Preliminary Draft

A Multicommodity Model of Futures Prices: Using Futures Prices of One Commodity to Estimate the Stochastic Process of Another

Gonzalo Cortazar
Carlos Milla
Felipe Severino

Ingeniería Industrial y de Sistemas
Pontificia Universidad Católica de Chile

January, 2007

Abstract

This article proposes a multicommodity model of futures prices to explain the stochastic behavior of more than one commodity. Jointly modeling more than one commodity has the advantage of being able to use long-maturity futures prices of one commodity to estimate futures prices for another commodity which only has short-maturity contracts.

The model considers that commodity prices have a set of common factors that explain the correlation among them, in addition to some commodity-specific factors. A procedure for choosing the number of common and commodity-specific un-observable-Gaussian factors is presented. Also, it is shown how commodities with and without seasonality may be modeled together and how to estimate the multicommodity model using a Kalman Filter.

Two empirical implementations of the proposed multicommodity model are presented: First a model for the WTI and the Brent oil contracts, with the former commodity having much longer maturity contracts than the latter, and second a model for the WTI and the Unleaded Gasoline in which the second commodity not only has shorter maturities, but also a strong seasonality.

Results for both model implementations show strong improvements over the traditional individual-commodity models, with much lower out-of-sample errors and better volatility estimates, even when using fewer factors.

I. Introduction.

There is an evolving literature on how to model the stochastic behavior of commodity futures prices. The relevance of these models is at least two-fold. First they are used to estimate prices for contracts for which there are no market prices, and second they provide an estimation of the volatility term structure required to value option-like derivatives or to estimate risk exposures.

Commodity models have been evolving in several aspects. First, the number of risk factors has been increasing from early one-factor models to two, three and four factor models, with considerable gain in flexibility and adjustment to different term structure behaviors (Brennan and Schwartz (1985), Gibson and Schwartz (1990), Cortazar and Schwartz (1994), Cortazar and Naranjo (2006)).

A second aspect in which models have been evolving is in how the drift and the factors are modeled. Initial models considered simple geometric Brownian motions while more advanced models included different specifications factor dynamics including seasonality and mean reversion (Laughton and Jacoby (1993), Schwartz (1997), Dai and Singleton (2000), Manoliu and Tompaidis (2002), Sorensen (2002)).

A third dimension in which models have been developing is on the estimation procedures, including simple cross section model calibration, traditional Kalman Filtering with complete data panels to Extended Kalman procedures with time-dependent number of daily observations which requires no data aggregation and makes a better use of all available data (Cortazar and Schwartz (2003), Sorensen (2002), Cortazar and Naranjo (2006))

A fourth dimension in which models differ is in the volatility specification. Most models consider a constant volatility specification while some propose a USV specification (Trolle and Schwartz (2006)) which seems to better fit volatility structures at the expense of some loss on term-structure fitting.

In this paper we propose a new dimension for model evolution which considers extending individual-commodity models into a multicommodity setting. The basic intuition is that for commodities with highly correlated returns, price variations on contracts for one commodity should be useful information for the other. This has the advantage of being able to use information on the behavior of long-maturity futures prices available for one commodity to estimate futures prices for another commodity which only has short-maturity contracts. Another advantage of using these multicommodity models arises when the spread between two commodities is of interest. By jointly modeling both commodities more stable spreads should be obtained compared to the alternative of individually modeling both commodities and subtracting one commodity estimate from the other one.

The proposed multicommodity model considers that commodity prices have a set of common factors that explain the correlation among them, in addition to some commodity-specific factors. A procedure for choosing the number of common and commodity-specific un-observable-Gaussian factors is presented. Also, it is shown how commodities with and without seasonality may be modeled together and how to estimate the multicommodity model using a Kalman Filter.

Two empirical implementations of the proposed multicommodity model are presented: First a model for the WTI and the Brent oil contracts, with the former commodity having much longer maturity contracts than the latter, and second a model for the WTI and the Unleaded Gasoline in which the second commodity not only has shorter maturities, but also a strong seasonality.

Results for both model implementations show strong improvements over the traditional individual-commodity models, with much lower out-of-sample errors and better volatility estimates, even when using fewer factors.

The paper is organized as follows. In Section 2 the general multicommodity model, with and without seasonality, is presented. Section 3 shows how to estimate this model using the Kalman Filter. Section 4 explains a model selection procedure for choosing the number of common and of commodity-specific factors. Section 5 shows the results of implementing the model for two pair of commodities: WTI-Brent and WTI-Unleaded Gasoline. The last section concludes.

II. The Multicommodity Model

II.1. *Model Definition*

The proposed multicommodity model is based on the canonical representation of Dai and Singleton (2000) for interest rates and represents an extension of Cortazar and Naranjo (2006) for more than one commodity.

First, an m-commodity $(p, k_1, \ldots, k_m)$ model is described, in which all commodities share p common factors and have k_i commodity-***i*** specific factors. Later the model is extended to include seasonality.

Let

$$\begin{bmatrix} Y_1 \\ \vdots \\ Y_i \\ \vdots \\ Y_m \end{bmatrix}_t = \log \begin{bmatrix} S_1 \\ \vdots \\ S_i \\ \vdots \\ S_m \end{bmatrix}_t \qquad i = 1, \ldots, m \quad \text{commodities} \tag{2.1}$$

Y_i is described as follow:

$$Y_{it} = \tilde{h}_i\,'\,X_{it} + c_{it}$$

Where $\tilde{h}_i$ is a vector of dimension $n \times 1$, c_{it} is a time-dependent function and X_{it} is a vector of n state variables which follows a multivariate Orstein-Uhlembeck (O-U) stochastic process.

Then,

$$\begin{bmatrix} Y_1 \\ \vdots \\ Y_i \\ \vdots \\ Y_m \end{bmatrix}_t = \begin{bmatrix} \tilde{h}_1 \\ \vdots \\ \tilde{h}_i \\ \vdots \\ \tilde{h}_m \end{bmatrix} X_t + C_t = \mathbf{h} \cdot X_t + C_t \tag{2.2}$$

Where $\mathbf{h}$ is a $m \times n$ matrix, X_t is a state variable vector $n \times 1$ and C_t is a time-dependent matrix $m \times 1$.

The dynamics of the state variables is defined as:

$$dX_t = (-\tilde{\mathbf{A}} X_t + \tilde{\mathbf{b}})dt + \tilde{\mathbf{\Sigma}} d\tilde{\mathbf{w}}_t \tag{2.3}$$

Where $\tilde{\mathbf{A}}$ is a $n \times n$ semi-positive matrix, $\tilde{\mathbf{b}}$ is a constant vector, $\tilde{\mathbf{\Sigma}}$ is a $n \times n$ matrix and $d\tilde{\mathbf{w}}_t$ is a vector of ***n*** correlated Brownian motion increments such that

$$(d\tilde{\mathbf{w}}_t)(d\tilde{\mathbf{w}}_t)' = \tilde{\mathbf{\Theta}} dt . \tag{2.4}$$

Assuming that X_t follows a non-stationary process and applying a linear transformation

$$\mathbf{T}(X_t) = \boldsymbol{\varphi} + \mathbf{L} X_t , \tag{2.5}$$

the canonical multicommodity model is defined as:

$$\begin{bmatrix} Y_1 \\ \vdots \\ Y_i \\ \vdots \\ Y_m \end{bmatrix}_t = \begin{bmatrix} \log S_1 \\ \vdots \\ \log S_i \\ \vdots \\ \log S_m \end{bmatrix}_t = \begin{bmatrix} \overbrace{1 \quad \ldots \quad 1}^{p} & \overbrace{1 \quad \ldots \quad 1}^{k_1} & \ldots & \overbrace{0 \quad \ldots \quad 0}^{k_i} & \ldots & \overbrace{0 \quad \ldots \quad 0}^{k_m} \\ \vdots \qquad\quad \vdots & \vdots \qquad\quad \vdots & & \vdots \qquad\quad \vdots & & \vdots \qquad\quad \vdots \\ \delta_{i1} \quad \ldots \quad \delta_{ip} & 0 \quad \ldots \quad 0 & \ldots & 1 \quad \ldots \quad 1 & \ldots & 0 \quad \ldots \quad 0 \\ \vdots \qquad\quad \vdots & \vdots \qquad\quad \vdots & & \vdots \qquad\quad \vdots & & \vdots \qquad\quad \vdots \\ \delta_{m1} \quad \ldots \quad \delta_{mp} & 0 \quad \ldots \quad 0 & \ldots & 0 \quad \ldots \quad 0 & \ldots & 1 \quad \ldots \quad 1 \end{bmatrix} X_t + \begin{bmatrix} 1 \\ \vdots \\ \delta_{i1} \\ \vdots \\ \delta_{m1} \end{bmatrix} \mu_t$$

(2.6)

Where p is the number of common state variables, k_i in the number of specific factors for commodity-***i***, $n = p + \sum_{i=1}^{m} k_i$ is the total number of state variables, X_t is a $n \times 1$ vector of state variables, μ is the long-term growth rate and δ_{ij} is the weight of state variable j for commodity ***i***.

