_t + \frac{(1 - e^{-k(T-t)})\alpha}{k}\varepsilon_t + A(T-t) \tag{5}$$

$A(T-t)$ denotes a deterministic function of time to maturity and the model's parameters.

Schwartz Model (1997):

$$dS_t = (r_t - \varepsilon_t)S_t dt + \sigma_1 S_t dW_1 \tag{6}$$

$$d\varepsilon_t = k(\alpha - \varepsilon_t)dt + \sigma_2 dW_2 \tag{7}$$

$$dr_t = a(m - r_t)dt + \sigma_3 dW_3 \tag{8}$$

W_1, W_2 and W_3 are standard Brownian motion with $dW_1 dW_2 = \rho_1 dt$, $dW_1 dW_3 = \rho_3 dt$, $dW_3 dW_2 = \rho_2 dt$.

$$\ln(F_{t,T}) = X_t + \frac{(1 - e^{-k(T-t)})\alpha}{k}\varepsilon_t + \frac{(1 - e^{-m(T-t)})a}{m}r_t + C(T-t) \tag{9}$$

Here, $C(T-t)$ is a deterministic function on $T-t$. Notice that the drivers in this relation $(1, \frac{(1-e^{-k(T-t)})\alpha}{k}, \frac{(1-e^{-\alpha(T-t)})\alpha}{k}, C(T-t))$ are not freely chosen, but obtained from non-arbitrage conditions.

Miltersen and Schwartz model (1998):

This model is an extension of the HJM 1992 framework for bonds. It uses all the information in the initial term structures of both interest rate and commodity futures prices.

$$f(t,T) = f(0,s) + \int_0^t \mu_f(u,s)du + \int_0^t \sigma_f(u,s)dW_u \tag{10}$$

$$\varepsilon(t,T) = \varepsilon(0,s) + \int_0^t \mu_\varepsilon(u,s)du + \int_0^t \sigma_\varepsilon(u,s)dW_u \tag{11}$$

$$S(t) = S(0) + \int_0^t \mu_s(u)S_u du + \int_0^t \sigma_s(u,s)S_u dW_u \tag{12}$$

W_u is a standard d-dimensional independent Brownian process for every differential equation. Correlations among the three processes come via the specification of the diffusion terms (σ's).

The conditions of no arbitrage completely determine the drift terms (μ's):

$$\mu_s(t) = f(t,t) - \varepsilon(t,t) \tag{13}$$

$$\mu_f(t) = \sigma_f(t,s) \int_t^s \sigma_f(t,v)dv \tag{14}$$

$$\begin{aligned} \mu_\varepsilon(t) &= \sigma_f(t,s)(\int_t^s \sigma_f(t,v)dv) + (\sigma_f(t,T) - \sigma_\varepsilon(t,T)) \\ &\times (\sigma_s(t) + \int_t^T (\sigma_f(t,s) - \sigma_\varepsilon(t,s))ds \end{aligned} \tag{15}$$

The expression for the futures prices of the commodity is given by:

$$F_{t,T} = S_t \exp(\int_t^T (f(t,s) - \varepsilon(t,s))ds) \tag{16}$$

This is a framework, thus it suffers from too much generality as does the Heath Jarrow and Merton (HJM) 1991 proposal for bonds; volatilities in the model have to be provided, so it does not proposed a specific model for capturing the term structure.

Schwartz-Smith Model (2000):

$$\ln(S_t) = \chi_t + \xi_t \tag{17}$$

$$d\chi_t = (-k\chi_t - \lambda_\chi)dt + \sigma_\chi dW_\chi \tag{18}$$

$$d\xi_t = (\mu_\xi - \lambda_\xi)dt + \sigma_\xi dW_\xi \tag{19}$$

Where W_χ and W_ξ are standard Brownian motions with $dW_\chi dW_\xi = \rho dt$. χ_t will be referred to as the short-term deviation in prices (temporary changes in prices that are not expected to persist) and ξ_t the equilibrium price level (fundamental changes that are expected to persist). These factors are almost orthogonal in their dynamics, which implies a small correlation between their stochastic increments.

$$\ln(F_{t,T}) = e^{-k(T-t)}\chi_t + \xi_t + A(T-t) \tag{20}$$

Urich Model (2000):

This model is an extension, to metals, of the Garbade (1996) work on the term structure of interest rates.

$$F_{t,T} = A_0(T-t) + \sum_{i=1}^{I} w_i(t)A_i(T-t) \tag{21}$$

$$A_i(T-t) = \sum_{j=1}^{J} b_{ij}(T-t)^{j-1} \tag{22}$$

$$A_0(T-t) = \sum_{j=1}^{J_1} b_{0j}(T-t)^{j-1} \tag{23}$$

In particular $F_{t,t} = S_t = b_{00}\sum_{i=1}^{I} b_{i0}w_i(t)$ which yields the spot price behaviour.

Here w_i are random walks with zero drift and unit variance per year; the random walks are statistically independent of each other. The functions A_i are referred to as modes of fluctuations (drivers). The degree of the underlying polynomials for forecasting purposes, could be anything, however for pricing purposes it should satisfy the following arbitrage free requirement.

$$I > 2,\ I = J,\ J_1 = 2J + 1 \tag{24}$$

Notice, it put important constraints on the factors w_i (independence).

2 Historical Measure Model.

The development of risk management methodologies for multivariate markets (e.g. futures prices, bond prices) relies on the assumption that some chosen underlying market factors follow diffusions (likely mean reverting processes, see previous section). Finding new suitable sets of underlyings (risk factors) and therefore handmade relations between curves and factors, appears to be a strong way to capture and handle complex behaviours in the curve [3].

On the other hand, while most efforts have been devoted to the explanation of the futures prices term structure in a risk-neutral world (Q-measure), little has been stated regarding the term structure behaviour under the P-measure (physical world). The modelling in this measure is vital for effective risk management of financial portfolios. For example, if we were interested in computing the h-days-horizon Value at Risk (see appendix A.I) of a portfolio made of future prices' derivatives, we would need to model the underlyings (future prices) using historical data of future prices, otherwise we would not be aware of the real risk of the portfolio, but rather of the mild risk from the Q-measure.

In this section we present and develop a multi-factor model for futures prices under the historical, observable measure. Let us assume that we have a filtered probability space, $(\Omega, \mathbb{F}, (\mathbb{F}_t)_{t\geq 0}, P)$. We also assume a finite time horizon T^* with $\mathbb{F} = \mathbb{F}_{T^*}$, all definitions and statements are understood to be valid only until this time horizon T^*. Let $E_t^P[\bullet]$ denote the conditional expectation under the measure

---

[3] This approach is equivalent to a nonlinear transformation of the original space, such that better behaved factors and dependency structure are found.

P measure conditional to the information at date t, $\mathbb{F}_t$. All equations between stochastic variables are to be understood as almost surely equations under the given probability measure. See appendix **??** for more details.

2.1 A Proposed Model

The model for $F_{t,T}$ should belong to a family capable of making a good fit to the observable curve (till actual time t). [4] In other words, we can try to fit the curve at every t by using a convenient function of $T-t$, for example a polynomial. Similar to the ideas of Urich 2000 (metals) and Garbade 1990 (interest rates), several changes are made in the structure of the model, leading to the following proposal:

$$F_{t,T} = S_t + \xi_t \cdot (T-t) + \eta_t \cdot (T-t)^2 + \chi_{t,T} \tag{25}$$

The model is specified by the coefficients (factors) $\chi_{t,T}, S_t, \xi_t, \eta_t$ (stochastic processes on t) which may be seen as the parameters of a linear regression ($\chi_{t,T}$ residual), (see estimation procedure), or the Taylor expansion of futures prices in terms of $T-t$, $F_{t,T} = F_{t,t} + \frac{\partial F_{t,T}}{\partial T}|_{T=t}(T-t) + \frac{\partial^2 F_{t,T}}{\partial^2 T}|_{T=t}(T-t)^2 + \chi_{t,T}$.

The most important reason for this selection is the meaning of those factors: slope and convexity. Besides, due to the fact that the factor ξ_t can be seen as the first derivative of futures prices with respect to $T-t$, it will give us an insight regarding the contango and backwardation movements (changes in the first derivative).

---

[4] This property holds regardless of the measure considered, either evolution of the whole curve (in t) under the historical measure or in the measure implied by the absence of arbitrage

The continuous setting would be, (Ito processes, see **??**):

$$d\chi_{t,T} = f_\chi(\xi_t, \eta_t, t, T)dt + \sigma_{0,t,T}dW_t \tag{26}$$

$$dS_t = f_S(\xi_t, \eta_t, t)dt + \sigma_{1,t}dW_t \tag{27}$$

$$d\xi_t = f_\xi(\xi_t, \eta_t, t)dt + \sigma_{2,t}dW_t \tag{28}$$

$$d\eta_t = f_\eta(\xi_t, \eta_t, t)dt + \sigma_{3,t}dW_t \tag{29}$$

where W_t is a d-dimensional vector of independent Brownian motions. Correlations come via the specification of the diffusions volatilities. The integrands $(f_\chi, f_S, f_\eta, f_\xi)$, $(\sigma_{0,t,T}, \sigma_{i,t})$ are predictable processes that are regular enough to allow for:

- Differentiation under the integral sign.
- Interchange of the order of integration.
- Partial derivatives with respect to the T-variable.
- Bounded for almost all $w \in \Omega$.

