WORKING PAPER SERIES-2013

WORKING PAPER 2013-0010

Economic Linkages, Relative Scarcity, and Commodity Futures Returns

Jaime Casassus, Peng Liu, and Ke Tang

January, 2013

Economic Linkages, Relative Scarcity, and Commodity Futures Returns[*]

Jaime Casassus

Pontificia Universidad Catolica de Chile

Peng Liu

Cornell University

Ke Tang

Renmin University of China

Economic Linkages, Relative Scarcity, and Commodity Futures Returns

Abstract

This paper shows that economic linkages among commodities create a source of long-term correlation between futures returns. We extend the Theory of Storage to a multi-commodity level and find that the convenience yield of a commodity depends not only on its own scarcity level, but also on its relative scarcity with respect to other economically-related commodities. This result implies a positive feedback effect from one commodity to another that is necessary to replicate the upward-sloping correlation term structure of futures returns observed from the related commodities. Our empirical Multi-Commodity Feedback Affine model (MCFA) allows for a flexible correlation term structure and validates our theoretical prediction. An out-of-sample test using short-maturity crack spread options data shows that our model considerably reduces the pricing error generated by traditional models.

Keywords: commodity futures, economic linkages, relative scarcity, correlation term structure, convenience yields, spread options.

JEL Classification: C0, G12, G13, D51, D81, E2.

Commodity markets have experienced dramatic up-and-down movements recently within a relatively short period of time. For example, the spot price of West Texas Intermediate (WTI) crude oil rallied from almost $50 per barrel in January 2007 to $145 per barrel in July 2008, the highest level in history since it began trading on the New York Mercantile Exchange (NYMEX). Surprisingly, only five months later, the oil price dropped to nearly $30 per barrel. Other energy commodities, industrial metals, agricultural, and livestock commodities have all experienced similar patterns. Since Keynes (1923), many scholars have studied the stochastic behavior of individual commodities, however, relationships involving multiple commodities have received little attention in theoretical modeling and commodity-related contingent-claim pricing.[1]

This paper shows that for productive commodities, their price dynamics depend not only on their own characteristics (i.e., prices, inventories), but also on the fundamentals of other economically related commodities. We show that the economic linkage between two commodities implies a source of long-term correlation between the futures returns. Examples of such economic linkages include, but are not restricted to, the following cases:

Production Relationships: One commodity can be produced from another commodity when the former is the output of a production process that uses the latter as an input factor. For example, the petroleum refining process "cracks" crude oil into heating oil and gasoline. A similar production relationship can be found in the soybean complex, where soybeans can be crushed into soybean meal and soybean oil.

Substitution Relationships: This relationship exists when two commodities are substitutes in consumption. Crude oil and natural gas are commonly viewed as substitute goods in the industrial

---

[1]Some of the prominent studies that document the stylized facts of single commodities are Gibson and Schwartz (1990), Brennan (1991), Bessembinder, Coughenour, Seguin, and Smoller (1995), Schwartz (1997), Richter and Sørensen (2002), Casassus and Collin-Dufresne (2005), and Trolle and Schwartz (2009)). Among the few empirical studies that consider more than one commodity are Malliaris and Urrutia (1996), Girma and Paulson (1999), Ai, Chatrath, and Song (2006), Paschke and Prokopczuk (2007), and Cortazar, Milla, and Severino (2008).

and electric generation sectors. According to the EIA Manufacturing Energy Consumption Survey (Energy Information Administration 2002), approximately 18 percent of the natural gas usage can be switched to petroleum products and up to 20 percent of power generation capacity is dual-fired. Another example of almost perfect substitutes is the case of WTI crude oil in North America and Brent crude oil from the North Sea.

Complementary Relationships: This relationship exists when two commodities share a balanced supply or are complementary in either consumption or production. Consider, for example, the case of industrial metals such as, lead, tin, zinc, and copper, that find most applications in the form of alloys. The equilibrium assemblage of mineral phases (i.e., paragenesis) gives these metals a natural complementary relationship in supply. Unleaded gasoline and heating oil offer another example of a complementary relationship. Whenever crude oil is cracked to supply gasoline, heating oil is also produced as a by-product.