The dynamics of the vector of state variables X_t is:

$$dX_t = (-KX_t)dt + \mathbf{\Sigma} d\mathbf{w}_t \qquad (2.7)$$

Where

$$K = \begin{pmatrix} 0 & 0 & \cdots & 0 \\ 0 & \kappa_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \kappa_n \end{pmatrix} \qquad \mathbf{\Sigma} = \begin{pmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_n \end{pmatrix}$$

$$\mathbf{\Theta} dt = d\mathbf{w}_t d\mathbf{w}_t{}' \qquad \mathbf{\Theta} = \begin{pmatrix} 1 & \rho_{12} & \cdots & \rho_{1n} \\ \rho_{21} & 1 & \cdots & \rho_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{n1} & \rho_{n2} & \cdots & 1 \end{pmatrix} \tag{2.8}$$

This is a non-stationary model for the log spot price. By assuming a constant risk premium λ, the risk adjusted process for the vector of state variables is:

$$dX_t = -(\lambda + KX_t)dt + \mathbf{\Sigma} d\mathbf{w}_t^* \tag{2.9}$$

Where λ is a $n \times 1$ vector of real constants.

II.2. *Futures Prices With and Without Seasonality*

One of the good properties of this model is that it has an analytic expression for the futures prices. Following Cox et al (1981), the price at time *t* of a futures contract with maturity *T* must be equal to the expected spot price for *T* under the risk-neutral measure:

$$F(S_t, t, T) = \mathrm{E}_t^{\mathrm{Q}}(S_T) \tag{2.10}$$

Thus

$$F(X_t, t, T)_i = \exp\left(\delta_{i1} X_1(t) + \sum_{j=2}^{n} \delta_{ij}\, \mathrm{e}^{-\kappa_j (T-t)} X_j(t) + \delta_{i1}\mu t + (\delta_{i1}\mu - \delta_{i1}\lambda_1 + \frac{1}{2}\delta_{i1}{}^2 \sigma_1^2)(T-t) \right.$$
$$\left. - \sum_{j=2}^{n} \delta_{ij} \frac{1 - \mathrm{e}^{-\kappa_j (T-t)}}{\kappa_j} \lambda_j + \frac{1}{2} \sum_{j \cdot l \neq 1} \delta_{ij} \delta_{il} \sigma_j \sigma_l \rho_{jl} \frac{1 - e^{-(\kappa_j + \kappa_l)(T-t)}}{\kappa_j + \kappa_l} \right) \tag{2.11}$$

There are several ways to include seasonality in the model. Following Manoliu and Tompadis (2002) we add a deterministic function $P(T)$ to the futures expression, thus

$$F(X_t,t,T)_i = P(T)\exp\left(\delta_{i1}X_1(t)+\sum_{j=2}^{n}\delta_{ij}\,\mathrm{e}^{-\kappa_j(T-t)}\,X_j(t)+\delta_{i1}\mu t+(\delta_{i1}\mu-\delta_{i1}\lambda_1+\frac{1}{2}\delta_{i1}{}^2\sigma_1^2)(T-t)\right.$$
$$\left.-\sum_{j=2}^{n}\delta_{ij}\frac{1-\mathrm{e}^{-\kappa_j(T-t)}}{\kappa_j}\lambda_j+\frac{1}{2}\sum_{j\cdot l\neq 1}\delta_{ij}\delta_{il}\sigma_j\sigma_l\rho_{jl}\frac{1-e^{-(\kappa_j+\kappa_l)(T-t)}}{\kappa_j+\kappa_l}\right) \tag{2.12}$$

With $P(T)$ being a periodic step-function, such that if T belongs to month m, then

$P(T) = S_m$

$$\prod_{m=1}^{12} s_m = 1 \tag{2.13}$$

In any case, the volatility of futures returns is:

$$\sigma_{F_i}^2(t,T)=\delta_{i1}^2\sigma_1^2+\sum_{j,l\neq 1}\delta_{ij}\delta_{il}\sigma_j\sigma_l\rho_{jl}e^{-(\kappa_j+\kappa_l)(T-t)} \tag{2.14}$$

III. Multicommodity Estimation Using the Kalman Filter

III.1. *Kalman Filter and Incomplete Data*

To adequately benefit from using the information on the prices of one commodity to calibrate the stochastic behaviour [of t]he prices of another, a joint estimation of both processes should be performed. In this section it is shown how the Kalman Filter can be used to estimate a model, even if data is incomplete and for some days there are no prices for a commodity (Cortazar and Naranjo (2006)).

Let the following measurement equation relate a vector of observable variables $\mathbf{z}_t$ with a vector of state variables, $\mathbf{x}_t$

$$\mathbf{z}_t = \mathbf{H}_t \mathbf{x}_t + \mathbf{d}_t + \mathbf{v}_t \qquad \mathbf{v}_t \sim N(\mathbf{0}, \mathbf{R}_t) \tag{3.1}$$

Where $\mathbf{H}_t$ is a $u_t \times n$ matrix, $\mathbf{d}_t$ is a $u_t \times 1$ vector and $\mathbf{v}_t$ is an $u_t \times 1$ of uncorrelated Gaussian disturbances with mean 0 and covariance matrix $\mathbf{R}_t$.

The transition equation describes the stochastic process followed by the state variables

$$\mathbf{x}_t = \mathbf{A}_t \mathbf{x}_{t-\Delta t} + \mathbf{c}_t + \mathbf{w}_t \qquad \mathbf{w}_t \sim N(0, \mathbf{Q}_t) \tag{3.2}$$

Where $\mathbf{A}_t$ is $n \times n$, $\mathbf{c}_t$ $n \times 1$ vector and $\mathbf{w}_t$ is a vector of uncorrelated Gaussian disturbances. The variance-covariance matrix of the estimation error, $\mathbf{P}_t$, is:

$$\mathbf{P}_t = \mathbf{E}_t (\mathbf{x}_t - \hat{\mathbf{x}}_t)(\mathbf{x}_t - \hat{\mathbf{x}}_t)' \tag{3.3}$$

So for a given $\hat{\mathbf{x}}_{t-1}$ and $\mathbf{P}_{t-1}$ the estimated state variables and variance-covariance error estimation matrix for t will be.

$$\hat{\mathbf{x}}_{t/t-1} = \mathbf{A}_t \hat{\mathbf{x}}_{t-1} + \mathbf{c}_t \tag{3.4}$$

The following one-step-ahead prediction of the observed variables can be obtained:

$$\hat{\mathbf{z}}_{t|t-1} = \mathbf{H}_t \hat{\mathbf{x}}_{t|t-1} + \mathbf{d}_t \tag{3.5}$$

When there is new information $\mathbf{z}_t$, a new estimation can be obtained

$$\hat{\mathbf{x}}_t = \hat{\mathbf{x}}_{t|t-1} + \mathbf{P}_{t|t-1} \mathbf{H}_t' \mathbf{F}_t^{-1} \left(\mathbf{z}_t - \hat{\mathbf{z}}_{t|t-1} \right) \tag{3.6}$$

$$\mathbf{P}_t = \mathbf{P}_{t|t-1} + \mathbf{P}_{t|t-1} \mathbf{H}_t' \mathbf{F}_t^{-1} \mathbf{H}_t \mathbf{P}_{t|t-1} \tag{3.7}$$

with

$$\mathbf{F}_t = \mathbf{H}_t \mathbf{P}_{t|t-1} \mathbf{H}_t' + \mathbf{R}_t \tag{3.8}$$

As Cortazar and Naranjo (2006) point out, the above optimal estimates can be obtained even if the number of observations varies with time. The Kalman Filter allows the dimension u_t of vectors $\mathbf{z}_t$, $\mathbf{d}_t$, $\mathbf{v}_t$ and of matrices $\mathbf{H}_t$ and $\mathbf{R}_t$, to be a function of time.

