



456

2014

Maximal Gaussian Affine Models for Multiple Commodities: A Note

Jaime Casassus, Peng Liu, Ke Tang.

Maximal Gaussian Affine Models for Multiple Commodities: A Note *

Jaime Casassus

Pontificia Universidad Catolica de Chile

jcasassus@uc.cl

Peng (Peter) Liu

Cornell University

pl333@cornell.edu

Ke Tang

Hanqing Advanced Institute of Economics and Finance

Renmin University of China

ketang@ruc.edu.cn

*Casassus acknowledges financial support from FONDECYT (grant No. 1110841) and from Grupo Security through FinanceUC. Liu acknowledges financial support from Institute of Social Science at Cornell University (small grant program). Tang acknowledges financial support from the National Natural Science Foundation of China (grant No. 71171194). Any errors or omissions are the responsibility of the authors. Please address any comments to Jaime Casassus, Instituto de Economia, Pontificia Universidad Catolica de Chile, email: jcasassus@uc.cl; Peng Liu, Cornell University, 465 Statler Hall, Ithaca, NY, 14850, email: pl333@cornell.edu; Ke Tang, Mingde Building, Hanqing Advanced Institute of Economics and Finance and School of Finance, Renmin University of China, Beijing, 100872, email: ketang@ruc.edu.cn.

Maximal Gaussian Affine Models for Multiple Commodities: A Note

Abstract

This paper extends the maximal affine models of single assets to a multi-commodity setup. We show that the correlated version of maximal affine models for a single commodity is no longer maximal for multiple commodities. In the maximal model, the convenience yield of a certain commodity could depend on the prices of other commodities, which is consistent with the structural model in our companion paper Casassus, Liu, and Tang (2013). This cross-commodity relationship is a feedback effect that may generate substantial comovement among long-run commodity prices, a fact that is consistent with many empirical studies.

Keywords: maximal affine models, Gaussian models, multiple commodities, futures prices.

JEL Classification: G13, C33

1 Introduction

Recently, there has been an increasing interest in studying multi-commodity pricing models. Cortazar, Milla, and Severino (2008) use long-maturity futures prices of one commodity to estimate the futures prices of another commodity that only has short-maturity contracts. Paschke and Prokopczuk (2009) present an integrated approach for multiple commodities and show its economic significance studying the risk management decisions of a refinery. Casassus, Liu, and Tang (CLT, 2013) show that economic linkages such as production relationships among commodities imply a new pattern of co-movement across commodity prices. Using a multi-commodity affine framework for futures prices, they show that the convenience yield of a certain commodity is determined by, among other things, the prices of related commodities.¹ However, none of the above-mentioned papers analyze if the proposed models are “maximal” or “maximally flexible”, in the sense that, conditional on observing the commodity prices, the model offers the maximum number of identifiable parameters (Dai and Singleton 2000). In this paper, we propose a maximal affine model for multiple commodities, and more importantly, study the implication of maximality in the co-movement of commodity prices.

The canonical “maximal” Gaussian affine multi-factor model proposed by Duffie and Kan (1996), Duffie, Pan, and Singleton (2000), and Dai and Singleton (2000), is defined for *single* assets V_i as follows:

$$V_i := \theta_0^i + \psi_Y^i Y, \tag{1}$$

where V_i denotes the (log) value of the i^{th} asset, θ_0^i is a constant, ψ_Y^i is a $1 \times m$ constant row vector, and Y is an $m \times 1$ column vector representing the state variables that follow Gaussian diffusion processes under the risk-neutral measure,

$$dY = -\Lambda Y dt + d\beta_Y^Q \tag{2}$$

¹For simplicity, we refer the convenience yield as the net convenience yield, i.e. the true convenience less the percentage storage cost.

where Λ is a lower triangular matrix and β_Y^Q is a vector of independent Brownian motions. This model is “maximal” in the Dai and Singleton (2000)’s sense. Casassus and Collin-Dufresne (CCD, 2005) employ the above framework to commodity futures pricing and propose a maximal affine model for a single commodity.

In this paper we extend the above-mentioned maximal Gaussian affine model to a multi-commodity framework.² The natural extension of the maximal affine model to multiple commodities is to consider equations (1) and (2) for each commodity and allow the β ’s to be correlated across commodities. We show that this correlated version of maximal affine models is no longer maximal for multiple commodities. We propose a $2n$ -factor canonical model for a system of n commodities that is maximal and that has some important characteristics consistent with recent multi-commodity studies. In particular, in our maximal model, the convenience yield of a certain commodity could depend on the prices of other commodities, which is consistent with the structural model in our companion paper Casassus, Liu, and Tang (2013). This cross-commodity relationship that relates the expected returns of the commodities is a new *feedback effect* not present in the natural extension of the traditional Gaussian affine models. The main contribution of this feedback effect is to allow for substantial co-movement among long-run commodity prices, a fact that is desirable given strong empirical evidence that points in this direction. We also extend our model to consider stochastic interest rates.