Example:

$$\begin{aligned} f_S &= k_S(\theta_{S,t} - S), & f_\xi &= k_\xi(\theta_{\xi,t} - \xi) \\ f_\eta &= k_\eta(\theta_{\eta,t} - \eta), & f_\chi &= a_0 + a_1 \cdot (T - t) + a_2 \cdot (T - t)^2 \end{aligned} \tag{30}$$

Regarding the discrete framework, we propose the following simple models, based on historical data of the underlyings (figure 3.4):

- $$\begin{aligned} \chi_{t,T} &= 0, & S_t &= a_0 + S_{t-1} + res_{t,0} \\ \xi_t &= a_1 + \xi_{t-1} + res_{t,1}, & \eta_t &= a_2 + \eta_{t-1} + res_{t,2} \end{aligned} \tag{31}$$
- The residuals follow a normal distribution, $res_t \sim N(0, \Sigma_*)$.

Remark 1. *This framework deals with a multivariate normal distribution, at any time t, instead of a lognormal for the future prices. The strength of this framework lies in the time dependency of both the mean and the covariance matrix, which allows a better fitting (smaller errors and better forecasting, see section 1.2.3) of the futures prices term structure than other frameworks.*

2.2 Estimation Procedure.

One of the difficulties in the empirical implementation of futures prices models is that frequently the factors or state variables of these models are not directly observable. For some commodities, the spot price is hard to obtain, and the futures contract closest to maturity is used as proxy for the spot price. The problems of estimating the instantaneous convenience yield are even more complex; normally, futures prices with different maturities are used to compute it. The instantaneous interest rate is also not directly observable.

In this section, we address the estimation procedure for our model. In this case, as in most practical cases, the underlyings are not directly observable (we can get a proxy by using a Taylor expansion interpretation), so they would have to be estimated as well. A robust method for estimating factors and parameters is the Kalman filter (see Schwartz-Smith 2000, Harvey 1989, Duffie-Stanton 2004). The Kalman filter is a recursive procedure for computing estimates of unobserved state variables based on observations that depend on these state variables. Given a prior distribution on the initial values of the state variables and a model describing the likelihood of the observations as a function of the true values, the Kalman filter generates updated posterior distributions for these state variables in accordance to Bayes' rule.

We perform a Kalman filter estimation of model 1.40 with the diffusion following 1.45-1.48 and constant (non-zero) volatilities, see appendix A for details. The reason for using this more elegant model will be better understood in the pricing section. It will allow us to compute the market price of risk and thus to determine completely these processes in the absence of arbitrage.

2.3 Empirical Results

We worked with 12 time series of oil futures prices from January 1990 until December 2000; each series corresponds to a specific maturity month e.g. we have the futures prices for $T_i - t$, (i=January,...,December) time to maturity for each time-t. A fitting of the future curve by a three factor exponential model is performed in this section. This fitting allows us to study the negligibility of the residual factor.

Fitting a Three Factor Model

We perform a regression of the future price curves (up to one year, which are actually the observable ones) on a quadratic polynomial. This regression is on the maturity time, T:

$$F(t,T) = \sum_{i=0}^{2} \chi_i(t) \cdot (T-t)^i + \varepsilon(t,T) \qquad (32)$$

This analysis will show the insignificance of the residual factor $\chi_{t,T}$. Note that $\varepsilon(t,T)$ is a good proxy for the residual $\chi_{t,T}$ form settings $1.1.44-1.48$ and $1.49-1.52$. We study the R-square of a quadratic polynomial fitting obtained from the regressions. The R-square, also called correlation coefficient, is a widely accepted measure of the quality of a least squares fit. The better the fit the more insignificant the error $\varepsilon(t,T)$, which is a reasonable proxy for the residual factor $\chi_{t,T}$.

We performed both fittings to 770 futures prices curves $(F_{t,T_1}, \ldots, F_{t,T_{12}})$, with dates (t) from 1991 to 2000. The following table (I) shows several statistics of the

R^2 (correlation coefficient) obtained from those fittings (QP-quadratic polynomial).

	mean	Min	0.01 percentile	0.05 percentile	0.95 percentile
QPol.	0.9750	0.3772	0.7156	0.8858	0.999

.

The average of the R^2, 0.97, but more importantly the 95 percent confidence interval $[0.88, 0.999]$ for the value of R^2 imply a good fitting. Thus, the residual is basically insignificant in the explanation of the future prices.

3 Arbitrage Free Model

In the previous section, we proposed a polynomial fitting for the term structure of commodity futures prices (1.40). As previously mentioned, the stochastic processes for the factors depends on the objective: risk management (forecasting) was covered in previous section; pricing derivatives is the focus of this section.

Derivative pricing is always a case of determining the price today, t, of a given function of the evolution of several underlying price factors, based on no-arbitrage conditions. We will focus on the particular case where a specific period for this evolution $[t, T]$ is given (contingent claim with date of maturity T, see appendix **??**). The stochastic process for the underlying factor(s), from t to T, in the absence of arbitrage, may be different from the one obtained by any sound statistical tools using historical (previous to t) data. This is why the key objective in derivative pricing is finding the factor's stochastic processes between t and T.

Let us assume that we have a filtered probability space, $(\Omega, \mathbb{F}, (\mathbb{F}_t)_{t\geq 0}, P)$. We also assume a finite time horizon T^* with $\mathbb{F} = \mathbb{F}_{T^*}$, all definitions and statements are understood to be only valid until this time horizon T^*. Let $E^Q[\bullet|\mathbb{F}_t]$ (which will also appear as $E^Q[\bullet|t]$ or $E^Q_t[\bullet]$) denote the conditional expectation under the measure Q measure conditional on the information at date t, $\mathbb{F}_t$. All equations between stochastic variables are to be understood as almost surely equations under the given probability measure. See appendix **??** for more details.

3.1 Equivalent Measure.

We have already seen that, under the real (historical) measure, the initial curve has a kind of polynomial shape (low order) on the time to maturity. In this section, we assume the model 1.40 for the futures prices and then we search for the conditions on the underlyings factors $(\chi_{t,T}, S_t, \xi_t, \eta_t)$ under the equivalent measure that make the futures prices arbitrage free.

We address this problem by using the standard procedure, which can be summarize as follows. The specification of the drifts for the underlying processes in the absence of arbitrage will follow from the standard arbitrage argument: the futures prices process must satisfy the following condition: $F_{s,T} = E_s^Q[F_{t,T}]$ under the equivalent measure Q (the futures prices is a martingale in the measure induced by bond prices as numeraire, Duffie 1992). The change of measure, from P to Q, which makes futures prices driftless, is induced by Girsanov's theorem. This would imply a drift change on each of the factors' processes involved in the futures prices' decomposition (1.40).

The following theorem provides the processes for the factors in the absence of arbitrage.

Theorem 1. *Assume the following model for the futures prices (1.40):*

$$F_{t,T} = S_t + \xi_t \cdot (T-t) + \eta_t \cdot (T-t)^2 + \chi_{t,T}, \tag{33}$$

the factors follow diffusion under the historical measure:

$$\begin{aligned} d\chi_{t,T} &= f_\chi(\xi_t,\eta_t,t,T)dt + \sigma_{0,t,T}dW_t, \quad & dS_t &= f_S(\xi_t,\eta_t,t)dt + \sigma_{1,t}dW_t \\ d\xi_t &= f_\xi(\xi_t,\eta_t,t)dt + \sigma_{2,t}dW_t, & d\eta_t &= f_\eta(\xi_t,\eta_t,t)dt + \sigma_{3,t}dW_t, \end{aligned} \tag{34}$$

where W_t *is a d-dimensional* ($d > 2$) *vector of independent Brownian motions.*

$$\begin{aligned} f_S &= k_S(\theta_{S,t} - S), \qquad & f_\xi &= k_\xi(\theta_{\xi,t} - \xi) \\ f_\eta &= k_\eta(\theta_{\eta,t} - \eta) \\ f_\chi &= \textstyle\sum_{i=0}^{d-1} a_i \cdot (T-t)^i, & \sigma_{0,t,T} &= \textstyle\sum_{i=0}^{d-1} b_i \cdot (T-t)^i \end{aligned} \tag{35}$$