Our paper finds that the economic linkage among commodities connects their price dynamics. Temporary deviations from a long-term relation between commodity prices (because of supply and demand imbalances caused by macro-economic factors, inventory shocks, etc.) are corrected over the long run. In particular, we show with a theoretical model that the convenience yield of a commodity, a benefit realized for holding inventories of the asset, depends on its "relative scarcity" with respect to other related commodities. This result extends the traditional Theory of Storage of Kaldor (1939), Working (1948), Brennan (1958), and Telser (1958), which connects the convenience yield of a commodity with its own scarcity level, to a multi-commodity level.

To highlight the effect of the relative scarcity on the joint dynamics of two commodities, consider the production linkage between crude oil and heating oil. Heating oil futures and expected spot prices decrease with the relative scarcity only because of the economic linkage between both commodities. Indeed, heating oil relative scarcity measured as the ratio of heating-to-crude-oil prices, increases

the producer's demand for crude oil, which increases the production of heating oil. More production implies higher *expected* heating oil inventories and lower *expected* heating oil prices in the next period.

The relative scarcity of heating oil also affects the correlation between the futures returns of both commodities. To understand this, consider how an increase in the crude oil stocks and a consequent fall in the crude oil price affect both futures prices. On the one hand, the absence of arbitrage between spot and futures prices implies a decrease in crude oil futures prices, while on the other hand, more crude oil stocks increase the relative scarcity of heating oil and, therefore, implies a decrease in heating oil futures prices. This simple mechanism shows that the production relationship between these commodities is a source of correlation between their futures returns.

Figure 1 shows the correlation term structure of weekly futures returns for the heating-crude oil and the WTI-Brent crude oil pairs from 2007.04 to 2010.09.[2] These commodity pairs are related by a production relationship and a substitution relationship, respectively. The plot shows upward-sloping correlation term structures for both commodity pairs. The economic linkage ties the prices of the commodities, which translates into higher long-term correlations. Interestingly, traditional commodity pricing models, such as correlated versions of the Gibson and Schwartz (1990; hereafter GS) and the Casassus and Collin-Dufresne (2005; hereafter CCD) models, are unable to match this evidence.

We also propose a multi-commodity feedback affine model (MCFA), where the expected change of each commodity price depends on the deviation from a long-term equilibrium with other commodities, i.e., a feedback effect that links the price dynamics of the related commodities. Furthermore, the feedback effect makes two commodity prices comove more tightly in the long run and hence is substantial in causing the upward-sloping correlation term structure observed in the data. The MCFA model nests the correlated GS and CCD models. The empirical results evidence a positive

---

[2]Note that, only after 2007, the maturity of heating oil can be greater than 24 months.

and significance effect of the relative scarcity of heating oil on its convenience yield. The model also considers feedback effects in the commodity risk premiums which allows to disentangle the correlation of futures returns from the correlation of commodity returns in the time-series dimension. The results suggest that the observed co-movement in the time series of the commodities pairs is mainly driven by the effect of the relative scarcity on the convenience yields rather than by the feedback effects in the risk premiums.

Since the correlation structure is crucial in the valuation of commodity spread options, the option prices implied by the traditional models have strong biases. We find that, for long-maturity spread options, the prices implied by our model are lower than the ones predicted by the traditional models, because the higher long-term correlation reduces the volatility of the spread. We show that the opposite is true for short-maturity options. An out-of-sample test using short-maturity crack spread options data shows that our model reduces the negative bias present in traditional models.

The rest of the paper is organized as follows. Section 1 presents an economic model for the case of two commodities that have a production relationship and shows how the relative scarcity of one of them affects their price dynamics.[3] Section 2 develops the empirical MCFA model that considers the predictions of the economic model. Section 3 describes the estimation of the empirical model and shows the results. Section 4 presents the valuation of spread options with the MCFA model and shows an out-of-sample comparison of several pricing models. Section 5 concludes.

1 The Economic Model

Commodity prices link two interconnected markets: the cash (or futures) market and the inventory market. For many commodities, physical ownership of units of that asset offers a benefit that,

---

[3]Given the brevity of this paper, we do not present structural models for the substitution or complementary relationships; interested readers can find those models in Appendix B of the earlier version of this paper, which can be found on SSRN.

in terms of a rate, is called the "convenience yield" (see Brennan 1991 and Schwartz 1997). The convenience yield of a commodity is attributed to the benefit of protecting regular production from supply shocks or by taking advantage of a rise in demand and price without resorting to revising the production schedule.