This missing observation problem may be increasingly important in multicommodity settings where not for all days and maturities there may be prices for all commodities in the model.

Model parameter estimations, $\hat{\boldsymbol{\psi}}$, are obtained maximizing the log-likelihood function of innovations.

$$\log L(\boldsymbol{\psi}) = -\frac{1}{2}\sum_t \log\left|\mathbf{F}_t\right| - \frac{1}{2}\sum_t \left(\mathbf{z}_t - \mathbf{z}_t^-\right)'\mathbf{F}_t^{-1}\left(\mathbf{z}_t - \mathbf{z}_t^-\right) \quad (3.9)$$

III.2. *Kalman Filter of the Multicommodity Model*

In this section it is shown how to make a state space representation of the multicommodity model to solve it using the Kalman Filter.

First all commodities from 1 to m are stacked in the following way:

$$\mathbf{z}_t = \begin{pmatrix} \log F^1{}_{1,t} \\ \vdots \\ \log F^1{}_{u_1,t} \\ \vdots \\ \log F^i{}_{1,t} \\ \vdots \\ \log F^i{}_{u_i,t} \\ \vdots \\ \log F^m{}_{1,t} \\ \vdots \\ \log F^m{}_{u_m,t} \end{pmatrix}$$

$$\textit{with} \quad (3.10)$$

$$u_t = \sum_{i=1}^{m} u_{i,t}$$

Where $F^{i}_{u_i,t}$ is the price of a futures contract of commodity i at time t with maturity corresponding to the u_i,t position.

Then, matrix $\mathbf{H}_t$ can be defined as:

$$\mathbf{H}_t = \overbrace{\qquad\qquad\qquad}^{p}\overbrace{\qquad\qquad\qquad}^{k_1}\cdots\overbrace{\qquad\qquad\qquad}^{k_i}\cdots\overbrace{\qquad\qquad\qquad}^{k_m}$$

$$\mathbf{H}_t = \begin{bmatrix}
1 & e^{-\kappa_2\tau^1_{1,t}} & \dots & e^{-\kappa_p\tau^1_{1,t}} & e^{-\kappa_{p+1}\tau^1_{1,t}} & \dots & e^{-\kappa_{p+k_1}\tau^1_{1,t}} & \dots & 0 & \dots & 0 & \dots & 0 & \dots & 0 \\
\vdots & \vdots & & \vdots & \vdots & & \vdots & & \vdots & & \vdots & & \vdots & & \vdots \\
1 & e^{-\kappa_2\tau^1_{u1,t}} & \cdots & e^{-\kappa_p\tau^1_{u1,t}} & e^{-\kappa_{p+1}\tau^1_{u1,t}} & \dots & e^{-\kappa_{p+k_1}\tau^1_{u1,t}} & \dots & 0 & \dots & 0 & \dots & 0 & & 0 \\
\vdots & \vdots & & \vdots & \vdots & & \vdots & & \vdots & & \vdots & & \vdots & & \vdots \\
\delta_{i1} & \delta_{i2}e^{-\kappa_2\tau^i_{1,t}} & \dots & \delta_{ip}e^{-\kappa_p\tau^i_{1,t}} & 0 & \dots & 0 & \dots & e^{-\kappa_{p+\sum_{j<i}k_j+1}\tau^i_{1,t}} & \dots & e^{-\kappa_{p+\sum_{j=1}^{i}k_j}\tau^i_{1,t}} & \dots & 0 & \dots & 0 \\
\vdots & \vdots & & \vdots & \vdots & & \vdots & & \vdots & & \vdots & & \vdots & & \vdots \\
\delta_{i1} & \delta_{i2}e^{-\kappa_2\tau^i_{ui,t}} & \dots & \delta_{ip}e^{-\kappa_p\tau^i_{ui,t}} & 0 & \dots & 0 & \dots & e^{-\kappa_{p+\sum_{j<i}k_j+1}\tau^i_{ui,t}} & \dots & e^{-\kappa_{p+\sum_{j=1}^{i}k_j}\tau^i_{ui,t}} & \dots & 0 & \dots & 0 \\
\vdots & \vdots & & \vdots & \vdots & & \vdots & & \vdots & & \vdots & & \vdots & & \vdots \\
\delta_{m1} & \delta_{m2}e^{-\kappa_2\tau^m_{1t}} & \dots & \delta_{mp}e^{-\kappa_p\tau^m_{1t}} & 0 & \dots & 0 & \dots & 0 & \dots & 0 & \dots & e^{-\kappa_{p+\sum_{j<m}k_m+1}\tau^m_{1,t}} & \dots & e^{-\kappa_{p+\sum_{j=1}^{m}k_j}\tau^m_{1,t}} \\
\vdots & \vdots & & \vdots & \vdots & & \vdots & & \vdots & & \vdots & & \vdots & & \vdots \\
\delta_{m1} & \delta_{m2}e^{-\kappa_2\tau^m_{u_m,t}} & \dots & \delta_{mp}e^{-\kappa_p\tau^m_{u_m,t}} & 0 & \dots & 0 & \dots & 0 & \dots & 0 & \dots & e^{-\kappa_{p+\sum_{j<m}k_m+1}\tau^m_{u_m,t}} & \dots & e^{-\kappa_{p+\sum_{j=1}^{m}k_j}\tau^m_{u_m,t}}
\end{bmatrix} \quad (3.11)$$

And

$$\mathbf{d}_t = \begin{pmatrix} v^1{}_{1,t} \\ \vdots \\ v^1{}_{u_1,t} \\ \vdots \\ v^i{}_{1,t} \\ \vdots \\ v^i{}_{u_i,t} \\ \vdots \\ v^m{}_{1,t} \\ \vdots \\ v^m{}_{u_m,t} \end{pmatrix} \qquad (3.12)$$

where for commodity *i*

$$v^i{}_{j,t} = \delta_{i1}\mu t + (\delta_{i1}\mu - \delta_{i1}\lambda_1 + \frac{1}{2}\delta^2{}_{i1}\sigma_1^2)\tau_{j,t} - \sum_{l=2}^{n}\delta_{il}\frac{1-\mathrm{e}^{-\kappa_l\tau_{j,t}}}{\kappa_l}\lambda_l + \frac{1}{2}\sum_{(l,q)\neq(1,1)}^{n}\delta_{il}\delta_{iq}\sigma_l\sigma_q\rho_{lq}\frac{1-e^{-(\kappa_l+\kappa_q)\tau_{j,t}}}{\kappa_l+\kappa_q} \quad (3.13)$$

With this state space representation, the multicommodity model may be estimated using the Kalman Filter.

IV. The Model Selection Procedure

IV.1. *The Basic Idea*

Previous sections have shown how to define and estimate a given $(p, k_1, \ldots, k_m)$ model for m commodities which shares p common factors and has k_i commodity-***i*** specific factors. In this section a procedure is proposed for choosing p and k_i, without resorting to estimating all possible models and comparing their performance which may be very expensive in computation time and not practical for real world applications.

In individual-commodity models it has long been recognized that using Principal Components methods is useful to select the number of factors that should be used to explain a given data variance percentage. This is the case in Litterman and Scheinkman (1991) for bonds and in Cortazar and Schwartz (1994) for commodity futures. Also, these individual principal components have been compared to analyze the behavior of correlated commodities (Tolmasky and Hindanov(2002)).

Multicommodity models have the difficulty of having to relate different variance-covariance matrices, finding common and commodity-specific factors. A similar problem can be found in the study of the evolution of biological species with the goal of finding their common components (Krzanowski (1979), Flury (1988) and Philips and Arnold (1999)).