Here k_i *are constants,* θ_i *are time dependent functions and* $a_i, b_i, \sigma_{0,t,T}, \sigma_{i,t}$*,* $i = 1,2,3$*) are* $1 \times d$ *time dependent vectors. The volatilities* ($\sigma_{i,t}$*,* $i = 0,\dots,m$*) are measurable and fulfill Lipschitz and growth conditions (such that an unique solution of the stochastic differential equations exists, see appendix A).*

Let us assume that the $d \times d$ *matrix* $\sigma(t)$ *is invertible for all t, where* $\sigma(t)$ *is defined by the following equation:*

$$\sum_{i=0}^{d-1} b_i \cdot (T-t)^i - \sigma_{1,t} - \sigma_{2,t} \cdot (T-t) - \sigma_{3,t} \cdot (T-t)^2] = \sigma(t) \cdot (1, T, \dots, T^{d-1})' \tag{36}$$

Then there exists an unique market price of risk vector λ*, and thus, an unique martingale measure* Q*, such that the futures prices are arbitrage free. Moreover, the factors' stochastic processes in the absence of arbitrage (under the Q-measure) are the following:*

$$dS_t = g_{1,t}dt + \sigma_{1,t}dW_t^Q \tag{37}$$

$$d\xi_t = g_{2,t}dt + \sigma_{2,t}dW_t^Q \tag{38}$$

$$d\eta_t = g_{3,t}dt + \sigma_{3,t}dW_t^Q \tag{39}$$

$$d\chi_{t,T} = (\xi_t + 2\eta_t(T-t) - g_{1,t} - g_{2,t}(T-t) - g_{3,t}(T-t)^2)dt + \sigma_{0,t,T}dW_t^Q, \tag{40}$$

where $g_{1,t} = f_S - \lambda \cdot \sigma_{1,t}$*,* $g_{2,t} = f_\xi - \lambda \cdot \sigma_{2,t}$ *and* $g_{3,t} = f_\eta - \lambda \cdot \sigma_{3,t}$ *denote the difference between the drift under P-measure and the market price of risk, d-dimensional vector* λ*. Moreover* g_i *are linear functions of the factors.*

Proof. Let us denote the process under the P-measure for the futures prices as:

$$dF_{t,T} = \mu_F(t,T)dt + \sigma_F(t,T)dW_t, \tag{41}$$

where dW is a d-dim vector of independent Brownian motions. As shown by Harrison and Kreps (1979), Harrison and Pliska (1981) and Duffie (1996), the absence of arbitrage in the future contracts market is equivalent to the existence of a d-dimensional column vector process $(\lambda(t), ..., \lambda_d(t))'$, (market price of risk) such that $\forall t$, $0 < t < T$,

$$\mu_F(t,T) = -\sigma_F(t,T) \cdot \lambda(t) \tag{42}$$

This lambda vector is the vector θ in Girsanov's theorem (see appendix A), therefore the change of measure would be:

$$dW_t = dW_t^Q + \lambda(t)dt \tag{43}$$

Let us denote the factors' processes as:

$$dX_s^i = f_i(s, X_s)ds + \sigma_i(s)dW_s \quad i = 1, 2, 3 \tag{44}$$

$$dR_{s,T} = f_R(s, T, X_s)ds + \sigma_R(s,T)dW_s, \tag{45}$$

where $R_{s,T} = \chi_{s,T}$, $X_s^1 = S_s$, $X_s^2 = \xi_s$ and $X_s^3 = \eta_s$. For convenience we denote $X_t^T = (X_{t,T}, X_t^1, X_t^2, X_t^3)$.

Applying Ito's lemma to 33 follows:

$$\mu_F(t,T) = -X_t^2 - 2X_t^3 \cdot (T-t) + f_1 + (T-t) \cdot f_2 + (T-t)^2 \cdot f_3 + f_R \tag{46}$$

$$\sigma_F(t,T) = \sigma_R - \sigma_1 - \sigma_2 \cdot (T-t) - \sigma_3 \cdot (T-t)^2] \tag{47}$$

The condition of no-arbitrage (1.90) and the previous expressions for the drift and volatility of the futures prices and the residual (1.94 and 1.95) provide a system of equations for λ (the market price of risk does not depend on maturity time T, hence the coefficients of the polynomial in T, $(1, T, T^2, ..., T^{d-1})$ have to be zero). To see this, we express the futures prices' drift and volatility as follows:

$$\mu_F(t,T) = \mu(t,X) \cdot \mathbf{T} \tag{48}$$

$$\sigma_F(t,T) = \sigma(t) \cdot \mathbf{T}, \tag{49}$$

where $\mathbf{T}$ is a $1 \times d$ vector of powers of T, $\sigma(t)$ is a $d \times d$ matrix and $\mu(t,X)$ is a $1 \times d$ vector, therefore equation 1.90 becomes:

$$(\lambda \cdot \sigma(t) - \mu(t,X)) \cdot \mathbf{T} = 0 \tag{50}$$

which implies:

$$\lambda \cdot \sigma(t) = \mu(t,X) \tag{51}$$

$$\lambda = \mu(t,X) \cdot \sigma^{-1}(t) \tag{52}$$

Note that λ would be a linear function of the factors X^i (because $\mu(t,X)$ is linear).

Let us assume that the drift of the stochastic processes for the factors (g_1, g_2 and g_3) are known under Q then we can compute the drift for the residual, such that the process for F(t,T) is a martingale, (zero drift under Q):

$$F(s,T) = E_s^Q(F_{t,T}) \ , \quad s < t < T \tag{53}$$

Let

$$dR_{s,T} = g_R(s,T,X_s)ds + \sigma_R(s,T)dW_s^Q \tag{54}$$

$$dX_s^i = g_i(s,X_s)ds + \sigma_i(s)dW_s^Q \quad i = 1,2,3 \tag{55}$$

As before W_s^Q is a $d - dim$ vector of independent Brownian motions. $X_t^T = (R_{t,T}, X_t^1, X_t^2, X_t^3)$. Let us denote:

$$u(s, X_s^T) = F_{s,T} = R_{s,T} + X_s^1 + X_s^2(T-s) + X_s^3(T-s)^2 \tag{56}$$

Applying Ito's lemma and the zero drift martingale condition for future prices:

$$0 = -\frac{\partial u}{\partial s}(s, X^T) + \frac{1}{2}\sum \sigma_i(s)\sigma_j(s)\frac{\partial^2 u}{\partial X^{i,T}\partial X^{j,T}}u + \sum g_i(s, T, X^T)\frac{\partial u}{\partial X^{i,T}}, \quad (57)$$

where

$$\frac{\partial u}{\partial s}(s, X^T) = 0 + X^2 + 2X^3(T-s) \quad (58)$$

$$\frac{\partial u}{\partial R_{,T}}(s, X^T) = 1; \frac{\partial u}{\partial X^1}(s, X^T) = 1; \frac{\partial u}{\partial X^2}(s, X^T) = T-s; \frac{\partial u}{\partial X^3}(s, X^T) = (T-s)^2 \quad (59)$$

Putting this back in equation 1.105:

$$0+X^2+2X^3(T-s) = g_R(s, T, X)+g_1(s, X)+g_2(s, X)\cdot(T-s)+g_3(s, X)\cdot(T-s)^2 \quad (60)$$

Which completely determines g_R as:

$$g_R(s, T, X) = X^2 + 2X^3(T-s) - g_1(s, X) - g_2(s, X)\cdot(T-s) - g_3(s, X)\cdot(T-s)^2. \quad (61)$$

□

Remark 2. *Note that firstly, the previous results apply as long as (*$\theta_i, k_i, a_i, b_i, f_\chi, \sigma_{0,t,T}, \sigma_{i,t}$*,* $i = 1, 2, 3$*) are measurable and fulfill Lipschitz and growth conditions (such that an unique solution of the stochastic differential equations exists), thus they might be functions of* t, S, ξ, η*. In this case, the factors drifts* $g_i, i = 1, 2, 3$ *under* Q *might not be linear on the factors (affine). Nevertheless we still require* σ_t*, which will now depend on the factors as well as time, to be invertible. Secondly, having calibrated the processes under* P*, we can easily obtain the processes under* Q *by simply computing* λ *from the system of equations 1.99 and then* g_1*,*g_2 *and* g_3 *follow.*

Remark 3. *We extend this result by considering powers greater than 2 in the drivers:*

$$F_{t,T} = \sum_{i=0}^{m} A_i(T-t) \cdot \chi_{i,t} + \chi_{t,T} \tag{62}$$

$$A_i(T-t) = (T-t)^i \tag{63}$$

The previous theorem remains intact, with the exception that the stochastic processes for factors and residual under the Q-measure would be as follow:

$$d\xi_{i,t} = g_{i,t}dt + \sigma_{i,t}dW_t^Q \tag{64}$$

$$d\chi_{t,T} = \{\sum_{i=0}^{m} \{g_{i,t} \cdot (T-t)^i - \chi_{i,t} \cdot (T-t)^{i-1}\}\}dt + \sigma_{t,T}dW_t^Q \tag{65}$$

g is the difference (by Girsanov theorem) between the observed drift and the market price of risk λ_t multiplied by the volatility, $g_i = \mu_{i,t} - \sigma_{i,t}\lambda$. In this case λ is unique as long as the residual drift and volatility are polynomials of order $d-1$ and $d > m$.