The remainder of the paper is organized as follows. Section 2 lays out the framework of maximal affine models for multiple commodities. Section 3 shows that under the multi-commodity maximal affine framework, commodity prices follow an error correction (feedback) relationship. Section 4 specifies the commodity futures price dynamics. Section 5 develops the model under the assumption that interest rate is stochastic. Section 6 concludes.

²The Gaussian nature of our model could be relaxed in favor of considering stochastic volatility and affine jumps, however, we follow recent single and multiple commodity studies and keep the joint distribution of the factors Gaussian (for example, Schwartz 1997, CCD 2005, Cortazar, Milla, and Severino 2008, Paschke and Prokopczuk 2009 and CLT 2013). This simplifies the analysis and allows us to highlight better the contribution of our model, which is the effect of maximality in the co-movement of the commodity prices. Moreover, as we will see later, “admissibility” in our multi-commodity framework would imply that the underlying factors cannot affect the volatility of the processes.

2 Maximal Affine Models for Multiple Commodities

We first define the maximal model for multiple assets as follows:

In a system of n assets which are governed by m factors, a model for the system is “maximal” if and only if every single asset in the system is modeled by an m -factor “maximal” model as defined in equation (1),

$$V := \theta_Y + \psi_Y Y, \quad (3)$$

where $V = (V_1, \dots, V_n)$ represent the (log) prices of n assets which are governed by Y in equation (2), $\theta_Y := (\theta_Y^1, \theta_Y^2, \dots, \theta_Y^n)^\top$ is an $n \times 1$ vector and $\psi_Y := (\psi_Y^1, \psi_Y^2, \dots, \psi_Y^n)^\top$ is an $n \times m$ matrix.

The later part of this section shows that a simple combination of maximal models for single commodities does not necessarily generate a maximal model in a multi-commodity system. Furthermore, Dai and Singleton (2000) and CCD models only allow for a *constant* θ_Y , while many commodity prices are subject to seasonal movements. Thus, we need to extend the model by letting θ_Y be time-varying. Therefore, we rewrite

$$\theta_Y^i(t) = \chi_Y^i + \omega_Y^i(t), \quad (4)$$

where $\omega_Y^i(t)$ is a periodical function and χ_Y^i is a constant.

In most commodity futures models, (log) spot prices and convenience yields are the main drivers of movements in commodity prices. Hence in this paper, without loss of generality, we assume that an n -commodity system is governed by $2n$ factors, n (log) spot price factors, and n convenience yield factors. Note that the interest rate r^f is constant in the setup of the model. Therefore, the log spot price and convenience yield vectors V and Δ are specified as $V = (x_1, x_2, \dots, x_n)^\top$ with $x_i := \ln S_i$, where S_i is the spot price for the i^{th} commodity; and $\Delta := (\delta_1, \delta_2, \dots, \delta_n)^\top$ with δ_i representing the convenience yield of the i^{th} commodity. Following Schwartz (1997) and CCD (2005), we assume the spot commodities are tradable,³ the absence

³If the spot price is actually not a traded asset (as would be the case for electricity futures, for example), then (5) needs not hold, however the process δ defined by equation (5) is still of interest, as it reflects, per definition, how much the spot price dynamics differ from that of a traded asset.

of arbitrage implies that under the risk-neutral measure the drift of the spot price of the i^{th} commodity must follow

$$E_t^Q[dS_i] = (r^f - \delta_i)S_i dt, \quad i = 1, \dots, n. \quad (5)$$

Applying Ito's lemma, we obtain the following expression for the convenience yield vector,

$$\begin{aligned} \Delta &= r^f \mathbf{1}_n - \frac{E_t^Q[dV] + \frac{1}{2}(\text{Var}_t[dx_1], \dots, \text{Var}_t[dx_n])^\top}{dt} \\ &= r^f \mathbf{1}_n + \psi_Y \Lambda Y - \frac{1}{2}((\psi_Y \psi_Y^\top)_{1,1}, \dots, (\psi_Y \psi_Y^\top)_{n,n})^\top \end{aligned} \quad (6)$$

where Var_t denotes the variance at time t , and $\mathbf{1}_n$ is an $n \times 1$ column vector with all elements equal to 1.

By writing equations (3) and (6) together, we have

$$Z := \begin{pmatrix} V \\ \Delta \end{pmatrix} = \begin{pmatrix} \theta_Y \\ \theta_c \end{pmatrix} + \begin{pmatrix} \psi \\ \psi_Y \Lambda \end{pmatrix} Y, \quad (7)$$

where $\theta_c := r^f \mathbf{1}_n - \frac{1}{2}((\psi_Y \psi_Y^\top)_{1,1}, \dots, (\psi_Y \psi_Y^\top)_{n,n})^\top$.