To perform the comparison between biological species represented by their variance-covariance matrices of dimension $r \times r$, two procedures have been proposed: Common Principal Components (CPC), in which all components are common to the different matrices, and Partial Common Principal Components (CPC(*p*), in which only *p* of the components are common.

Principal components are based on an eigenvalue-eigenvector decomposition of variance-covariance matrices. The first procedure (CPC) searches for a set of principal components that simultaneously explains the variance of all matrices. In this case the eigenvectors are equal for all matrices, being the eigenvalues specific for each one. The second procedure (CPC(p)) extends this approach allowing p eigenvectors to be identical in all groups while the remaining $r - p$ eigenvectors are specific for each group.

The model selection procedure proposed considers aggregating all futures data into a fixed number (r) of maturity-classes. Second, estimating different factor representations (CPC(*p*)), computing the likelihood, eigenvalues and eigenvectors for each representation. Finally, in order to penalize representations with a high number of parameters, the model is chosen using the Schwarz Information Criteria (SIC).

The benefits of this model selection procedure is that it is computationally efficient and multiple model representations can be easily explored, before engaging on the much more demanding Kalman Filtering of the chosen model.

IV.2. *A More Detailed Description of CPC-CPC(p) and SIC*

The Common Principal Components approach assumes a level of similarity among m covariance matrices $\Psi_1, \ldots, \Psi_m$ of dimension $r \times r$, assuming that all Ψ_i are positive definite.

Then, m covariances matrices have common principal components if

$$\Psi_i = \beta \Lambda_i \beta' \quad i = 1,\ldots,m \tag{4.1}$$

Where β is an orthogonal $r \times r$ matrix and

$$\Lambda_i = diag(\lambda_{i1},\ldots,\lambda_{ir}) \tag{4.2}$$

The number of parameters is $r(r-1)/2$ for the orthogonal matrix β plus $m \cdot r$ for the diagonal matrices Λ_i.

Assuming that all CPCs are well defined, that is, for each $j \in \{1,\ldots,r\}$ there is a least one population i in which the characteristic root λ_{ij} is distinct.

Being the sample covariance matrices $S_i \sim W_r(n_i, \Psi_i / n_i)$, the joint likelihood function of $\Psi_1,\ldots,\Psi_m$ given $S_1,\ldots,S_m$ is

$$L(\Psi_1,\ldots,\Psi_m) = C \times \prod_{i=1}^{m} etr\left(-\frac{n_i}{2}\Psi_i^{-1}S_i\right)(\det \Psi_i)^{-n_i/2} \tag{4.3}$$

Where the factor C does not depend on the parameters. Maximizing (4.3) the estimation of common principal components analyzes the similarity of the different matrices (Flury (1988)).

However, if the two “species” are not actually the same, even though they do have some common factors, the CPC model may still be rejected. The Partial Principal Components Model (CPC(p)) model takes care of this problem by allowing p components to be identical for all m matrices while the remaining $r-p$ components are specific. Formally, the hypothesis of partial CPCs (of order p) is

$$\Psi_i = \beta^{(i)} \Lambda_i \beta^{(i)\prime} \quad i = 1,\ldots,m \tag{4.4}$$

Where

$$\Lambda_i = diag(\lambda_{i1},\dots,\lambda_{ir}) \tag{4.5}$$

and

$$\beta^{(i)} = \left(\beta_c, \beta_s^{(i)}\right) \tag{4.6}$$

All $\beta^{(i)}$ are orthogonal $r \times r$ matrices. Then β_c, of dimension $r \times p$, is common to all groups, while $\beta_s^{(i)}$, of dimension $r \times (r-p)$, is specific. By orthogonality $CPC(r-1)$ implies $CPC(r)$ which is the ordinary CPC model. We therefore restrict p to the range $1 \le p \le r-2$. This means that the partial CPC model requires a dimension r of at least 3.

In this case β_c and $\beta_s^{(i)}$ will be written as

$$\beta_c = \left(\beta_1,\dots,\beta_p\right) \tag{4.7}$$

And

$$\beta_s^{(i)} = \left(\beta_{p+1}^{(i)},\dots,\beta_r^{(i)}\right) \qquad i = 1,\dots,m \tag{4.8}$$

Just like in the CPC approach, we start with m independent sample matrices $S_i \sim W_r(n_i, \Psi_i / n_i)$. Assuming the $\mathrm{CPC(p)}$ model, maximizing the likelihood is equivalent to minimizing the function.

$$g(\beta_c, \beta_s^{(1)}, \ldots, \beta_s^{(m)}, \Lambda_1, \ldots, \Lambda_m) = \sum_{i=1}^{m} n_i \left[\sum_{j=1}^{r} \log \lambda_{ij} + \sum_{j=1}^{p} \frac{\beta_j^{'} S_i \beta_j}{\lambda_{ij}} + \sum_{j=p+1}^{r} \frac{\beta_j^{(i)'} S_i \beta_j^{(i)}}{\lambda_{ij}} \right] \quad (4.9)$$

To minimize this function under orthogonal constraints for all $\beta^{(i)}$:

$$\beta_h^{'} \beta_j = \begin{cases} 0, & if\ h \neq j \\ 1, & if\ h = j \end{cases} 1 \leq h, j \leq p$$

$$\beta_h^{(i)'} \beta_j^{(i)} = \begin{cases} 0, & if\ h \neq j \\ 1, & if\ h = j \end{cases} p \leq h, j \leq r,\ i = 1, \ldots, m$$

$$\beta_h^{'} \beta_j^{(i)} = 0, \qquad i = 1, \ldots, m\ \ 1 \leq h \leq p < j \leq r, \quad (4.10)$$

This is equivalent to minimizing the function

$$G = g - \sum_{h=1}^{p} \gamma_h (\beta_h^{'} \beta_h - 1) - 2 \sum_{1 \leq h < j \leq p} \gamma_{hj} \beta_h^{'} \beta_j$$
$$- \sum_{i=1}^{m} \left[\sum_{h=p+1}^{r} \gamma_h^{(i)} (\beta_h^{(i)'} \beta_h^{(i)} - 1) - 2 \sum_{p < h < j \leq r} \gamma_{hj}^{(i)} \beta_h^{(i)'} \beta_j^{(i)} + 2 \sum_{1 \leq h \leq p < j \leq r} \delta_{hj}^{(i)} \beta_h^{(i)'} \beta_j^{(i)} \right] \quad (4.11)$$

Where the $\gamma_h, \gamma_{hj}, \gamma_h^{(i)}, \gamma_{hj}^{(i)}$ and $\delta_{hj}^{(i)}$ are the $\left[p(p+1) + m(r-p)(r+p+1) \right] / 2$ Lagrange multipliers.

To avoid the shortcomings of multiple testing when comparing several competing models, we propose using the Schwarz Information Criteria (SIC) which penalizes models with many parameters.

$$SIC = -2 \log L + q \log(n) \quad (4.12)$$

Where L is the likelihood function, q is the number of parameters and n is the sample size.

Suppose the existence of c models, where $q_1 < q_2 < \ldots < q_c$ denote the number of parameters estimated, and $L_1 \leq L_2 \leq \ldots \leq L_c$ are the values of the likelihood function at the maxima. Then minimizing 4.12 is equivalent to choosing the *ith* model if:

$$SIC(i) = -2\log\frac{L_i}{L_c} + (q_i - q_1)\log(n) \qquad \forall i, \quad n_i = n \tag{4.13}$$

is a minimum.

V. Model Implementation and Results for the WTI-Brent and the WTI-Unleaded Gasoline models

V.1. *The Data*

Two empirical implementations of the proposed multicommodity model are presented. Model CLCB studies the WTI OIL (CL) and the BRENT OIL (CB) futures contracts, while Model CLHU analyzes the WTI (CL) and the Unleaded Gasoline (HU) contracts. In both models the WTI has much longer-maturity contracts than the other commodity. In addition, in Model CLHU, prices of the unleaded gasoline contract exhibit seasonality.