The Theory of Storage of Kaldor (1939), Working (1948), and Telser (1958) predicts that the return on purchasing a commodity and selling it for delivery (using futures) equals the interest forgone less the convenience yield net of storage costs. This condition implies that futures prices are decreasing in convenience yields. A similar result applies for expected spot commodity prices. All other things being equal, the higher the convenience yield of a commodity, the lower the expected spot price and therefore, the lower the expected spot return of that asset.[4]

The traditional presentation of the Theory of Storage proposes that the marginal benefit for holding inventories increases with the absolute scarcity of a commodity (see Pindyck 2001 and Routledge, Seppi, and Spatt 2000). If we consider only the market for any *single* commodity and use the spot price as a proxy for scarcity, the statement indicates: (1) The convenience yield is an increasing function of the spot price and (2) there is a positive correlation between incremental changes in the spot price and the convenience yield. This paper extends the traditional Theory of Storage by considering a multi-commodity framework and proposes a positive relation between the convenience yield of a commodity and its relative scarcity with respect to other commodities with which it shares an economic relationship. Using the price ratio of two related commodities as a proxy for the relative scarcity of one of them, our paper proposes a new third prediction for the convenience yield of a commodity that is relatively scarce: (3) The convenience yield of a commodity is an increasing

---

[4]To see this note that the expected commodity spot price return is:

$$\mathbb{E}_t^{\mathbb{P}}[dS_{i,t}] = (r_t + \pi_{i,t} - \delta_{i,t})S_{i,t}dt \qquad (1)$$

where r_t is the instantaneous risk-free rate, $\pi_{i,t}$ is the spot price risk premium and $\delta_{i,t}$ is the convenience yield (see Schwartz 1997 and CCD).

function of the price-level ratio between it and other related commodities.[5] Several empirical studies support the first two predictions, which are derived from the traditional Theory of Storage. For example, CCD explicitly models the positive dependence of the convenience yield of a commodity on its spot price and the instantaneous positive correlation between the spot price and the convenience yield. However, the third prediction, which connects the convenience yield of a commodity with its relative scarcity has not been studied.

To demonstrate the importance of our prediction, let us consider the following two-period two-commodity example. Consider commodities that are in a long-term equilibrium production relationship: heating oil (a downstream product) and crude oil (an upstream product). Further assume that at time 0 the prices of heating and crude oil are $20 and $15, respectively, while at time 1 their prices move to $22 and $21, respectively. If we look only at the heating oil market, the Theory of Storage predicts that heating oil will have a higher convenience yield at time 1 than at time 0, since heating oil is more expensive in the second period. However, if we look at both markets, we observe that, at time 0, heating oil is *relatively scarce* compared with crude oil, because of the higher heating-to-crude oil price ratio.[6] Indeed, since heating oil is refined from crude oil (and not the other way around), a high price ratio (i.e., high production profit), indicates that the refining capability cannot satisfy the strong demand for heating oil. In contrast to the traditional prediction of the Theory of Storage, we propose that the heating oil convenience yield will be lower at time 1 than at time 0. Indeed, a higher price ratio at time 0 also implies an increase in the cracking of crude oil and, therefore, both an increase in expected heating oil stocks and a drop in the price of heating oil in the next period. This decrease in the expected heating oil price due to its current relative scarcity is consistent with an increase in today's convenience yield of heating oil. Note that the dependence of the convenience

---

[5]In the next subsection we show the formal relationship between the relative scarcity of a commodity with respect to another commodity and their spot price ratio. We also derive an equilibrium relationship between the relative scarcity and convenience yield of the commodities.

[6]In a recent study, Ahn and Kogan (2011) use the price difference between heating oil and crude oil to decompose oil shocks into demand and supply components.

yield of a certain commodity on its relative scarcity is an extension of the traditional Theory of Storage.