Proposition 1 *Assume that the $2n$ affine factors, specified by n log spot prices x_i and n convenience yields δ_i ($i = 1, \dots, n$), drive the movement of futures prices in an n -commodity system. The maximal model under the risk-neutral measure is as follows:*

$$dZ = \bar{U} dt + \bar{\Psi} Z dt + d\beta_Z^Q \quad (8)$$

where $d\beta_Z^Q$ is the Wiener-process increment vector with covariance matrix $\bar{\Omega}$,

$$\bar{\Omega} = \begin{pmatrix} \bar{\sigma}_1^2 & \bar{\rho}_{1,2} \bar{\sigma}_1 \bar{\sigma}_2 & \dots & \bar{\rho}_{1,2n} \bar{\sigma}_1 \bar{\sigma}_{2n} \\ \bar{\rho}_{1,2} \bar{\sigma}_1 \bar{\sigma}_2 & \bar{\sigma}_2^2 & \ddots & \bar{\rho}_{2,2n} \bar{\sigma}_2 \bar{\sigma}_{2n} \\ \vdots & \ddots & \ddots & \vdots \\ \bar{\rho}_{1,2n} \bar{\sigma}_1 \bar{\sigma}_{2n} & \dots & \bar{\rho}_{2n-1,2n} \bar{\sigma}_{2n-1} \bar{\sigma}_{2n} & \bar{\sigma}_{2n}^2 \end{pmatrix}. \quad (9)$$

$\bar{U} = (\bar{R}, \bar{L})^\top$ and

$$\bar{R} : = (r^f - \frac{1}{2}\bar{\sigma}_1^2, r^f - \frac{1}{2}\bar{\sigma}_2^2, \dots, r^f - \frac{1}{2}\bar{\sigma}_n^2)^\top \quad (10)$$

$$\bar{L} : = (\bar{\theta}_1(t), \bar{\theta}_2(t), \dots, \bar{\theta}_n(t))^\top, \bar{\theta}_i(t) = \bar{\chi}_i + \bar{\omega}_i(t), \quad (11)$$

and

$$\bar{\Psi} := \begin{pmatrix} 0 & -I_{n \times n} \\ \bar{A} & \bar{B} \end{pmatrix}, \quad (12)$$

with

$$\bar{A} = \begin{pmatrix} \bar{a}_{1,1} & \bar{a}_{1,2} & \dots & \bar{a}_{1,n} \\ \bar{a}_{2,1} & \bar{a}_{2,2} & \ddots & \bar{a}_{2,n} \\ \vdots & \ddots & \ddots & \vdots \\ \bar{a}_{n,1} & \bar{a}_{n,2} & \dots & \bar{a}_{n,n} \end{pmatrix}, \bar{B} = \begin{pmatrix} \bar{b}_{1,1} & \bar{b}_{1,2} & \dots & \bar{b}_{1,n} \\ \bar{b}_{2,1} & \ddots & \bar{b}_{2,n-1} & \bar{b}_{2,n} \\ \vdots & \ddots & \ddots & \vdots \\ \bar{b}_{n,1} & \dots & \bar{b}_{n,n-1} & \bar{b}_{n,n} \end{pmatrix} \quad (13)$$

Proof. Equation (7) can be seen as an invariant transformation from Z to Y (see also Dai and Singleton, 2000, and CCD). This transformation rotates the state variables, but all the initial properties of the model are maintained, that is, the resulting model is still a “maximal” $2n$ factor Gaussian model. Furthermore, we apply Ito’s lemma to obtain the specific relationships between the model parameters specified in equation (18) and those specified in (2) and (3); appendix A shows the derivation in a greater detail. ■

Proposition 1 shows that when matrix \bar{A} is non diagonal the dynamics of the convenience yield of a certain commodity and therefore, the convenience yield itself, depends on other commodity prices. This is consistent with the main conclusion of our companion paper CLT (2013), in which the relative scarcity, represented by the price difference between a certain commodity and a related commodity, is an important part of the convenience yield. This results extends the existing commodity literature in which the convenience yield of a certain commodity *only* depends on its own characteristics (e.g., Gibson and Schwartz 1995, Schwartz 1997, and CCD

2005).

A corollary of Proposition 1 is that the combination of correlated maximal models for a single commodity is not maximal in a multi-commodity setting.⁴ Indeed, the combination of maximal models for single commodities restricts $\bar{a}_{i,j \neq i} = 0$ and $\bar{b}_{i,j \neq i} = 0$ ($i = 1, \dots, n, j = 1, \dots, n$). In other words, the maximal model for multiple commodities and nests the combination of maximal models for single commodities.