Each of the two models is tested using three different data sets. First, for the **in-sample testing**, a data set is used for parameter and state-variable estimation. Second, in what we call **traditional out-of-sample testing** a data set is used only for state-variable estimation, but without re-estimating model parameters. Finally in the **extreme out-of-sample testing** model prices are compared with those of a data set that is neither used for parameter nor for state variable estimation.

For <u>Model CLCB</u> the **in-sample data** consists of daily prices for CL and CB contracts traded at NYMEX and ICE, respectively, between 2001 and 2004. The longest-maturity CL-contract used is 7 years, while the longest-maturity CB-contract used is 2.5 years. The **traditional out-of-sample data** includes the same contracts traded daily between January 2005 and December 2006. Given that new long-maturity CB contracts (over 2.5 years) are issued starting on February 2005, we reserve this data to be used only for the **extreme out-of-sample testing** to compare these prices with model estimations that did not consider this information at all.

For <u>Model CLHU</u> the **in-sample data** consists of daily prices for CL and HU contracts traded at NYMEX between 2000 and 2004. The longest-maturity CL-contract used is 7 years, while the longest-maturity HU-contract is 1 year. The **traditional out-of-sample data** includes the same contracts traded daily between January 2005 and December 2006. Given that during 2006 the HU contracts are being been faced out (with the last contract maturing on December 2006), and a new reformulated gasoline contract (RB) was introduced, prices from July to December of 2006 for this new RB-contract (with maturities that extend to 2007) are used for the **extreme out-of-sample testing**.

V.2. ***Model Selection***

In this section the question of how many common factors to include in each of the multicommodity models is addressed.

Contracts for each of the three commodities (CL, CB and HU) in the in-sample data set are aggregated into 6 groups, (Litterman and Scheinkmann (1991), Cortazar and Schwartz (1994)).

Then, for each pair of commodities (CL-CB and CL-HU), Partial Common Principal Components are estimated, using different number of common factors (from cero to 5), and the Schwarz Information Criteria (SIC) number (which takes into consideration both the likelihood and the number of model parameters) is computed. The number of common factors chosen for each model will be the one that minimizes the SIC.

The following Tables show that for the CLCB Model, this procedure proposes 3 common factors and for the CLHU Model, 2 common factors.

Table: Model Selection for the CL-CB case.

Minimizing the SIC for different number of Partial Common Principal Components CPC(p), in which p is the number of common factors. Note that given that the data set is aggregated into 6 groups, CPC(5) = CPC(6)

	ChiSqr	Number of parameters (q_i)	(q_i-q_1)	SIC(i)
CPC(5)	82,67	27	0	82,67
CPC(4)	76,30	28	1	82,97
CPC(3)	38,13	30	3	**58,14**
CPC(2)	20,71	33	6	60,74
CPC(1)	11,08	37	10	77,80
CPC(0)		42	15	100,08
n	790			

Table: Model Selection for the CL-HU case.

Minimizing the SIC for different number of Partial Common Principal Components CPC(p), in which p is the number of common factors. Note that given that the data set is aggregated into 6 groups, CPC(5) = CPC(6)

	ChiSqr	Number of parameters (q_i)	(q_i-q_1)	SIC(i)
CPC(5)	301,93	27	0	301,93
CPC(4)	277,28	28	1	284,40
CPC(3)	69,74	30	3	91,10
CPC(2)	38,25	33	6	**80,96**
CPC(1)	17,15	37	10	88,33
CPC(0)		42	15	106,78
n	1234			

To allow for each model to have at least one commodity-specific factor and also for each commodity to have at least 3 factors to adequately represent its dynamics, we finally choose (3, 0, 1) factors, for the CLCB Model, and (2, 1, 1) factors for the CLHU Model, in which (p, k_1, k_2) defines a model with p common factors and k_i commodity-specific factors for commodity i.

V.3. *Model Estimation*

Using the Kalman Filter procedure with incomplete data panels described earlier, 2 multicommodity models (CLCB and CLHU) plus their 4 individual commodity models to be used as benchmarks, are calibrated.

The performance of Model CLCB, which is specified with (3, 0, 1) factors, will be compared with an individual CL model with 3 factors and an individual CB model with 4 factors.

The performance of Model CLHU, which is specified with (2, 1, 1) factors and 12 seasonality factors, will be compared with an individual CL model with 3 factors and an individual HU model with 3 factors and 12 seasonality factors.

The next Table shows the parameter estimates for all the models

Table

Parameter Estimates for Multicommodity and Individual-Commodity Models

	CLCB	CL	CB	CLHU	CL	HU
κ1	-	-	-	-	-	-
κ2	0,384	0,318	0,460	0,414	1,081	1,000
κ3	0,911	1,106	0,697	1,184	0,455	4,001
κ4	0,435		6,951	1,104		
σ1	0,196	0,424	0,252	0,217	0,239	0,333
σ2	0,178	0,688	0,885	0,122	0,662	0,398
σ3	0,336	0,977	1,000	0,325	0,286	0,362
σ4	0,080	-	0,121	0,355		
ρ21	- 0,389	- 0,902	- 0,567	- 0,408	0,424	- 0,604
ρ31	0,370	0,892	0,525	0,152	- 0,483	0,488
ρ41	- 0,192	-	- 0,055	0,010		
ρ32	- 0,589	- 0,948	- 0,965	- 0,026	- 0,760	- 0,603
ρ42	0,180	-	0,275	0,044		
ρ43	- 0,087	-	- 0,310	0,954		
λ1	0,022	0,079	0,051	0,018	0,037	0,006
λ2	0,046	0,054	- 0,068	- 0,010	0,179	0,004
λ3	- 0,003	0,137	0,213	0,064	0,074	0,007
λ4	0,017	-	- 0,179	- 0,170		
μ	0,002	0,001	0,000	0,000	0,007	0,003
ξ1	0,005	0,005	0,003	0,006	0,006	0,017
ξ2	0,008	-	-	0,022		
s1				0,963		0,969
s2				0,978		0,974
s3				1,021		1,021
s4				1,063		1,056
s5				1,063		1,051
s6				1,051		1,042
s7				1,035		1,030
s8				1,011		1,010
s9				0,971		0,975
s10				0,955		0,962
s11				0,949		0,958
s12				0,951		0,960
δ11	1,000	1,000	1,000	1,000	1,000	1,000
δ12	1,000	1,000	1,000	1,000	1,000	1,000
δ13	1,000	1,000	1,000	1,000	1,000	1,000
δ14	-	-	1,000	-		
δ21	1,000			0,998		
δ22	1,000			0,985		
δ23	1,000			-		
δ24	1,000			1,000		
log L	**182458**	**123558**	**71267**	**229490**	**185474**	**48174**

V.4. ***Results***

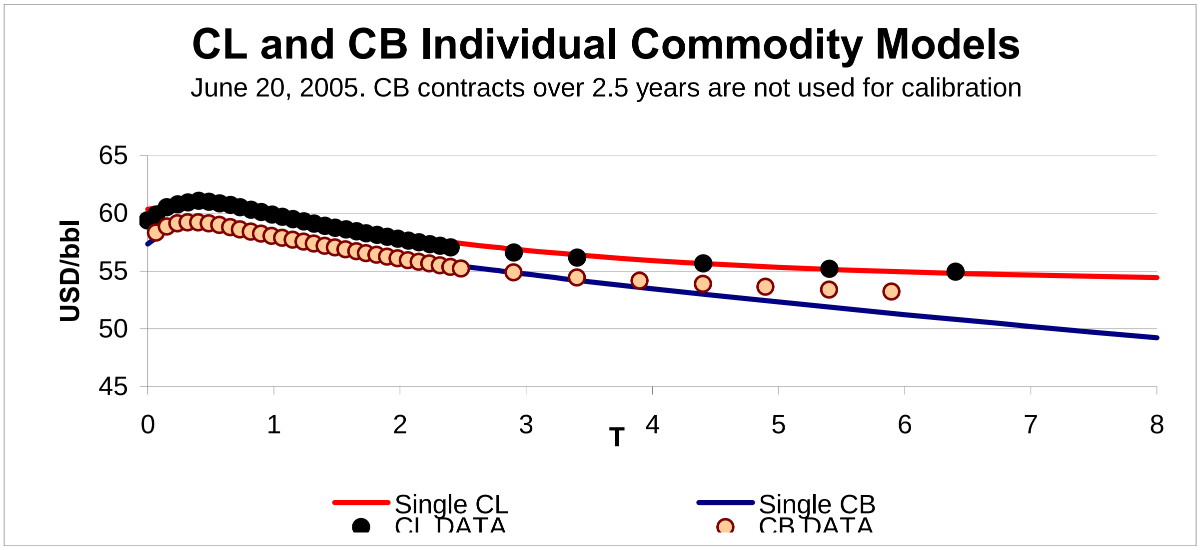
To analyze the performance of the multicommodity models we compare each of them with individual commodity models specified with same number of factors. For example the (3, 0, 1) CLCB Model is compared with the 3-factor CL model and the 4-factor CB Model. This is a conservative comparison in the sense that the multicommodity model has 4 different factors, while the individual models use altogether 7. Even with this somewhat unfair comparison the multicommodity model performs in many ways much better than the individual commodity models, as will be shown in what follows.