Remark 4. *Note that the drift of the spot price under the Q-measure is the factor ξ. This result from using the well-known result that the drift of the spot price under non-arbitrage conditions is:*

$$F(t,t) \cdot \frac{\partial \ln(F_{t,T})}{\partial T}|_{T=t}. \tag{66}$$

(see Duffie 2001) It is easy to verify that

$$\frac{\partial \ln(F_{t,T})}{\partial T}|_{T=t} = \frac{\xi_t}{S_t}. \tag{67}$$

Remark 5. *We could also work with a model where the underlying factors are functions of time and maturity time:*

$$F_{t,T} = \sum_{i=1}^{n} A_i(T-t) \cdot \chi_{i,t,T}. \tag{68}$$

This variant would avoid the notion of residual. However we would have to assume the factors processes are known under the arbitrage-free measure and then compute the feasible expressions for the drivers A_i. See Chapter 2, section 2.2.4 for an analysis of this framework.

Remark 6. *We may always select the drivers and factors we want, arbitrage can be avoided by taking a convenient residual factor (either deterministic or random). The residual's volatility, $\sigma_{0,t,T}$ may very well be zero; this would be the case where the factors completely explain the futures volatility.*

Note that the idea of our model could be extended even further by considering:

$$\Phi(F_{t,T}) = \sum_{i=1}^{m} A_i(T-t) \cdot \chi_{i,t} + R_{t,T} \tag{69}$$

where Φ could be taken as a positive function, e.g. ln. In this case, Schwartz 1997 model, where three factors are used, can be obtained as a particular case of this setting. e.g.

3.2 Fitting of Volatilities and Correlations

We are interested in modelling not only the term structure of futures prices, but also the term structure of volatilities and correlations [5].

Proposition 1. *The instantaneous volatility of the futures prices $F_{t,T}$, under the*

---

[5] 4 Each model considered in previous articles (see Schwartz 1997, 1998, Urich 2000) has different implications not only for the term structure of futures prices but also for the term structure of volatilities of futures prices and term structure of correlations of futures prices

assumptions of theorem 1, is given by:

$$\begin{aligned} V[F_{t,T}] = {} & \|\sigma_{3,t}\|^2 \cdot (T-t)^4 + \|\sigma_{2,t}\|^2 \cdot (T-t)^2 + \|\sigma_{0,t,T}\|^2 + \|\sigma_{1,t}\|^2 \\ & + 2\sigma_{0,t,T} \cdot \sigma_{3,t} \cdot (T-t)^2 + 2\sigma_{2,t} \cdot \sigma_{3,t} \cdot (T-t)^3 \\ & + 2\sigma_{0,t,T} \cdot \sigma_{2,t} \cdot (T-t) + 2\sigma_{1,t} \cdot \sigma_{3,t} \cdot (T-t)^2 \\ & + 2\sigma_{1,t} \cdot \sigma_{2,t} \cdot (T-t) + 2\sigma_{0,t,T} \cdot \sigma_{1,t} \end{aligned} \tag{70}$$

Proof. It follows from computing the variance of a linear combination of random variables in equation 1.40 (Notice that the instantaneous volatility was denoted $\sigma_F(t,T)$, see 1.89):

$$\begin{aligned} V[F_{t,T}] = {} & V[\eta_t] \cdot (T-t)^4 + V[\xi_t] \cdot (T-t)^2 + V[S_t] + V[\chi_{t,T}] \\ & + 2Cov[\chi_{t,T}, \eta_t] \cdot (T-t)^2 + 2Cov[\chi_{t,T}, \xi_t] \cdot (T-t) \\ & + 2Cov[\xi_t, \eta_t] \cdot (T-t)^3 + 2Cov[S_t, \eta_t] \cdot (T-t)^2 \\ & + 2Cov[S_t, \xi_t] \cdot (T-t) + 2Cov[\chi_{t,T}, S_t] \end{aligned} \tag{71}$$

□

In the case of constant volatilities, this would be a fourth degree polynomial; which should provide a good fit to the volatility term structure of futures prices; the correlations between the factors will determine whether this curve is increasing or decreasing. In practice, a decreasing pattern will characterize most commodities. (see Schwartz 1998).

It is also interesting to check whether this model captures the term structure of correlations (which is decreasing in time-to-maturity for most commodities, see Schwartz 1997).

Proposition 2. *The instantaneous correlations of the future price $F_{t,T}$ at time t, maturing at T and T_1, under the assumptions of theorem 1, are given by:*

$$Corr[F_{t,T}, F_{t,T_1}] = \frac{Cov[F_{t,T}, F_{t,T_1}]}{V[F_{t,T}], V[F_{t,T_1}]^{\frac{1}{2}}} \tag{72}$$

Where the covariance is given by:

$$\begin{aligned}Cov[F_{t,T}, F_{t,T_1}] = {} & \|\sigma_{0,t,T}\|^2 + \|\sigma_{1,t}\|^2 + (T-t)\cdot\sigma_{0,t,T}\cdot\sigma_{2,t} + (T-t)\cdot\sigma_{1,t}\cdot\sigma_{2,t} \\ & \sigma_{0,t,T}\cdot\sigma_{1,t} + (T_1-t)\cdot\sigma_{0,t,T}\cdot\sigma_{2,t} + (T_1-t)\cdot\sigma_{1,t}\cdot\sigma_{2,t} \\ & + (T-t)^2\cdot\sigma_{0,t,T}\cdot\sigma_{3,t} + (T_1-t)^2\cdot\sigma_{0,t,T}\cdot\sigma_{3,t} + (T-t)^2\cdot\sigma_{1,t}\cdot\sigma_{3,t} \\ & + (T_1-t)^2\cdot\sigma_{1,t}\cdot\sigma_{3,t} + \|\sigma_{2,t}\|^2(T-t)(T_1-t) \\ & + [(T-t)(T_1-t)^2 + (T_1-t)(T-t)^2]\cdot\sigma_{2,t}\cdot\sigma_{3,t} + (T-t)^2(T_1-t)^2\cdot\|\sigma_{3,t}\|^2.\end{aligned} \tag{73}$$

Proof. The expression for the covariance follows from using the model 1.40 and the fact that covariance is a bilinear function. □

This is a quotient of polynomials in two variables of order 4 (for constant volatilities); the correlation on the factors will determine whether this structure is increasing or decreasing.

Proposition 3. *In the setting of Theorem 1, the distribution of $F_{t,T}$, given a σ-algebra $\mathbb{F}_0$, is Normal under the Q-measure. Its mean is $F_{0,T}$ and its variance is as follows:*

$$\begin{aligned}V[F_{t,T}|0] = {} & \int_0^t ||\sigma_{3,y}||^2 dy \cdot (T-t)^4 + 2\int_0^t (\sigma_{2,y}\cdot\sigma_{3,y})\, dy\cdot(T-t)^3 \\ & + \int_0^t (||\sigma_{2,y}||^2 + 2\sigma_{0,y,T}\cdot\sigma_{3,y})dy\cdot(T-t)^2 + 2\int_0^t (\sigma_{0,y,T}\cdot\sigma_{2,y})\, dy\cdot(T-t) \\ & + 2\int_0^t \sigma_{1,y}\cdot\sigma_{3,y})dy\cdot(T-t)^2 + 2\int_0^t (\sigma_{1,y}\cdot\sigma_{2,y})\, dy\cdot(T-t) + 2\int_0^t (\sigma_{0,y,T}\cdot\sigma_{1,y})\, dy \\ & + \int_0^t ||\sigma_{0,y,T}||^2 dy + \int_0^t ||\sigma_{1,y}||^2 dy.\end{aligned}$$

Proof. The conditional normality comes from the multivariate normality of the underlyings (see Bjork 1998). It is well known that n-dimensional linear stochastic differential equations,

$$dX_t = (A_t \cdot X_t + b_t)dt + \sigma_t dW_t \tag{74}$$

where A_t is an $n \times n$ matrix deterministic function, b and σ are R^n value deterministic function, are multivariate normal distributed. The solution of this SDE is (see Ikeda-Watanabe 1981):

$$X_t = \Psi_{t,0} \cdot x_0 + \int_0^t \Psi_{t,s} \cdot \sigma_s dW_s, \tag{75}$$

where the deterministic matrix function $\Psi_{\cdot,\cdot}$ is called the fundamental matrix of $A_\cdot$, it satisfies:

$$\frac{\Psi_{t,s}}{dt} = A_t \cdot \Psi_{t,s} \tag{76}$$

$$\Psi_{s,s} = I. \tag{77}$$

It follows that $F_{t,T}$ is a linear combination of normal distributions.