For example, in a two-commodity system (say, heating and crude oil) the maximal affine model is:

$$dx_1 = (r^f - \frac{1}{2}\bar{\sigma}_1^2 - \delta_1)dt + \bar{\sigma}_1 dW_1^Q \quad (14)$$

$$dx_2 = (r^f - \frac{1}{2}\bar{\sigma}_2^2 - \delta_2)dt + \bar{\sigma}_2 dW_2^Q \quad (15)$$

$$d\delta_1 = (\bar{\theta}_1(t) + \bar{a}_{1,1}x_1 + \boxed{\bar{a}_{1,2}}x_2 + \bar{b}_{1,1}\delta_1 + \boxed{\bar{b}_{1,2}}\delta_2)dt + \bar{\sigma}_3 dW_3^Q \quad (16)$$

$$d\delta_2 = (\bar{\theta}_2(t) + \boxed{\bar{a}_{2,1}}x_1 + \bar{a}_{2,2}x_2 + \boxed{\bar{b}_{2,1}}\delta_1 + \bar{b}_{2,2}\delta_2)dt + \bar{\sigma}_4 dW_4^Q \quad (17)$$

However, the combination of maximal models for single commodities is not maximal because it restricts the parameters in the boxes to be zero. The following section shows how these restricted parameters affect the joint dynamics of commodity prices and, in particular, how they are related to the long-term co-movement of the prices.

3 Co-movements of Commodity Prices

In this section, we show that under the setup of the maximal model, the commodity prices can follow an error correction (feedback) relationship. This is also what we find in our companion paper, i.e. commodity prices follow error correction because of long-term economic relationships. The feedback effect makes two commodity prices co-move more tightly in the long run and hence is substantial in analyzing the co-movements of the multi-commodity prices. The error correction

⁴The correlated versions of maximal models for single commodities correspond to cases in which spot prices and convenience yields of separate commodities are correlated with each other.

between two commodities is different with correlation between them; normally error correction comes into play in the long run, but correction is normally temporary.

Proposition 2 *The “maximal” model specified in Proposition 1 is identical to the following model:*

$$dX = Udt + \Psi Xdt + d\beta^Q \quad (18)$$

where $X := (x_1, \dots, x_n, \eta_1, \dots, \eta_n)^\top$ with η_i follows a mean-reverting process,

$$d\eta_i = (\theta_i(t) - k_i\eta_i)dt + \sigma_{n+i}dW_{n+i}^Q \quad \text{where } n = 1, \dots, n \quad (19)$$

$U = (R, L)^\top$ and

$$R : = (r^f - \frac{1}{2}\sigma_1^2, r^f - \frac{1}{2}\sigma_2^2, \dots, r^f - \frac{1}{2}\sigma_n^2)^\top, \quad (20)$$

$$L : = (\theta_1(t), \theta_2(t), \dots, \theta_n(t))^\top, \theta_i(t) = \chi_i + \omega_i(t). \quad (21)$$

$$\Psi := \begin{pmatrix} B & A \\ 0 & D \end{pmatrix} \text{ with}$$

$$B = \begin{pmatrix} b_{1,1} & b_{1,2} & \dots & b_{1,n} \\ b_{2,1} & \ddots & b_{2,n-1} & b_{2,n} \\ \vdots & \ddots & \ddots & \vdots \\ b_{n,1} & \dots & b_{n,n-1} & b_{n,n} \end{pmatrix}, A = \begin{pmatrix} -1 & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & -1 & \ddots & a_{2,n} \\ \vdots & \ddots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \dots & -1 \end{pmatrix} \quad (22)$$

$$D = \text{diag}(-k_1, -k_2, \dots, -k_n). \quad (23)$$

$d\beta^Q$ is a Wiener-process increment vector with the covariance matrix Ω specified as,

$$\Omega = \begin{pmatrix} \sigma_1^2 & \rho_{1,2}\sigma_1\sigma_2 & \dots & \rho_{1,2n}\sigma_1\sigma_{2n} \\ \rho_{1,2}\sigma_1\sigma_2 & \sigma_2^2 & \ddots & \rho_{2,2n}\sigma_2\sigma_{2n} \\ \vdots & \ddots & \ddots & \vdots \\ \rho_{1,2n}\sigma_1\sigma_{2n} & \dots & \rho_{2n-1,2n}\sigma_{2n-1}\sigma_{2n} & \sigma_{2n}^2 \end{pmatrix}. \quad (24)$$

Proof. Let $\eta := (\eta_1, \eta_2, \dots, \eta_n)^\top$, from the notion that the drift of spot prices must equal risk-free rate less the convenience yield in the risk-neutral measure (see equation 5), we thus can show that the convenience yield vector Δ has to follow

$$\Delta = -BV - A\eta, \quad (25)$$

and therefore,

$$X = \begin{pmatrix} V \\ \eta \end{pmatrix} = \begin{pmatrix} I_{n \times n} & 0 \\ -A^{-1}B & -A^{-1} \end{pmatrix} Z \quad (26)$$

which can be seen as an invariant transform from Z to X . As with Proposition 1, we apply Ito's lemma to compare the parameters in (8) and (26) show that they are identical. Appendix B shows the derivation in detail. ■