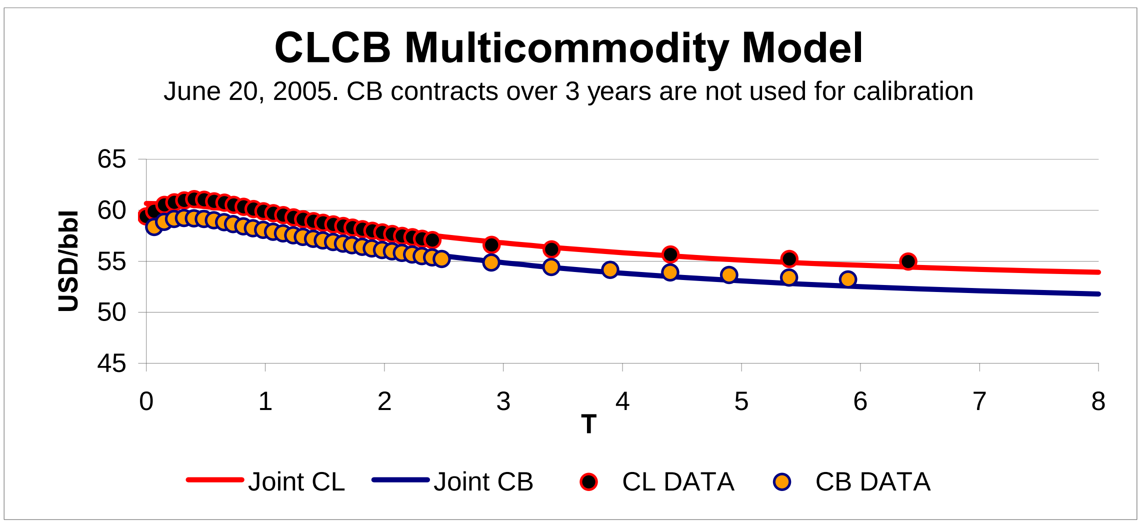
There are at least three measures of model performance that could be used to validate the use of a multicommodity, over an individual-commodity, model. First, Price Adjustment, second Volatility Adjustment, and third Spread Stability (between two commodities). As expected, our preliminary results show much more stable spreads on multicommodity models, but we do not report them in this version of the paper and restrict ourselves to the first two performance measures.

V.4.1. Price Adjustments

The following Figure shows individual model estimations for the CL and CB contracts on June 20, 2005. All CL and CB contracts with less than 2,5-year maturities are included in the traditional out-of-sample data and are very similar to model estimates. CB contracts with more than 2,5-year maturities belong to the extreme out-of-sample data set (not used for parameter or state-variable estimation) and exhibit strong differences with model estimates.

As can be seen in the next Figure, the multicommodity model is able to fit very well this extreme out-of-sample data set by using the correlation structure between both commodities to infer prices from commodities without transactions.



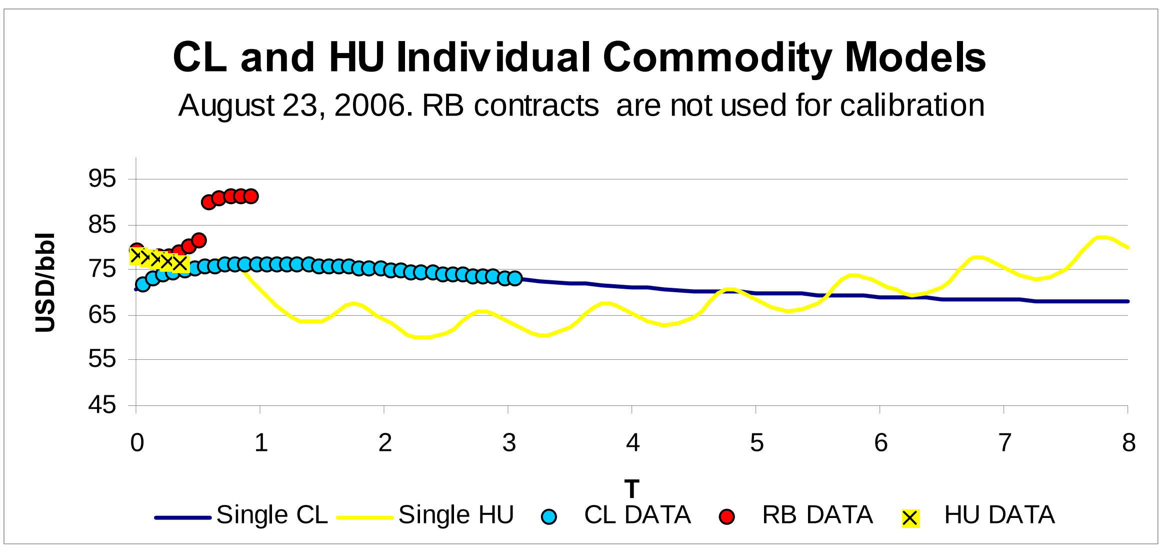


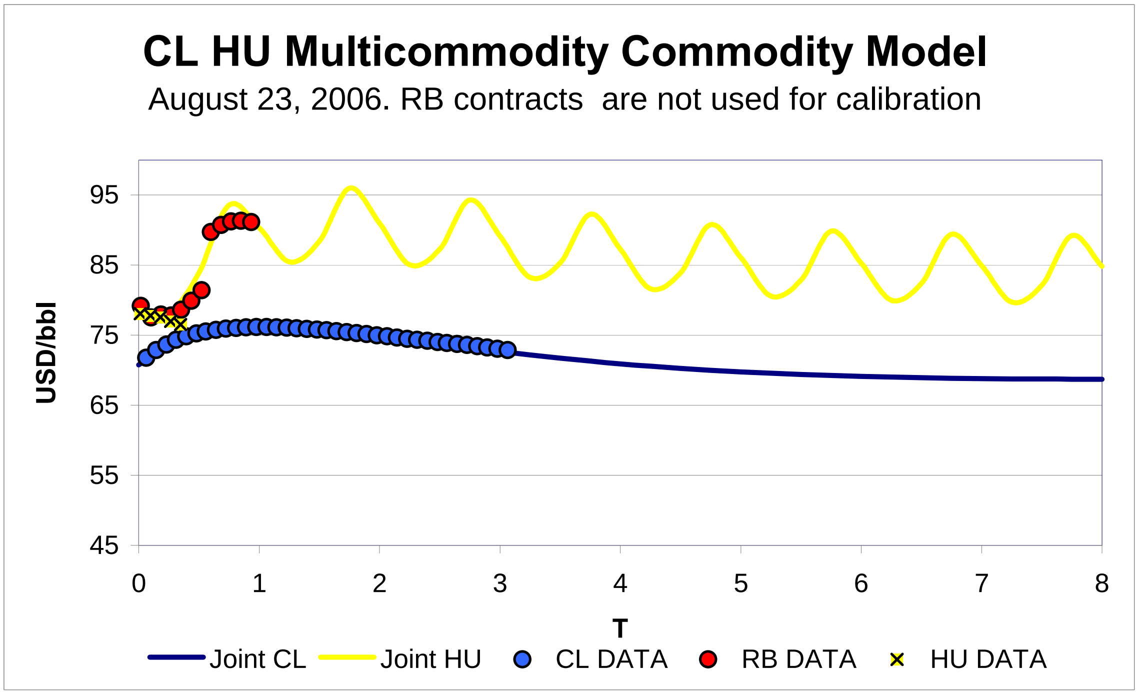
The next Table shows the in-sample and out-of-sample ME and RMSE errors for the CL and CB contracts, using the individual and the multicommodity models. For the CB contract the extreme out-of-sample testing using CB contracts of more than 2,5-year maturity contracts, is also reported.