Now, we compute the expectation and variance. From the martingale property of futures prices under the Q-measure, it follows:

$$E^Q[F_{t,T}|0] = E^Q[\chi_{t,T}/0] + E^Q[S_t|0] + E^Q[\xi_t|0] \cdot (T-t) + E^Q[\eta_t|0] \cdot (T-t)^2 = F_{0,T} \tag{78}$$

$$\begin{aligned}\sigma^2(t,T) = V[F_{t,T}/0] &= V[\chi_{t,T}/0] + V[S_t/0] + V[\xi_t/0] \cdot (T-t)^2 \\ &+ 2(T-t)Cov[\chi_{t,T}, \xi_t|0] + 2(T-t)Cov[S_t, \xi_t|0] + 2Cov[\chi_{t,T}, S_t|0] \\ &+ V[\eta_t|0] \cdot (T-t)^4 + 2(T-t)^3 Cov[\xi_t, \eta_t|0] \\ &+ 2(T-t)^2 Cov[\chi_{t,T}, \eta_t|0] + 2(T-t)^2 Cov[S_t, \eta_t|0].\end{aligned}$$

The expectation, variance and covariance given time 0, for each of those factors is the following:

$$E^Q[(\chi_{t,T}, S_t, \xi_t, \eta_t)/0] = \Psi_{t,0} \cdot (\chi_{0,T}, S_0, \xi_0, \eta_0) \tag{79}$$

$$Cov^Q[(\chi_{t,T}, S_t, \xi_t, \eta_t)/0] = \int_0^t \Psi_{t,s} \cdot \sigma_s \cdot \sigma_s' \cdot \Psi_{t,s}' ds. \tag{80}$$

□

4 Applications. Commodities Derivatives.

In this section, we first take advantage of the normality of our model in order to price well known derivatives. In the second part we pursue the creation of derivatives whose underlyings represent mathematical features of the term structure. We present a new family of derivatives that can be named as conditional derivatives, the payoff conditioning a set of random variables on the value of a second correlated set of random variables.

In order to find the price at time zero of a contingent claim on a future contract starting at t with maturity T, we need to find the conditional expectation and volatility, under the Q-measure, given information at a given initial point 0. This is what was provided in previous proposition.

Remark 7. *Due to the normality framework, we can obtain closed form solutions for several well known derivatives as European Option on a future contract and Arithmetic Asian Options on spot or future contracts (where the payoff at t is $\left\{\int_0^t F_{u,T-u} du - K\right\}^+$).*

4.1 Derivatives Based on the Proposed Model.

In this section we propose derivatives that aim to protect and take advantage of contango (Co) and backwardation (Ba) movements. These derivatives will be based on the new underlyings created, and will keep track of those movements.

We basically create payoff, using spot or futures prices, conditioning on the sign or the magnitude of the new underlyings' factors (slope, curvature). The idea is that a company may want protection against some inconvenient movements of the term structure curve while it benefits from other convenient movements. [6]

Examples:

1. A contract with maturity date T, which gives the right to sell (buy) a future contract, with maturity $T_1 > T$, if the market will be in contango, $\xi > 0$. These kinds of derivatives will protect the owner against changes from Co to Ba.

2. Option, maturity date T. It gives the right to sell a future contract, with maturity $T_1 > T$, if the spot prices continue decreasing ($\xi < 0$). These kinds of derivatives will protect from sudden spot prices arises.

$$\begin{cases} F_{T,T_1} & \xi_T < 0 \\ S_T & \xi_T \geq 0 \end{cases} \neq (S_T - F_{T,T_1})^+ \tag{81}$$

$$O_t = E_t^Q \left[e^{-r\cdot(T-t)} \cdot \left\{ F_{T,T_1} \cdot 1_{\{\xi(T)<0\}} + S_T \cdot 1_{\{\xi(T)\geq 0\}} \right\} \right] \tag{82}$$

3. The following derivative protects against a concave downward curve. Notice that we can create a very similar derivative by conditioning onto ξ instead of

---

[6] The idea of creating derivatives respect to the underlyings may be extended to other models, i.e Schwartz 1997. But there are two problems with this extension: firstly the meaning of the underlyings may not be clear (affecting the usefulness of the derivatives); secondly these derivatives can be constructed by using already available derivatives from Futures and Bonds, so there is no need for creating them.

η (protecting against a decreasing period on the curve).

$$\begin{cases} F_{T,T_1} & \eta_T > 0 \\ S_T + \xi_T \cdot (T_1 - T) & \eta_T \leq 0 \end{cases} \tag{83}$$

$$O_t = E_t^Q \left[e^{-r\cdot(T-t)} \cdot \left\{ F(T,T_1) \cdot 1_{\{\eta(T)<0\}} + [S_T + \xi_T \cdot (T_1 - T)] \cdot 1_{\{\eta(T)\geq 0\}} \right\} \right] \tag{84}$$

4. Lookback Option (payoff):

$$\left(\max_{u\leq T} \{S_u + \xi_u \cdot (T_1 - u)\} - F_{T,T_1} \right)^+ . \tag{85}$$

   It chooses between the future price and the value of a future at the time of its maximum, assuming no curvature.

5. Lookback Option (payoff):

$$= \left(\max_{u\leq T} \{S_u\} + (T_1 - T) \cdot \max_{u\leq T} \{\xi_u\} + (T_1 - T)^2 \cdot \max_{u\leq T} \{\eta_u\} - F_{T,T_1} \right)^+ . \tag{86}$$

   It takes the best slope and curvature in a period instead of the observed future price.

6. Lookback conditional Option (payoff):

$$\begin{cases} S_T + (T_1 - T) \cdot \max_{u\leq T} \{\xi_u\} + (T_1 - T)^2 \cdot \eta_T & \max_{u\leq T} \{\xi_u\} > a \\ S_T + (T_1 - T) \cdot a + (T_1 - T)^2 \cdot \eta_T & \max_{u\leq T} \{\xi_u\} \leq a \end{cases} \tag{87}$$

   It considers the best slope in a period regardless of the curvature (zero).

7. Lookforward Option (payoff):

$$\left(\max_{T_1 \geq u \geq T} \{F_{T,u}\} - F_{T,T_1} \right)^+ \tag{88}$$

   It provides the best future price available for the period $[T, T_1]$.

8. American slope (contango) Option: it gives the right to buy an underlying future price (with fixed maturity day) as soon as ξ (or η, or both) becomes negative.

This sequence can be enlarged depending on the particular needs of a company. It is aimed as a starting point.

A Estimation Details.

This appendix provides a method to estimate the parameters and state variables of our model 1.40.

Kalman Filter Estimation

In order to ease notation, we will assume the following diffusion processes for the factors:

$$d\chi_{t,T} = (a_0 + a_1 \cdot (T-t) + a_2 \cdot (T-t)^2)dt + \sigma_0 dW_{0,t} \tag{89}$$

$$dS_t = (B_S - K_S \cdot S)dt + \sigma_1 dW_{1,t} \tag{90}$$

$$d\xi_t = (B_\xi - K_\xi \cdot \xi)dt + \sigma_2 dW_{2,t} \tag{91}$$

$$d\eta_t = (B_\eta - K_\eta \cdot \eta)dt + \sigma_3 dW_{3,t}, \tag{92}$$

where the Brownian motions $W_{i,t}$ are correlated. We proceed by first finding the mean vector and covariance matrix for a discrete-time approximation of the above processes and then take the limit as the time steps (length $\triangle t = \frac{t}{n}$) are made infinitesimally small (see Harvey 1989, Schwartz and Smith 2000).

$$E[(\chi_{t,T}, S_t, \xi_t, \eta_t)|0] = \tag{93}$$

$$\begin{pmatrix} \chi_{0,T} + (a_0 + a_1 \cdot T + a_2 \cdot T^2) \cdot t - (a_1 + 2a_2 \cdot T - 0.5) \cdot t^2 + (a_2) \cdot t^3 \\ e^{-K_S \cdot t} \cdot S_0 + B_S \cdot \frac{1-e^{-K_S \cdot t}}{K_S} \\ e^{-K_\xi \cdot t} \cdot \xi_0 + B_\xi \cdot \frac{1-e^{-K_\xi \cdot t}}{K_\xi} \\ e^{-K_\eta \cdot t} \cdot \eta_0 + B_\eta \cdot \frac{1-e^{-K_\eta \cdot t}}{K_\eta} \end{pmatrix} \tag{94}$$

$$V[(\chi_{t,T}, S_t, \xi_t, \eta_t)|0] = \tag{95}$$

$$\begin{pmatrix} \sigma_0^2 \cdot t & \frac{(1-e^{-K_S \cdot t})\rho_{01} \cdot \sigma_0 \cdot \sigma_1}{K_S} & \frac{(1-e^{-K_\xi \cdot t})\rho_{02} \cdot \sigma_0 \cdot \sigma_1}{K_\xi} & \frac{(1-e^{-K_\eta \cdot t})\rho_{03} \cdot \sigma_0 \cdot \sigma_1}{K_\eta} \\ - & (1 - e^{-2K_S \cdot t})\frac{\sigma_1^2}{2K_S} & \frac{(1-e^{-K_S \cdot t})(1-e^{-K_\xi \cdot t})\rho_{12}\sigma_1 \cdot \sigma_2}{K_S \cdot K_\xi} & \frac{(1-e^{-K_S \cdot t})(1-e^{-K_\eta \cdot t})\rho_{13}\sigma_1 \cdot \sigma_3}{K_S \cdot K_\eta} \\ - & - & (1 - e^{-2K_\xi \cdot t})\frac{\sigma_1^2}{2K_\xi} & \frac{(1-e^{-K_\xi \cdot t})(1-e^{-K_\eta \cdot t})\rho_{23}\sigma_3 \cdot \sigma_2}{K_\eta \cdot K_\xi} \\ - & - & - & (1 - e^{-2K_\eta \cdot t})\frac{\sigma_1^2}{2K_\eta} \end{pmatrix}. \tag{96}$$