Note that Proposition 2 shows that the movement of a certain commodity might depend on the prices of other commodities when B is not a diagonal matrix. For example, in a two-commodity system, the maximal model can be written as

$$dx_1 = (r^f - \frac{1}{2}\sigma_1^2 + b_{1,1}x_1 + \overline{b_{1,2}}x_2 - \eta_1 + \overline{a_{1,2}}\eta_2)dt + \sigma_1 dW_1^Q \quad (27)$$

$$dx_2 = (r^f - \frac{1}{2}\sigma_2^2 + \overline{b_{2,1}}x_1 + b_{2,2}x_2 + \overline{a_{2,1}}\eta_1 - \eta_2)dt + \sigma_2 dW_2^Q \quad (28)$$

$$d\eta_1 = (\theta_1(t) - k_1\eta_1)dt + \sigma_3 dW_3^Q \quad (29)$$

$$d\eta_2 = (\theta_2(t) - k_2\eta_2)dt + \sigma_4 dW_4^Q \quad (30)$$

As before, when the parameters in boxes are restricted to be zero, the maximal model for multiple commodities reduces to the combination of maximal models for single commodities. Furthermore, equation (27) can be rewritten as

$$dx_1 = [r^f - \frac{1}{2}\sigma_1^2 + b_{1,2}(x_2 - cx_1 - d) + (b_{1,2}c + b_{1,1})x_1 - \eta_1 + a_{1,2}\eta_2 + b_{1,2}d]dt + \sigma_1 dW_1^Q$$

where $x_2 = cx_1 + d$ represents the long-term equilibrium relationship between x_1 and x_2 , which is assumed to be known in advance. Thus, when $b_{1,2} \neq 0$, there is an error-correction relationship between x_1 and x_2 . Specifically, when $b_{1,2} > 0$, a positive error in the long-term economic relationship tends to feed a positive correction back on x_1 ; hence, x_1 and x_2 move together; when $b_{1,2} < 0$, x_1 and x_2 move in opposite directions. Similarly, the movement of x_2 can be determined following the same logic. In the econometrics literature, $(x_1, x_2)^\top$ follows a vector error-correlation model; moreover, when $b_{1,2}c + b_{1,1} = 0$ and $b_{2,1} + cb_{2,2} = 0$, x_1 and x_2 are cointegrated with each other (see Engle and Granger 1987).

Note that when the parameters in the boxes are restricted to be zero, the model cannot capture the error-correlation or cointegration relationship. However, many researchers have found cointegration (error-correction) in multiple commodities, such as Malliaris and Urrutia (1996), Girma and Paulson (1999) and Paschke and Prokopczuk (2009).⁵ This shows that the maximal model for multiple commodities is necessary to capture the error-correction relationship among commodities. CLT (2013) utilize the models in (27) to (30) estimate the heating oil and WTI crude oil pair, the WTI and Brent crude oil pair, and the heating oil and unleaded gasoline pair. The results for all three pairs show that $b_{2,1}$ or $b_{1,2}$ is strongly different from zero. This is because these commodity pairs follow a long-term economic relationships.

⁵The definition of cointegration is that different time series depend on one or more non-stationary factor and other stationary factors. Starting from this point, Paschke and Prokopczuk (2009) assumes energy commodities are cointegrated under the physical measure, therefore, the time series of their prices depend on many factors, one of them being non-stationary. Typically, only the non-stationary time series has the feature of cointegration, while under the risk neutral measure, this normally is not the case. For example, both CCD (2005) and CLT (2013) find that in the risk neutral measure the commodity prices are much more mean-reverting than the that in the physical measure.

The error-correction characteristic present among commodity prices has important implications for the underlying process of our $2n$ -factor maximal model. Indeed, the error-correction relationship implies that the convenience yield cannot be guaranteed to be positive, because it depends not only on its own (log) price but also on the price of other commodities.⁶ This implies that convenience yields, as well as log prices, can take both positive and negative signs (something that it is actually consistent with the data). If our $2n$ -factor maximal model were non Gaussian, one would expect that prices and/or convenience yields may affect the volatility of the underlying processes. If this were the case, some “admissibility” restrictions, again in the Dai and Singleton (2000)’s sense, would have to be imposed in order to have well-defined futures prices in our multi-commodity framework. Since both log prices and convenience yields cannot be forced to stay with one sign, “admissibility” would imply that the coefficients of these variables in the volatility of the processes should be zero, taking the model back to the Gaussian case.

4 Futures Pricing

Since the setup of the model is an affine structure, we know that the futures prices is an exponentially affine function of the factors. Denote the futures price $F_i(X, \tau)$ for the i^{th} commodity with time to maturity of $\tau := T - t$, where T is the futures maturity.