	ME CL Indiv Mod	ME CL Multi Mod	RMSE CL Indiv	RMSE CL Multi
IN SAMPLE	-0,0011%	-0,0012%	0,48%	0,54%
OUT SAMPLE	-0,0012%	-0,0129%	0,33%	0,45%

	ME CB Indiv	ME CB Multi	RMSE CB Indiv	RMSE CB Multi
IN SAMPLE	0,0029%	-0,0080%	0,31%	0,80%
OUT SAMPLE	1,1668%	0,3625%	2,86%	1,06%
EXTREME OUT SAMPLE	4,2567%	1,5560%	5,15%	1,74%

The following Figures and Tables compare the performance of individual and multicommodity models for the CL and HU contracts.





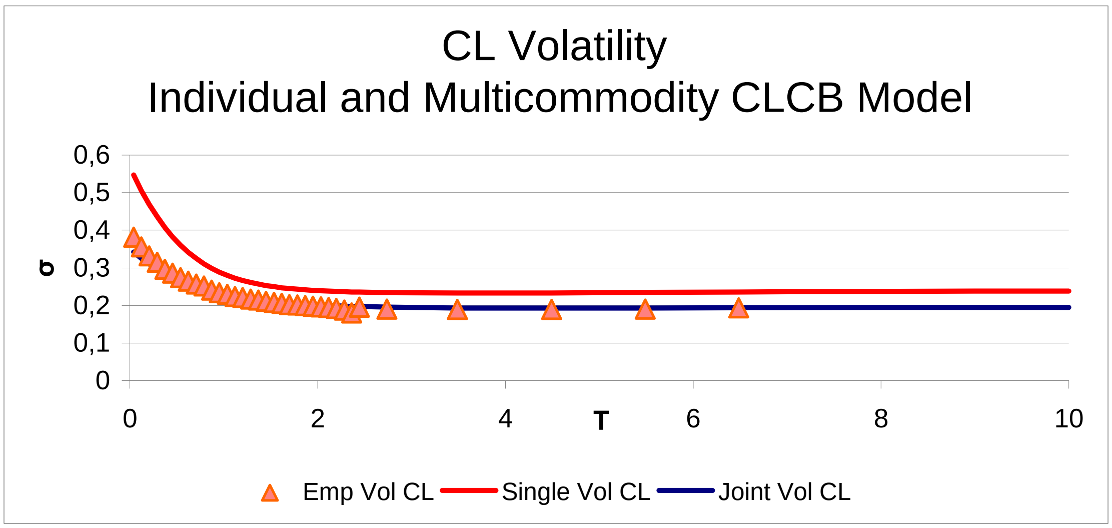
	ME CL Ind Mod	ME CL Multi Mod	RMSE CL Indiv	RMSE CL Multi
IN SAMPLE	-0,0083%	-0,0038%	0,54%	0,54%
OUT SAMPLE	0,0076%	0,0054%	0,38%	0,38%

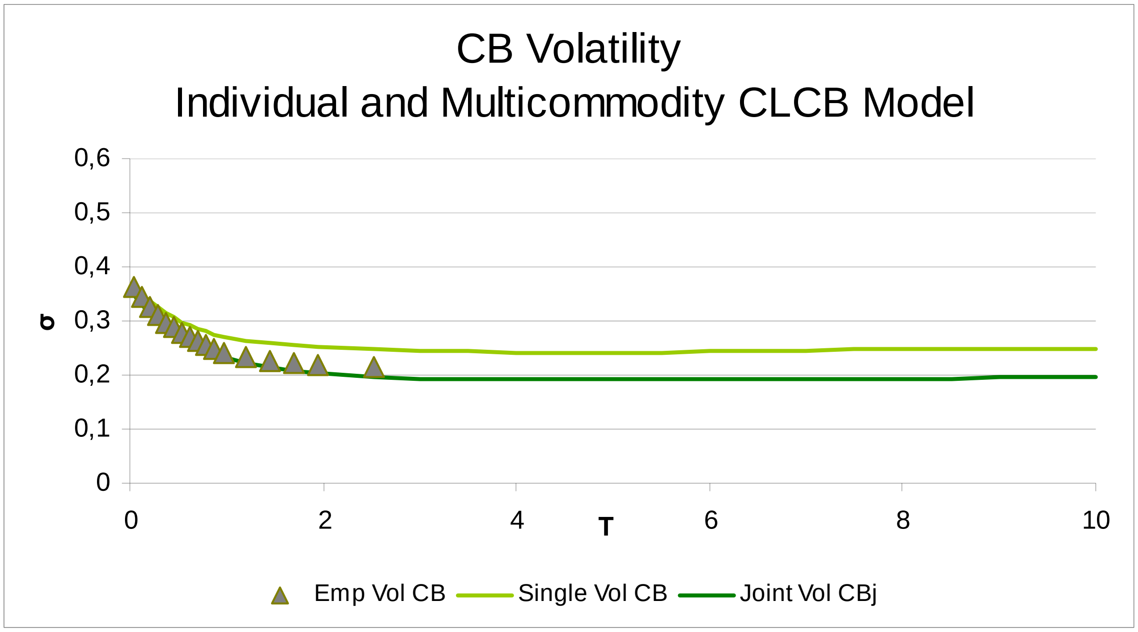
	ME HU Ind	ME HU Multi	RMSE HU Ind	RMSE HU Multi
IN SAMPLE	0,0082%	0,1913%	1,56%	1,87%
OUT SAMPLE	0,0225%	-0,6392%	0,81%	2,19%
EXTREME OUT SAMPLE	8,6271%	-0,3812%	10,38%	3,73%

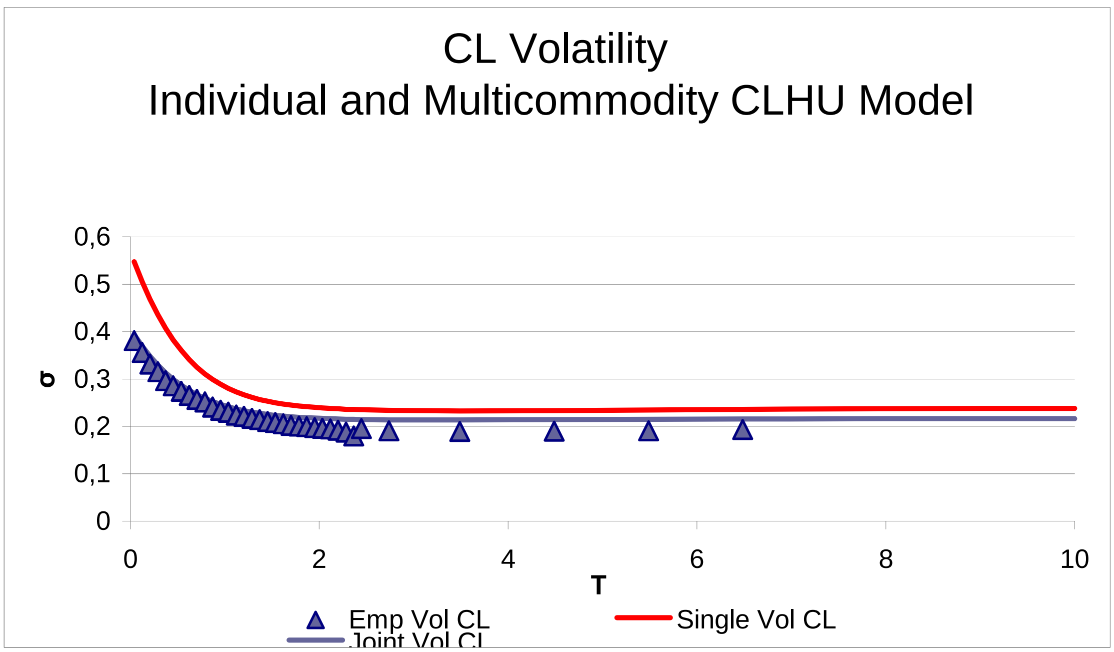
An analysis of the above Tables and Figures show a much better behavior of multicommodity models on extreme out of sample testing for extrapolating long-maturity prices for a commodity with only short-maturity contracts. The figures also show better spread estimates using multicommodity models. Finally, given that individual-commodity models have more independent factors, it is not surprising that their in-sample errors are smaller.