The discrete evolution of the factors (state variables) is described by the **transition equations**, which from B.6 and B.8 can be written as:

$$x_{t,T} = c + G \cdot x_{t-\triangle t,T} + w_t \tag{97}$$

$$t = 1, \ldots, n_T, \tag{98}$$

where

- $x_{t,T} \equiv [\chi_{t,T}, S_t, \xi_t, \eta_t]$
- $c \equiv [\triangle t, B_S \cdot \triangle t, B_\xi \cdot \triangle t, B_\eta \cdot \triangle t]$
- $G \equiv \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & e^{-K_S \cdot \triangle t} & 0 & 0 \\ 0 & 0 & e^{-K_\xi \cdot \triangle t} & 0 \\ 0 & 0 & 0 & e^{-K_\eta \cdot \triangle t} \end{pmatrix}$

- $\triangle t \equiv$ the length of the time steps.
- w_t is a 4×1 vector of serially uncorrelated, normally distributed disturbances with $E[w_t] = 0$ and $Var[w_t] = W \equiv Cov[x_{\triangle t}]$ given by B.8.
- $n_T \equiv$ the number of time periods in the data set.

The **measurement equation** describes the relationship between the state variables and the observed prices, from equation 1.40, this is:

$$y_{t,T} = q_t' \cdot x_{t,T} + e_{t,T} \tag{99}$$

$$t = 1, \ldots, n_T \tag{100}$$

where:

- $y_{t,T}$ observed futures prices with maturity T_i. where $i = 1, \ldots, n$.
- $q_t \equiv \begin{pmatrix} 1 & 1 & T-t & (T-t)^2 \end{pmatrix}$ an 4×1 vector.
- $e_t = [e_{t,T_1}, \ldots, e_{t,T_n}]$ a n×1 vector of serially uncorrelated, normally distributed disturbances with: $E[e_t] = 0$, $Cov[e_t] = V$.

Everything in this formulation is derived directly from our model with the exception of the introduction of the measurement errors (e_t). These (e_t) have two common interpretations, representing errors in the reporting of prices (perhaps due to price quotes) or, as errors in the model's fit to observed prices. Given these equations and a set of observed futures prices ($y_{t,T}$, $t = 1, ..., n_T$, $T = T_1, \ldots, T_n$), the Kalman filter is run recursively beginning with a prior distribution on the initial values of the state variables ($x_{0,T} = [\chi_{0,T}, S_0, \xi_0, \eta_0]$). We assume the prior distribution is multivariate normal, the mean vector and covariance matrix are $m_{0,T}$, $C_{0,T}$ respectively. In each subsequent period, the observation $y_{t,T}$ and the previous periods mean vector and covariance matrix are used to calculate the posterior mean vector and covariance matrix for the then-current state variables. The mean and covariance of the state variables conditioned on all of the information available at time t are given by:

$$E[x_{t,T}] = m_{t,T} \equiv a_{t,T} + A_{t,T} \cdot (y_{t,T} - f_{t,T}) \tag{101}$$

$$V[x_{t,T}] = C_{t,T} \equiv R_{t,T} - A_{t,T} \cdot Q_{t,T} \cdot A_{t,T} \tag{102}$$

where

- $a_{t,T} \equiv c + G \cdot m_{t-1,T}$ is the mean of $x_{t,T}$ based on what is know at period $t-1$.
- $R_{t,T} \equiv G_t \cdot C_{t-1,T} \cdot G_t + W$ is the covariance of $x_{t,T}$ based on information at period $t-1$.
- $f_{t,T} \equiv q_t \cdot x_{t,T}$ is the mean of the period-t futures prices given what is known at period $t-1$.
- $Q_{t,T} \equiv q_t \cdot R_{t,T} \cdot q_t + V$ is the covariance of the period-t futures prices given what is known at period $t-1$.
- The matrix $A_{t,T} = R_{t,T} \cdot q_t \cdot Q_{t,T}^{-1}$ defines a correction to the predicted state variables ($a_{t,T}$) based on the difference between the (log) prices observed at time t, $y_{t,T}$, and the predicted time-t price vector, $f_{t,T}$.

Estimation of State Variables.

After running the Kalman filter for a while, the variance in the state variable estimates [7] (given by B.14) will approach an asymptotic value that is independent of the particular price sequence observed or the assumed prior distributions. If there

---

[7] The accuracy with which the state variables can be estimated depends on the kind and quality of information observed.

is uncertainty about the level of the state variables, the forecasts of the state variables must be augmented to reflect this additional uncertainty. If we let $\widehat{x}_{0,T}$ denote the mean of the current state variable (given by equation (B.13) as $m_{t,T} = \widehat{x}_{0,T}$) and $\widehat{\sigma}_i^2, \widehat{\rho}_{i,j}$ the corresponding variances and correlation coefficient (defined by the covariance matrix $C_{t,T}$ given by equation B.14), then, from equation (B.6, B.8), the mean and variance for the state variables at time t are given by:

$$E[(\chi_{t,T}, S_t, \xi_t, \eta_t)|0] = \tag{103}$$

$$\begin{pmatrix} \widehat{\chi}_{0,T} + (a_0 + a_1 \cdot T + a_2 \cdot T^2) \cdot t - (a_1 + 2a_2 \cdot T - 0.5) \cdot t^2 + (a_2) \cdot t^3 \\ e^{-K_S \cdot t} \cdot \widehat{S}_0 + B_S \cdot \frac{1-e^{-K_S \cdot t}}{K_S} \\ e^{-K_\xi \cdot t} \cdot \widehat{\xi}_0 + B_\xi \cdot \frac{1-e^{-K_\xi \cdot t}}{K_\xi} \\ e^{-K_\eta \cdot t} \cdot \widehat{\eta}_0 + B_\eta \cdot \frac{1-e^{-K_\eta \cdot t}}{K_\eta} \end{pmatrix} \tag{104}$$

$$V[(\chi_{t,T}, S_t, \xi_t, \eta_t)|0] = \tag{105}$$

$$\begin{pmatrix} \sigma_0^2 \cdot t & \frac{(1-e^{-K_S \cdot t})\rho_{01} \cdot \sigma_0 \cdot \sigma_1}{K_S} & \frac{(1-e^{-K_\xi \cdot t})\rho_{02} \cdot \sigma_0 \cdot \sigma_1}{K_\xi} & \frac{(1-e^{-K_\eta \cdot t})\rho_{03} \cdot \sigma_0 \cdot \sigma_1}{K_\eta} \\ - & \left(1 - e^{-2K_S \cdot t}\right)\frac{\sigma_1^2}{2K_S} & \frac{(1-e^{-K_S \cdot t})(1-e^{-K_\xi \cdot t})\rho_{12}\sigma_1 \cdot \sigma_2}{K_S \cdot K_\xi} & \frac{(1-e^{-K_S \cdot t})(1-e^{-K_\eta \cdot t})\rho_{13}\sigma_1 \cdot \sigma_3}{K_S \cdot K_\eta} \\ - & - & \left(1 - e^{-2K_\xi \cdot t}\right)\frac{\sigma_1^2}{2K_\xi} & \frac{(1-e^{-K_\xi \cdot t})(1-e^{-K_\eta \cdot t})\rho_{23}\sigma_3 \cdot \sigma_2}{K_\eta \cdot K_\xi} \\ - & - & - & \left(1 - e^{-2K_\eta \cdot t}\right)\frac{\sigma_1^2}{2K_\eta} \end{pmatrix} \tag{106}$$