Proposition 3 *In model setup (18), futures prices $F_i(X, \tau)$ are determined by*

$$F_i(X, \tau) = \exp(m_i(\tau) + G_i(\tau)^\top X), \quad (31)$$

where

$$\begin{aligned} G(\tau) &= \exp(\Psi^\top \tau) \\ m_i(\tau) &= \int_0^\tau \left[G_i(s)^\top U + \frac{1}{2} (G(s)^\top \Omega^\top \Omega G(s))_{i,i} ds \right] \end{aligned}$$

⁶To see this, recall how the convenience yield affect the expected log spot price change under the risk neutral measure, i.e., $E^Q [dx_i] = (r^f - \frac{1}{2}\sigma_i^2 - \delta_i) dt$.

$G_i(\tau)$ denotes the i^{th} column of the $G(\tau)$ matrix.

Proof. See Appendix C. ■

Since the log of futures prices can be expressed as an affine structure of factors in X , it is natural to use the Kalman filter to estimate the maximal model. For the empirical estimation of multiple commodity systems, we refer readers to our companion paper CLT (2013).

5 Maximal Affine Models with Stochastic Interest Rates

One important extension to our previous $2n$ factor model is to allow the interest rate to be stochastic. Considering stochastic interest rates is specially important if the objective is to correctly back out the convenience yield from the futures prices. For this reason, any researchers have included stochastic interest rates as a factor in commodity pricing, for example Schwartz (1997) and CCD (2005).⁷

When we enhanced our model to consider stochastic interest rates, the total number of factors is $2n + 1$. Following Schwartz (1997) and CCD (2005) we assume that the risk-free interest rate follows an autonomous one factor Ornstein-Uhlenbeck process⁸ in the risk neutral measure:

$$dr_t^f = k_r(\theta_r - r_t^f)dt + \sigma_r dW_r^Q.$$

Following similar procedure of CCD (2005) and Appendix A, the maximal form of model (8) to (13) do not change except that \bar{L} in (11) change to

$$\bar{L} := (\bar{\theta}_1(t) + \bar{c}_1 r_t^f, \bar{\theta}_2(t) + \bar{c}_2 r_t^f, \dots, \bar{\theta}_n(t) + \bar{c}_n r_t^f)^\top, \bar{\theta}_i(t) = \bar{\chi}_i + \bar{\omega}_i(t). \quad (32)$$

where $\bar{c}_1, \bar{c}_2, \dots, \bar{c}_n$ are constants. Similarly, the only different part for the maximal model of (18)

⁷We thank the referee for pointing this out.

⁸This is because interest rates are determined by the bond markets, which we assume to be autonomous. Also, the interest rate could also follow a Cox-Ingersoll-Ross type of process, however, our main results would remain unchanged.

to (24) when adding in a stochastic interest rate is that R in (20) changes to

$$R := (c_1 r_t^f - \frac{1}{2} \sigma_1^2, c_2 r_t^f - \frac{1}{2} \sigma_2^2, \dots, c_n r_t^f - \frac{1}{2} \sigma_n^2)^\top, \quad (33)$$

where c_1, c_2, \dots, c_n are constants.

6 Conclusions

In this paper, we propose a “maximal” affine model for a multi-commodity setup as an extension of Dai and Singleton (2000) and Casassus and Collin-Dufresne (2005). We show that the combination of maximal affine models for single commodities is no longer maximal in the multiple commodity system. Under the setup of the maximal model, the convenience yield of a certain commodity can depend on the price of other commodities in the system, and different commodities can follow a feedback (or error-correction) relationship with each other.

References

- [1] Casassus, J. and Collin-Dufresne, P. (2005), Stochastic Convenience Yield Implied from Commodity Futures and Interest Rates, *Journal of Finance*, 2283-2331
- [2] Casassus, J., Liu, P. and Tang, K. (2013), Economic Linkages, Relative Scarcity, and Commodity Futures Returns, *Review of Financial Studies*, 26, 1324-1362.
- [3] Cortazar, G., Milla, C. and Severino, F. (2008), A Multicommodity Model of Futures Prices: Using Futures Prices of One Commodity to Estimate the Stochastic Process of Another, *Journal of Futures Markets*, 28, 537-560.
- [4] Dai, Q. and Singleton, K. (2000), Specification Analysis of Affine Term Structure Models, *Journal of Finance*, 55, 1943-1978.
- [5] Duffie, D., and Kan, R. (1996), A yield-factor model of interest rates, *Mathematical Finance* 6, 379-406.
- [6] Duffie, D., Pan, J. and Singleton, K. (2000), Transform analysis and asset pricing for affine jump-diffusions, *Econometrica* 68, 1343-1376.
- [7] Engle, R., and Granger, C. (1987), Co-integration and error correction: Representation, estimation, and testing, *Econometrica* 55, 251-276.
- [8] Gibson R. and Schwartz, E. (1990), Stochastic Convenience Yield and the Pricing of Oil Contingent Claims, *Journal of Finance*, 45, 959-976.
- [9] Girma, P., and Paulson, A. (1999), Risk arbitrage opportunities in petroleum futures spreads, *Journal of Futures Markets* 19, 931-955.
- [10] Higham, N. and Kim, H. (2001), Solving a Quadratic Matrix Equation by Newton-Method with Exact Line Searches, *SIAM Journal Matrix Analysis and Application*, 23, 303-316.