1.1 The Convenience Yield and Relative Scarcity of a Commodity

In what follows, we use a general equilibrium model to illustrate the mechanism that connects a commodity convenience yield to its relative scarcity with respect to a related commodity. We propose a model that extends the single-commodity equilibrium models of Routledge, Seppi, and Spatt (2000) and Casassus, Collin-Dufresne, and Routledge (2008).[7] We consider a production economy that has a capital sector (K_t) and two storable commodity sectors that share a long-run equilibrium relation. For simplicity, we again assume they are crude oil (with stocks denoted as $Q_{1,t}$) and heating oil (with stocks denoted as $Q_{2,t}$). There are infinite resources of crude oil, but to make them available, an investment ($I_{1,t}$) is needed. Heating oil is produced from crude oil with commodity input quantity of q_t and capital input of $I_{2,t}$. An infinitely-long-lived representative agent derives utility from consumption of the following goods: the two commodities plus the standard consumption good from the capital sector, which is used as the numeraire. The representative agent maximizes expected utility with respect to consumption of capital, crude oil, and heating oil ($C_{K,t}, C_{1,t}, C_{2,t}$, respectively), demand of crude oil for production purposes (q_t), and investments in the commodity sectors ($I_{1,t}$ and $I_{2,t}$):

$$\sup_{\{C_{K,t}, C_{1,t}, C_{2,t}, q_t, I_{1,t}, I_{2,t}\} \in \mathcal{A}} \mathbb{E}_0^{\mathbb{P}} \left[\int_0^\infty e^{-\theta t} u\left[C_{K,t}, C_{1,t}, C_{2,t}\right] dt \right] \tag{2}$$

where $\mathcal{A}$ is the set of admissible strategies. The utility function $u[C_K, C_1, C_2]$ satisfies the standard conditions. The optimization problem is subject to the following processes that describe the dynamics

---

[7]Our model is also similar in spirit to the cross-commodity model of Routledge, Seppi, and Spatt (2001).

of capital, crude oil, and heating oil stocks, respectively:

$$dK_t = (\alpha_K K_t - C_{K,t} - I_{1,t} - I_{2,t})dt + \sigma_K K_t dW_{K,t}^{\mathbb{P}} \tag{3}$$

$$dQ_{1,t} = (f_1[I_{1,t}; Q_{1,t}] - q_t - C_{1,t})dt \tag{4}$$

$$dQ_{2,t} = (f_2[I_{2,t}, q_t; Q_{2,t}] - C_{2,t})dt \tag{5}$$

where $f_1[I_1; Q_1]$ is the crude oil production rate and $f_2[I_2, q; Q_2]$ is the heating oil production rate and both are increasing and concave. As mentioned before, the traditional Theory of Storage proposes that agents benefit from commodity inventories. For simplicity, we include this in an ad-hoc way by assuming that the production functions depend positively on their own commodity stocks. In particular, the benefit for the agents manifests insofar as the marginal productivity of the factors is higher when the commodity stock is higher, i.e., $\frac{\partial^2 f_i}{\partial I_i \partial Q_i} > 0$ and $\frac{\partial^2 f_2}{\partial q \partial Q_2} > 0$.[8] This could be thought of as a reduction in adjustment costs or economies of scale. We also assume that capital investment and crude oil are complementary inputs for the production of heating oil, i.e., $\frac{\partial^2 f_2}{\partial I_2 \partial q} > 0$.

Both commodities are consumption goods, hence, equilibrium commodity prices are the marginal rate of substitution of each commodity for the numeraire, i.e., $S_{i,t} = \frac{u_{i,t}}{u_{K,t}}$ with $u_{i,t} \equiv \frac{\partial u_t}{\partial C_{i,t}}$. To validate the heating-to-crude oil price ratio, $\frac{S_{2,t}}{S_{1,t}} = \frac{u_{2,t}}{u_{1,t}}$, as a measure for the relative scarcity of heating oil, $\frac{Q_{1,t}}{Q_{2,t}}$, we show that the marginal utility price of each commodity, $u_{i,t}$, is decreasing in its own stocks. Let $J_t \equiv J[K_t, Q_{1,t}, Q_{2,t}, t]$ be the indirect utility function for the optimization problem in equations (2)-(5). Since J_t inherits the concavity of the utility function $u[.]$ (see for example, Benveniste and Scheinkman 1979), the envelope condition for the consumption of commodity i (i.e., $e^{-\theta t} u_{i,t} = \frac{\partial J_t}{\partial Q_{i,t}}$) implies that a decrease in the stocks $Q_{i,t}$ increases the marginal utility $u_{i,t}$.

The following proposition shows the main predictions of the equilibrium model:

---

[8] Recall that the objective of our model is to inspect the mechanism that connects the convenience yield of a certain commodity with its relative scarcity with respect to other related commodities. This ad hoc assumption does not affect our cross-commodity results and greatly facilitates the solution.