$$+ \begin{pmatrix} \widehat{\sigma}_0^2 \cdot t & \frac{(e^{-K_S \cdot t})\widehat{\rho}_{01} \cdot \widehat{\sigma}_0 \cdot \widehat{\sigma}_1}{K_S} & \frac{(e^{-K_\xi \cdot t})\widehat{\rho}_{02} \cdot \widehat{\sigma}_0 \cdot \widehat{\sigma}_1}{K_\xi} & \frac{(e^{-K_\eta \cdot t})\widehat{\rho}_{03} \cdot \widehat{\sigma}_0 \cdot \widehat{\sigma}_1}{K_\eta} \\ - & \left(e^{-2K_S \cdot t}\right)\frac{\widehat{\sigma}_1^2}{2K_S} & \frac{(e^{-K_S \cdot t})(e^{-K_\xi \cdot t})\widehat{\rho}_{12}\widehat{\sigma}_1 \cdot \widehat{\sigma}_2}{K_S \cdot K_\xi} & \frac{(e^{-K_S \cdot t})(e^{-K_\eta \cdot t})\widehat{\rho}_{13}\widehat{\sigma}_1 \cdot \widehat{\sigma}_3}{K_S \cdot K_\eta} \\ - & - & \left(e^{-2K_\xi \cdot t}\right)\frac{\widehat{\sigma}_1^2}{2K_\xi} & \frac{(e^{-K_\xi \cdot t})(e^{-K_\eta \cdot t})\widehat{\rho}_{23}\widehat{\sigma}_3 \cdot \widehat{\sigma}_2}{K_\eta \cdot K_\xi} \\ - & - & - & \left(e^{-2K_\eta \cdot t}\right)\frac{\widehat{\sigma}_1^2}{2K_\eta} \end{pmatrix}. \tag{107}$$

Comparing this with Equation (B.6, B.8), we see that the uncertainty about the current state variables serves to increase uncertainty about their future values by adding terms to the covariance matrix. Although there may be considerable uncertainty about the values of the state variables at any time, this uncertainty has relatively little impact on forecasts and futures prices. If we observe prices for a vector of futures contracts with varying maturities, there will typically be little uncertainty about the state variables. In fact, if we observe prices for two contracts with different maturities and have zero measurement error, we can invert equation (B.11) and estimate the state variables exactly. With multiple contracts and measurement errors for all contracts, we cannot estimate the state variables exactly, but if the measurement errors are small, there will be little uncertainty in the state variable estimates. In the empirical results of the next section, there is essentially zero error in the state variable estimates.

Parameter Estimation.

The Kalman filtering procedure allows us to estimate the state variables over time given particular assumptions about the parameters of the process; all of the previous probabilistic results assumed that the parameters of the process were known. The Kalman filtering paradigm also allows one to efficiently calculate the likelihood of a set of observations given a particular set of parameters (see, Harvey 1989, Chapter 3.4 for details). By varying the parameters and rerunning the Kalman filter for each set of parameters, we can identify the set of parameters that maximizes this likelihood function. In our model, there are several model parameters to be estimated $(a_0, a_1, a_2, B_S, B_\xi, B_\eta, K_S, K_\xi, K_\eta)$, $(\sigma_0, \sigma_1, \sigma_2, \sigma_3, \rho_{ij})$, $i \neq j = 0, \ldots, 3$ plus the terms in the covariance matrix for the measurement errors (V). In general, there are $(n+1)n/2$ free variables in the covariance matrix, where n is the number of futures contracts whose prices are observed (the matrix must be symmetric). We simplify the estimation problem by assuming that V is diagonal with diagonal elements $(s_1^2, \dots, s_n^2)$. In all cases, we started the Kalman filter with a prior mean $(m_{0,T})$ and covariance matrix $(C_{0,T})$ based on the observed means and covariance in the data. Although the likelihood scores vary somewhat, the estimated state variables and parameters did not appear to be very sensitive to the assumed initial mean and covariance.

FIGURES and TABLES

Exhibit 1, Oil futures prices. Contango and Backwardation

Exhibit 2, Behavior through time to maturity. Increasing, decreasing, concave upward and downward.

Exhibit 3, Model's factors, spot, slope and curvature.

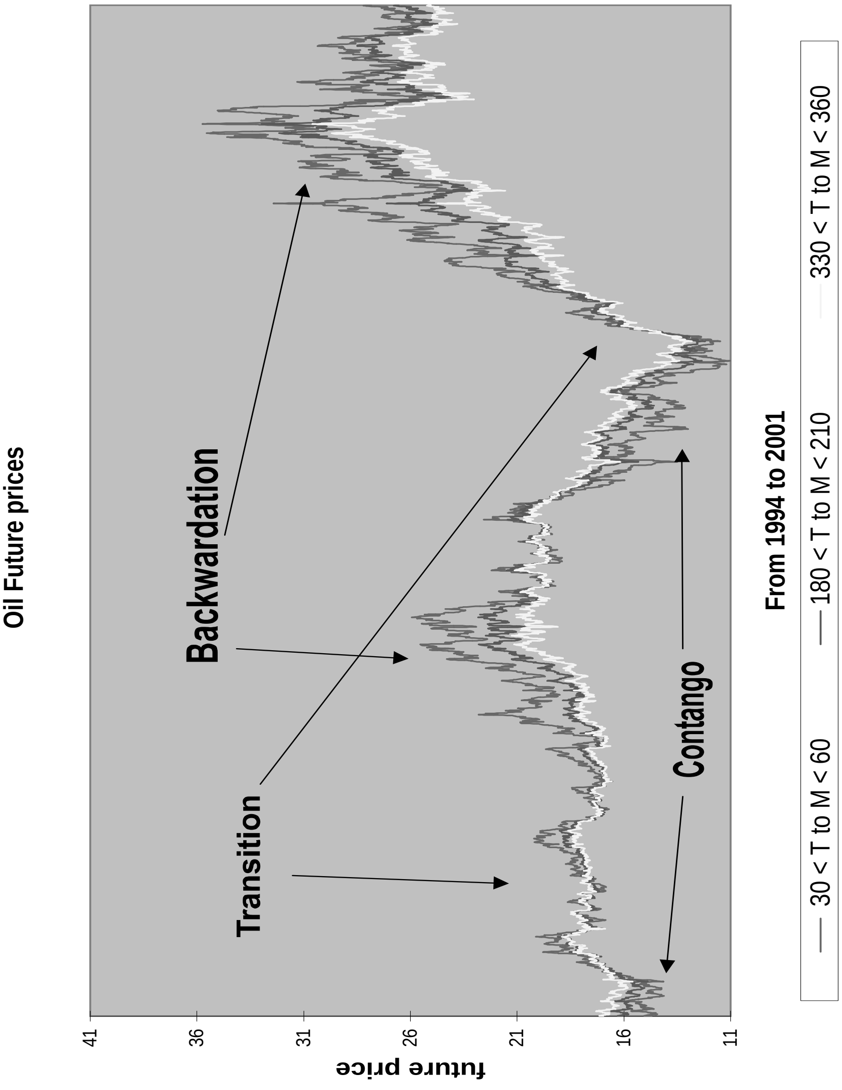

Figure 1: Oil futures prices.

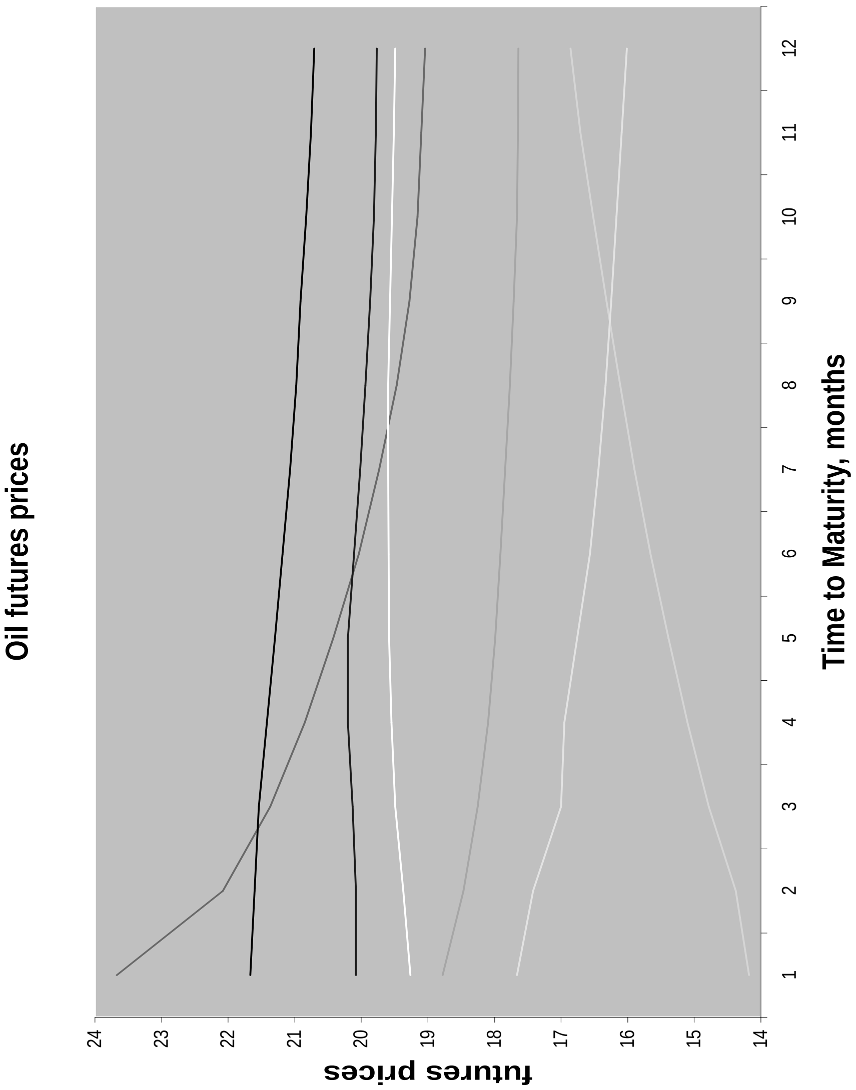

Figure 2: Behavior through time to maturity. Contango (concave upward) and backwardation (concave downward).