- [11] Malliaris, A., and Urrutia, J. (1996), Linkages between agricultural commodity futures contracts, *Journal of Futures Markets* 16, 595-609.
- [12] Paschke, R., and Prokopczuk, M. (2009), Integrating multiple commodities in a model of stochastic price dynamics, *Journal of Energy Markets*, 2, 47-82.
- [13] Schwartz, E.S, (1997), The Stochastic Behaviour of Commodity Prices: Implications for Valuation and Hedging, *Journal of Finance*, 52 923-973.
- [14] Smith, H., Singh, R. and Sorensen, D. (1995) Formulation and solution of the non-linear,damped eigenvalue problem for skeletal systems, *International Journal of Numerical Methods of Engineering*, 38, 3071-3085.

Appendixes:

A. Proof of Proposition 1

Equation (7) specifies a unique transformation from the latent variables Y to Z . Thus the Z processes in (7) preserves the “maximal” specification of the model. Letting $\Gamma_0 := \begin{pmatrix} \theta_Y \\ \theta_c \end{pmatrix}$, $\Gamma_Y := \begin{pmatrix} \psi_Y \\ \psi_Y \Lambda \end{pmatrix}$ and employing Ito’s lemma on (7) we see that

$$dZ = \Gamma_Y \Lambda \Gamma_Y^{-1} \Gamma_0 dt - \Gamma_Y \Lambda \Gamma_Y^{-1} Z dt + \Gamma_Y d\beta_Y^Q, \quad (34)$$

Denoting $\psi_Y = \begin{pmatrix} \psi_1 & \psi_2 \end{pmatrix}$ and $\Lambda = \begin{pmatrix} \Lambda_1 & 0 \\ \Lambda_2 & \Lambda_3 \end{pmatrix}$ where $\psi_1, \psi_2, \Lambda_1, \Lambda_2, \Lambda_3$ are all $n \times n$ matrixes and, comparing this with (8), we have,

$$\begin{aligned} \bar{\Omega} &= \psi_Y^\top \psi_Y + K^\top \psi_Y^\top \psi_Y \Lambda \\ \bar{B} &= (\psi_2 \Lambda_3 \psi_2^{-1} \psi_1 - \psi_1 \Lambda_1 - \psi_2 \Lambda_2)^{-1} (\psi_1 \Lambda_1^2 + \psi_2 \Lambda_2 \Lambda_1 + \psi_2 \Lambda_3 \Lambda_2 - \psi_2 \Lambda_3^2 \psi_2^{-1} \psi_1) \\ \bar{A} &= (\psi_1 \Lambda_1 + \psi_2 \Lambda_2 - \psi_2 \Lambda_3 \psi_2^{-1} \psi_1)^{-1} (\psi_1 \Lambda_1^2 + \psi_2 \Lambda_2 \Lambda_1 + \psi_2 \Lambda_3 \Lambda_2 - \psi_2 \Lambda_3^2 \psi_2^{-1} \psi_1) \psi_2 \Lambda_3 \psi_2^{-1} - \psi_2 \Lambda_3^2 \psi_2^{-1} \end{aligned} \quad (35)$$

Note that there are, in total, $n + 2n^2$ parameters in Λ , and $2n^2$ in ψ_Y . Also, there are, in total, $n + 2n^2$ parameters in $\bar{\Omega}$ and $2n^2$ in \bar{A} and \bar{B} . Thus, there is a one-one relationship from (Λ, ψ_Y) to $(\bar{\Omega}, \bar{A}, \bar{B})$. Given \bar{B} and \bar{A} , \bar{R} can be determined easily from $\bar{\Omega}$; for \bar{L} we have

$$\bar{L} = -(\bar{A}\theta_Y + \bar{B}\theta_c), \quad (36)$$

Specifically, $\bar{\chi}_i = \sum_{k=1}^n \bar{A}_{i,k} \chi_Y^i + \sum_{k=1}^n \bar{B}_{i,k} (\theta_c)_k$, and $\bar{\omega}_i(t) = \sum_{k=1}^n \bar{A}_{i,k} \omega_Y^i(t)$.

Therefore (8) is the “maximal” specification.