Proposition 1 *The convenience yield of crude oil and heating oil are:*

$$\delta_{i,t} = \frac{\partial f_{i,t}}{\partial Q_{i,t}} \quad for \quad i = 1, 2 \tag{6}$$

The fact that crude oil is used to produce heating oil (and not the other way around) implies that relative scarcity affects the heating oil convenience yield, but not the crude oil convenience yield:

$$\frac{\partial \delta_{2,t}}{\partial \left(\frac{S_{2,t}}{S_{1,t}}\right)} > 0 \quad and \quad \frac{\partial \delta_{1,t}}{\partial \left(\frac{S_{2,t}}{S_{1,t}}\right)} = 0 \tag{7}$$

Proof See Appendix A.1. □

The first part of the proposition shows that the convenience yield of a commodity, $\delta_{i,t}$, is the marginal productivity of that commodity in its own sector. This result arises because this rate corresponds to the marginal benefit for storing the commodity and postponing its consumption until the next period.[9] This is consistent with the notion that the convenience yield is a benefit for holding inventories of the commodity.

Proposition 1 also shows that an increase in today's relative scarcity of heating oil increases the convenience yield of this commodity. The mechanism that generates this relation is the production linkage between the commodities. To see this consider the first-order condition for the representative agent's problem with respect to crude oil demand, q_t:

$$S_{1,t} = S_{2,t} \frac{\partial f_{2,t}}{\partial q_t} \quad or \quad \frac{\partial f_{2,t}}{\partial q_t} = \left(\frac{S_{2,t}}{S_{1,t}}\right)^{-1} \tag{8}$$

Since the production functions are concave, equation (8) implies that an increase in the relative scarcity of heating oil increases the demand for crude oil, q_t, and, thus, the production of heating

---

[9]Indeed, the fact that the representative agent can freely consume the commodity or store it and get the net convenience yield for holding inventories, determines an intertemporal Euler equation that relates today's marginal utility of consumption (i.e., marginal cost of postponing consumption of the commodity) with the convenience yield and tomorrow's marginal utility of consumption (i.e., marginal benefit of consuming the commodity in the next period).

oil.[10] The assumption $\frac{\partial^2 f_2}{\partial q \partial Q_2} > 0$ and the definition of the convenience yield in equation (6) imply that an increase in q_t increases the heating oil convenience yield. The intuition behind this result is that a fraction of the new production of heating oil will be stored to take advantage of the higher marginal productivity of the factors and the rest will be consumed. This higher expected heating oil stocks decreases the expected heating oil price implying a higher heating oil convenience yield. The effect of relative scarcity in the dynamics of the heating oil price confirms the notion that heating oil producers determine their production schedule based not only on the price of heating oil but also on the crude oil market. On the other hand, changes in the stocks of heating oil, $S_{2,t}$ and, therefore, changes in the relative scarcity, does not affect the convenience yield of crude oil. This prediction of the model implies that relative scarcity links the dynamics of both crude oil and heating oil prices only through the convenience yield of heating oil. In the empirical section of the paper we test this hypothesis and use the results to validate our model. The price dynamics of both commodities are linked closely together because of the production relationship between them. It is worth noting that the previous results are valid for any utility function, since the convenience yields are related to the production technologies and not to the preferences of the representative agent.[11]

Relative scarcity could also affect the expected commodity returns through the risk premium of the commodity prices. Unfortunately, the current model is too general to shed some light on this direction. The commodity risk premiums depend on the covariance between the prices and the marginal utility $u_{K,t}$, and to obtain them we need to actually solve the model in an explicit way. The magnitude of the risk premia and their possible relation to prices and relative scarcity will depend on the preferences of the representative agent. We leave this relationship to the empirical model where we allow the risk premia to vary over time.

---

[10]Since crude oil and investment are complementary inputs for the production of heating oil, an increase in the demand for crude oil also implies a higher investment in the heating oil sector. The complementary assumption in the inputs, although not necessary for our main prediction, increases the effect of the relative scarcity in the production of heating oil.

[11]Of course, commodity prices and relative scarcity depend on the representative agent's utility, but the results in Proposition 1 remain the same for any preference specification.