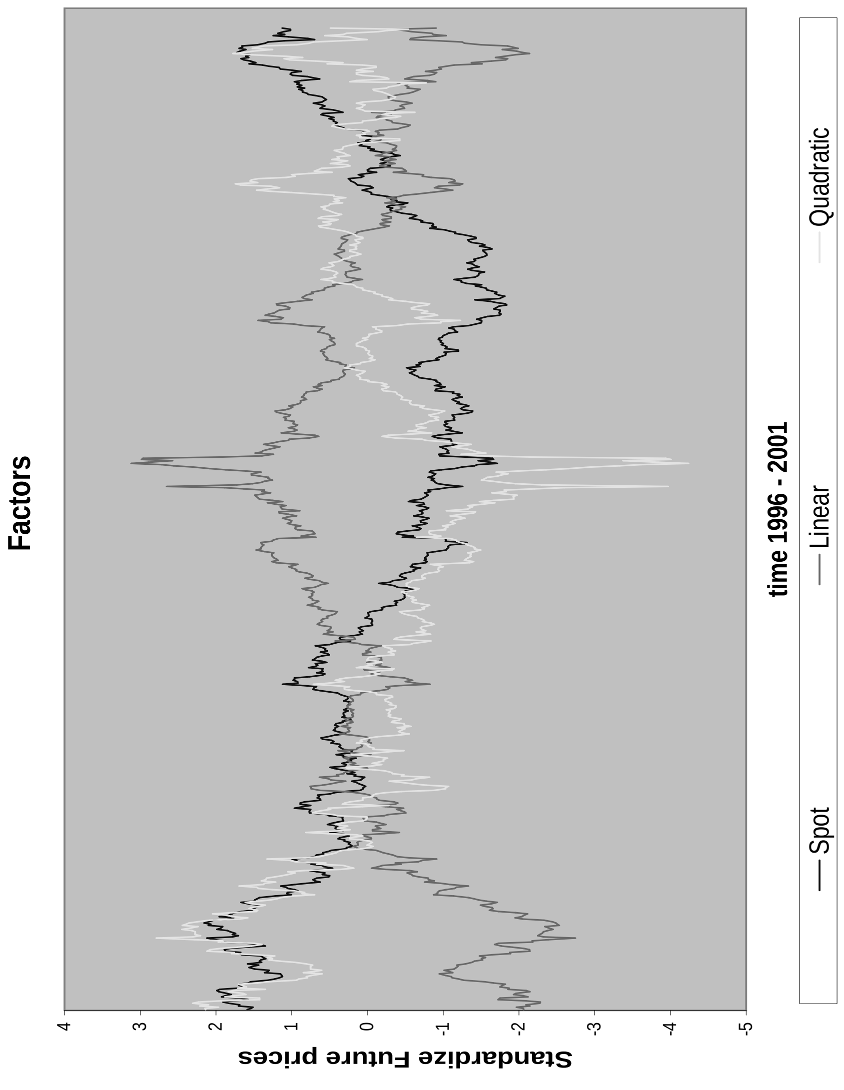

Figure 3: Time movements of model's factors, spot, slope and curvature.

References

[1] A survey of commodity risk management instruments Report by the UNCTAD secretariat 1998.

[2] Bjork, T. 1998 Arbitrage theory in continuous time. Oxford University Press.

[3] Bjork, T. 2002 On the term structure of futures and forward prices. Working paper.

[4] Braker, K. and Smith, K. 1999 Detecting and modeling changing volatility in the cooper futures markets. Journal of futures Markets, 19, 79-100.

[5] Breeden, D.T., 1980, Consumption risks in futures markets, Journal of Finance, 35(2): 503-520, May.

[6] Brennan, M.J., 1958, The supply of storage, American Economic Review, 48: 50-72, March.

[7] Brennan, M.J. Schwartz, E. S. 1985 Evaluating natural resource investment. Journal of Business, Vol. 58, 135-157.

[8] Brennan, M. J., 1991, The price of convenience and the evaluation of commodity contingent claims. Eds.: Stochastic Models and Option Values, North Holland.

[9] Cortazar, G. and E. Schwartz, 1994, The valuation of commodity-contingent claims, The Journal of Derivatives, 1(4): 27-39.

[10] Cox, J. Ingersoll, J. Ross, S. 1981 The relation between forward prices and futures prices. Journal of Financial Economics 9, 321-346.

[11] Duffie, D. 2001 Dynamic Asset Pricing Theory 3nd edition. Princeton University Press.

[12] Duffie, D. and R. Kan (1996) A yield factor model of interest rates. Mathematical finance 6, 379-406.

[13] Dusak, K. 1973 Futures trading and investor returns: An investigation of commodity market risk premiums. Journal of Political Economy, 81 (6) 1387-1406.

[14] Fama, E. and French, K. 1987 Commodity futures prices: Some evidence on forecast power, premiums and the theory of storage. Journal of Business, January.

[15] Filipovic, D., 2002, Separable Term Structure and the Maxima Degree problem, Mathematical Finance 12, 4 341-349 .

[16] Garbade, K., 1996, Fixed Income Analysis, Cambridge, MA: MIT Press.

[17] Garbade, K. Silver, W. 1983 Cash settlement of futures contracts: An economic analysis. Journal of Futures Markets, 1983, v3(4) 451-472.

[18] Gibson, R. and E. S. Schwartz, 1990, Stochastic convenience yield and the pricing of oil contingent claims, Journal of Finance, 45(3): 959-976.

[19] Harrison, J. M. Kreps, D. 1979 Martingales and arbitrage in multiperiod securities markets. Journal of Economic theory. 20, 381-408.

[20] Harrison, J. M. Pliska, S. 1981 Martingales and stochastic integrals in the theory of continuous trading. Stochastic Process and Their Applications 11, 215-260.

[21] Harvey, A. 1989, Forecasting structural time series models and the kalman filter. Cambridge University Press.

[22] Hazuka, T.B., 1984, Futures markets: Consumption betas and backwardation in commodity markets, Journal of Finance, 39(3): 647-655, July.

[23] Heath, D. Jarrow, R. Morton, A. 1992 Bond pricing and the term structure of interest rates: A new methodology. Econometrica 60, 77-105.

[24] Hull, J. 2000, Options, Futures and Other Derivatives. New Jersey: Prentice Hall.

[25] Karatzas, I. Shreve, S. 1988 Brownian Motion and stochastic calculus. Springer.

[26] Levin, A. 2004. Mean reverting and co-integrated futures curve models. Working Paper

[27] Manoliu, M. Tompaidis, S. (2002) Energy Futures prices: Term structure model with Kalman filter estimation, Applied mathematical finance 9, 21-43.

[28] Miltersen, K. Schwartz, S. 1998, Pricing of options on commodity futures with stochastic term structures of convenience yield and interest rate, Journal of Financial and Quantitative Analysis, 33(3): 33-59.

[29] Page, S.; Hewitt A. 2001 World commodity prices: Still a problem for developing countries?. Overseas Development Institute.

[30] Pilipovic, D. 1998 Energy Risk. New York: McGraw-Hill.

[31] Ross, S.A., 1995, Hedging long-run commitments: Exercises in incomplete market pricing, Preliminary draft.

[32] Schwartz, E. S., 1997, The stochastic behavior of commodity prices: Implications for valuation and hedging, Journal ofFinance, 52(3): 923-973.

[33] Schwartz, E., 1998, Valuing long-term commodity assets, Financial Management, 27(1): 57-66.

[34] Schwartz, E. and J. Smith, 2000, Short-term variations and long-term dynamics in commodity prices, Management Science,46(7): 893-911.

[35] Urich, T., 2000, Modes of fluctuation in metal futures prices, The Journal of Futures Markets 20(3): 219-241.

[36] Sterling, Y. 2002, Valuation of Commodity derivatives in a new Multi-Factor model Review of Derivatives Research 5, 251-271.