B. Proof of Proposition 2

(26) specifies a unique transformation from the latent variables Z to X . Denote $\Gamma_Z = \begin{pmatrix} I_{n \times n} & 0 \\ -A^{-1}B & -A^{-1} \end{pmatrix}$. Applying Ito's lemma to equation (26) yields

$$dX = \Gamma_Z \bar{U} + \Gamma_Z \bar{\Psi} \Gamma_Z^{-1} X + \Gamma_Z d\beta_Z^Q. \quad (37)$$

By comparing the parameters in (37) and those in (18), we find that if the following equations hold, the two models are identical.

$$0 = B^2 - \bar{B}B + \bar{A} \quad (38)$$

$$\bar{B} - B = ADA^{-1} \quad (39)$$

$$\Omega = (\Gamma_Z)^\top \bar{\Omega} \Gamma_Z \quad (40)$$

(38) is a quadratic matrix equation, which has been studied quite often (e.g., Smith, Singh, and Sorensen, 1995). In most cases, there is no analytical solution for the quadratic matrix equation, but it can be solved by numerical methods such as the Newton method (*c.f.* Higham and Kim, 2001). After obtaining B , we can solve (39). Since A and D can be seen as the eigen value and eigen matrix of $(\bar{B} - B)$, we can first obtain D by calculating the eigen values of $(\bar{B} - B)$, and then normalizing the i^{th} eigen vector to make its i^{th} element equal to one. A is just the collection of those eigen vectors. After obtaining A and B , we can easily obtain Ω by equation (40). Note that there are, in total, $2n^2$ parameters in \bar{A} and \bar{B} , and also $2n^2$ parameters in A , B , and D . Thus, (38) and (39) provide a mapping from (\bar{A}, \bar{B}) to (A, B, D) . Also, it is easy to show that $R = \bar{R}$, and

$$L = -(A^{-1}B\bar{R} + A^{-1}\bar{L}). \quad (41)$$

Specifically, $\bar{\chi}_i = \sum_{k=1}^n (A^{-1}B)_{i,k} \bar{R}_k + \sum_{k=1}^n (A^{-1})_{i,k} (\bar{\theta})_k$, and $\omega_i(t) = \sum_{k=1}^n (A^{-1})_{i,k} \bar{\omega}_k(t)$.

C. Proof of Proposition 3

In the risk-neutral measure, the i^{th} futures prices $F_i(X, \tau)$ need to satisfy,

$$F_i(X, \tau) = E_t^Q[S_i(T)] \quad (42)$$

Thus, for the maximal model (18), $F_i(X, \tau)$ should satisfy the following vector-based Feynman-Kac equation:

$$-\frac{\partial F_i}{\partial \tau} + (U + \Psi X)^\top \frac{\partial F_i}{\partial X} + \frac{1}{2} \left(\frac{\partial^2 F_i}{\partial X^2} \Omega \Omega^\top \left(\frac{\partial F_i}{\partial X} \right)^\top \right)_{i,i} = 0 \quad (43)$$

with boundary condition $F_i(X, 0) = \exp(X_i)$.

Assume that

$$\ln(F_i(X, \tau)) = m_i(\tau) + G_i(\tau)^\top X, \quad (44)$$

where $m_i(\tau)$ is the i^{th} element of the $m(\tau)$ vector, and $G_i(\tau)$ is the i^{th} column of the $G(\tau)$ matrix.

By plugging (44) into (43), we have two ordinary differential equations

$$\begin{aligned} -\frac{\partial m_i(\tau)}{\partial \tau} + U^\top G_i + \frac{1}{2} (G_i(\tau)^\top \Omega \Omega^\top G_i(\tau))_{i,i} &= 0 \\ \frac{\partial G_i(\tau)}{\partial \tau} - \Psi^\top G_i(\tau) &= 0 \end{aligned} \quad (45)$$

with boundary condition

$$\begin{aligned} m_i(0) &= 0 \\ G_{i,i}(\tau) &= 1 \\ G_{j,i}(\tau) &= 0 \quad (i \neq j) \end{aligned}$$

Thus, the solution for (43) is

$$\begin{aligned}
 G(\tau) &= \exp(\Psi^\top \tau) \\
 m_i(\tau) &= \int_0^\tau \left[G_i(s)^\top U + \frac{1}{2} (G(s)^\top \Omega^\top \Omega G(s))_{i,i} ds \right]
 \end{aligned}$$

$G_i(\tau)$ denotes the i^{th} column of the $G(\tau)$ matrix. When Ψ is diagnosable,

$$G(\tau) = \Xi \text{diag}(\exp(\lambda_1 \tau), \dots, \exp(\lambda_{2n} \tau)) \Xi^{-1}$$

where Ξ is the matrix composed of eigen vectors of Ψ^\top and λ_k ($k = 1, \dots, 2n$) are the eigen values of Ψ^\top ; otherwise $G(\tau)$ can be calculated by Taylor expansion, i.e., $G(\tau) = I + \frac{1}{2}(\Psi^\top \tau)^2 + \frac{1}{6}(\Psi^\top \tau)^3 \dots$