1.2 An Example of a Cobb-Douglas Economy

In this section we assume a standard Cobb-Douglas economy to better understand the relation between the convenience yield of heating oil and the relative scarcity of this commodity. The production functions are:

$$f_1[I_1; Q_1] = \alpha_1(Q_1)I_1^{\beta_1} \tag{9}$$

$$f_2[I_2, q; Q_2] = \alpha_2(Q_2)I_2^{\beta_2}q^{\gamma} \tag{10}$$

in which $0 < \beta_i, \gamma < 1$ and the total productivity factors, $\alpha_i(Q_i)$, are positive, increasing and concave functions in Q_i. Here, $\alpha_i' = \frac{\partial \alpha_i}{\partial Q_i} > 0$ represents the exogenous incentive to hold commodity stocks in line with the prediction of the Theory of Storage.

The next proposition shows the equilibrium convenience yields for the Cobb-Douglas production technologies.

Proposition 2 *In a Cobb-Douglas economy, the convenience yields for crude oil and heating oil are*

$$\delta_{1,t} = \alpha_{1,t}' I_{1,t}^{\beta_1} \tag{11}$$

$$\delta_{2,t} = \alpha_{2,t}' I_{2,t}^{\beta_2} q_t^{\gamma} \tag{12}$$

respectively. We can also express the convenience yields in terms of the commodity prices and, for the case of the heating oil, in terms also of its relative scarcity:

$$\delta_{1,t} = \alpha_{1,t}' (\alpha_{1,t}\beta_1 S_{1,t})^{\frac{\beta_1}{1-\beta_1}} \tag{13}$$

$$\delta_{2,t} = \alpha_{2,t}' (\alpha_{2,t}\beta_2 S_{2,t})^{\frac{\beta_2}{1-\beta_2-\gamma}} \left(\alpha_{2,t}\gamma \frac{S_{2,t}}{S_{1,t}} \right)^{\frac{\gamma}{1-\beta_2-\gamma}} \tag{14}$$

Proof See Appendix A.2. $\square$

As expected, equation (12) shows that the heating oil convenience yield increases with the demand

11

for crude oil, q_t. Since this variable depends on the relative scarcity, equation (14) confirms that the convenience yield of heating oil depends on the ratio $\frac{S_{2,t}}{S_{1,t}}$. The impact of the relative scarcity on this convenience yield increases with the input share of oil for the production of heating oil, γ. The feedback effect of the crude oil price $S_{1,t}$ on the heating oil convenience yield $\delta_{2,t}$ is present as long as $\gamma > 0$. Indeed, if this parameter is zero, only absolute scarcity, measured by the spot price $S_{2,t}$, matters.

Also, equations (11) and (12) show that the convenience yields increase with the investment rates. This implies that the convenience yields $\delta_{i,t}$ increase with the spot prices $S_{i,t}$, and therefore, absolute scarcity (see equations (13) and (14)). The strength of these effects depend on the elasticities β_i. The convenience yields also depend on the stocks $Q_{i,t}$ through the productivity factors $\alpha_i(Q_i)$. Since we assume that these functions are concave, the correlation between the convenience yields and spot price for the same commodity is positive, a prediction that it is line with the Theory of Storage.[12]

1.3 Implications for the Correlation of Futures Returns

Our main equilibrium result from the previous section is that the heating oil convenience yield increases with its relative scarcity (see Proposition 1). This implies that a decrease in the crude oil price increases the relative scarcity of heating oil, which in turn increases the heating oil convenience yield and thus decreases expected heating oil prices (see equations (1) and (14)). This mechanism, which is present only because of the production linkage between the two commodities, creates a positive feedback between current crude oil prices and expected heating oil prices. Moreover, this mechanism also affects the correlations of futures returns in a positive way. Indeed, the absence of arbitrage implies that a decrease in the crude oil spot price implies a negative change in crude oil futures prices. On the other hand, the increase in relative scarcity of heating oil implies a

---

[12]To see this positive correlation, note that a decrease in the commodity stock will imply an increase in both the convenience yield and the spot price for that asset.

higher convenience yield for this asset, which has a negative effect on heating oil futures prices. In other words, our model predicts that a crude oil price shock impacts both crude oil and heating oil futures prices in the same direction; thus, the economic linkage between these commodities creates a new source of positive correlation of futures returns across commodities. We will show later in an affine reduced-form model that the positive feedback effect from crude oil to heating oil implies an upward-sloping correlation term structure of futures returns. In the next section we present a multi-commodity affine model and use it, among other things, to prove the existence of the positive feedback effect for the crude and heating oil pair.