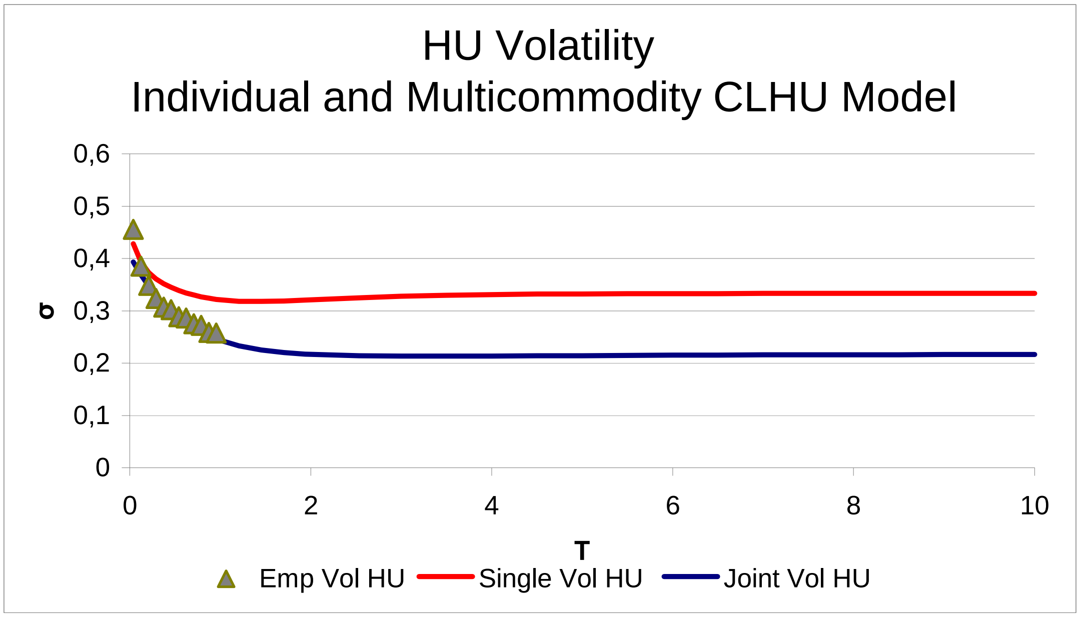
V.4.2. Volatility Adjustments

In this section we show how model-volatility compares to empirical-volatility for each commodity.




CB Volatility
Individual and Multicommodity CLCB Model
0,6
0,5
0,4
σ
0,3
0,2
0,1
0
0
2
4
T
6
8
10
Emp Vol CB
Single Vol CB
Joint Vol CBj


CL Volatility
Individual and Multicommodity CLHU Model
0,6
0,5
0,4
σ
0,3
0,2
0,1
0
0
2
4
6
8
10
T
Emp Vol CL
Single Vol CL
Joint Vol CL



For all the above cases, it can be seen that multicommodity volatility estimates track much more closely empirical data volatilities. This should induce commodity spreads to be much more stable.

VI. CONCLUSION

This article proposes a multicommodity model of futures prices to explain the stochastic behavior of more than one commodity. Jointly modeling more than one commodity has the advantage of being able to use long-maturity futures prices of one commodity to estimate futures prices for another commodity which only has short-maturity contracts.

The model considers that commodity prices have a set of common factors that explain the correlation among them, in addition to some commodity-specific factors. The multicommodity model is based on the canonical representation of Dai and Singleton (2000) for interest rates and represents an extension of the individual-commodity model in Cortazar and Naranjo (2006).

A procedure for choosing the number of common and commodity-specific un-observable-Gaussian factors is presented. Also, it is shown how commodities with and without seasonality may be modeled together and how to estimate the multicommodity model using a Kalman Filter.

A first empirical implementation of the proposed multicommodity model is presented for the WTI (CL) and the Brent oil (CB) contracts. A (3, 0, 1) CLCB model is chosen, with three common factors and one commodity-specific factor for the CB contract.

A second implementation for the WTI (CL) and the Unleaded Gasoline (HU) is discussed. A (2, 1, 1) CLHU model is chosen, with two common factors and one commodity-specific factor for each of the two commodities. In addition, the HU contract is assumed to have 12 monthly constants to fit its seasonal behavior.

Results for both model implementations show strong improvements over the traditional individual-commodity models, with much lower out-of-sample errors and better volatility estimates, even when using fewer factors.

The advantages of using these multicommodity models is specially clear when model estimates are compared with data not used at all in model calibration, in what is call extreme out-of-sample testing. Also if spreads between two commodities are of interest, using multicommodity models provides much more stable estimates, as will be shown in the next version of this paper.

VII. BIBLIOGRAPHY

BRENNAN, M.J. and SCHWARTZ, E.S. (1985) Evaluating natural resources investments. Journal of Business, Vol. 58, N° 2, 135-157.

CASASSUS, J., AND COLLIN-DUFRESNE, P. (2005). Stochastic Convenience Yield implied from Commodity Futures and Interest Rates. The Journal of Finance, 60, 2283-2331.

CORTAZAR, G., AND NARANJO, L. F. (2006). An N-Factor Gaussian Model of Oil Futures. Journal of Futures Markets, Vol 26, 243-268.

CORTAZAR, G. and SCHWARTZ, E.S. (1994) The valuation of commodity-contingent claims. The Journal of Derivatives, Vol. 1, 27-35.

CORTAZAR, G., AND SCHWARTZ, E. S. (2003). Implementing a stochastic model for oil futures prices. Energy Economics, 25, 215-238.

DAI, Q. and SINGLETON, K.J. (2000) Specification analysis of affine term structure models. The Journal of Finance, Vol. 55, N° 5, 1943-1978.

FLURY(1988). Common principal components and related multivariate models. Wiley, New York.

GIBSON, R. and SCHWARTZ, E.S. (1990) Stochastic convenience yield and the pricing of oil contingent claims. The Journal of Finance, Vol. 45, N° 3, 959-976.

KRZANOWSKI J. (1979). Between Groups Comparison of Principal Components. Journal of the American Statistical Association, Vol.74, Nº367,703-707.

LAUGHTON D. AND JACOBY H. (1993). Reversion, Timing Options, and Long-Term Decision-Making. Financial Management vol. 22(3), Fall.

LITTERMAN R. AND SCHEINKMAN R. (1991). Common Factors Affecting Bond Returns. Journal of Fixed Income.

MANOLIU M. AND TOMPAIDIS S. (2002). Energy futures prices: term structure models with Kalman filter estimation. Applied Mathematical Finance, 9, 21-43.

PHILIPS P. AND ARNOLD S. (1999). Hierarchical Comparison of Genetic Variance-Covariance Matrices I. Using the Flury Hierarchy. Evolution, Vol. 53,Nº5,1506-1515.

SCHWARTZ, E.S. (1997) The stochastic behaviour of commodity prices: Implications for valuation and hedging. The Journal of Finance, Vol. 52, N° 3, 923-973.

SØRENSEN, C. (2002) Modeling seasonality in agricultural commodity futures. Journal of Futures Markets, Vol. 22, 393-426.

TOLMASKY C.AND HINDANOV D. (2002). Principal Components Analysis for correlated curves and seasonal commodities: The Case of Petroleum Market. Journal of Futures Markets, Vol. 22, Nº11. 1019.

TROLLE A. AND SCHWARTZ S.(2006). Unspanned Stochastic Volatility and the Pricing of Commodity Derivatives. NBER Working Paper No. 12744.