2   The Empirical Model

Guided by the economic model presented in Section 1, we develop a reduced-form model that is consistent with the stylized facts about economically related commodities (i.e., upward-sloping correlation term structure, stochastic convenience yields, mean-reversion, etc.). Furthermore, we distinguish two sources of co-movement across commodities: 1) a short-term effect associated with the correlation of instantaneous changes in commodity prices, and 2) a long-term feedback effect that is a consequence of a multi-commodity equilibrium economic relationship. The feedback effect manifests insofar as the dynamics of one commodity being related to the prices of the other commodities in the economy. In particular, we allow the convenience yield of a commodity to depend on its relative scarcity with respect to the other commodities in the economy. For simplicity, we consider an *affine* relationship among the convenience yields and the risk factors, and hence name the empirical model as multi-commodity feedback affine model (MCFA).

2.1 The Multi-Commodity Feedback Affine Model (MCFA)

Following CCD, we first describe the processes of the MCFA model under the risk-neutral measure ($\mathbb{Q}$) and then specify the risk premium to determine the processes under the physical measure ($\mathbb{P}$).

Assume a system of n commodities that share an economic linkage. Denote

$$x_{i,t} = \log(S_{i,t}) \quad \text{for} \quad i = 1, \ldots, n \tag{15}$$

where $S_{i,t}$ is the spot price of commodity i. The log spot price price follows a Gaussian process

$$dx_{i,t} = (\mu_i^{\mathbb{Q}} - \delta_{i,t})dt + \sigma_i dW_{i,t}^{\mathbb{Q}} \quad \text{for} \quad i = 1, \ldots, n \tag{16}$$

where $\delta_{i,t}$ is the convenience yield of commodity i, and σ_i is the instantaneous volatility of the commodity returns. Absence of arbitrage implies that $\mu_i^{\mathbb{Q}} = r - \frac{1}{2}\sigma_i^2$ where r is the risk-free interest rate which is assumed to be constant.[13] We also assume that $W_{i,t}^{\mathbb{Q}}$ ($i = 1, \ldots, n$) are correlated Brownian motions.

Motivated by our theoretical framework, the convenience yield of commodity i, $\delta_{i,t}$, is a function of its spot price (as in CCD) and its relative scarcity with respect to the other commodities in the economy. To keep our empirical model affine, we use the log of the price ratio (instead of the price ratio itself) to represent the relative scarcity. Specifically, we use the log-price difference ($x_{i,t} - x_{j,t}$) to represent the relative scarcity of the ith commodity with respect to the jth commodity. We also consider n extra latent factors, $\eta_{i,t}$ ($i = 1, \ldots, n$), affecting the n convenience yields. Therefore,

$$\delta_{i,t} = h_i^{\mathbb{Q}} x_{i,t} + \sum_{j=1,j\neq i}^{n} c_{i,j}^{\mathbb{Q}}(x_{i,t} - x_{j,t}) + \eta_{i,t} - \sum_{j=1,i\neq j}^{n} a_{i,j}\eta_{j,t} \tag{17}$$

where $c_{i,j}^{\mathbb{Q}}$ represents the extent to which the convenience yield of commodity i depends on the relative scarcities with other commodities, and $h_i^{\mathbb{Q}}$ and $a_{i,j}$ are constants. If setting $b_{i,i}^{\mathbb{Q}} \equiv -\left(h_i^{\mathbb{Q}} + \sum_{j=1,j\neq i}^{n} c_{i,j}^{\mathbb{Q}}\right)$

---

[13]It is straightforward to extend our model to consider stochastic interest rates as in Schwartz (1997).

and $b^{\mathbb{Q}}_{i,j\neq i} \equiv c^{\mathbb{Q}}_{i,j}$, equation (17) thus can be rewritten as

$$\delta_{i,t} = -\sum_{j=1}^{n} b^{\mathbb{Q}}_{i,j} x_{j,t} + \eta_{i,t} - \sum_{j=1,i\neq j}^{n} a_{i,j} \eta_{j,t} \tag{18}$$

The latent factors $\eta_{i,t}$ $(i = 1, \ldots, n)$ follow mean-reverting processes under the risk-neutral measure

$$d\eta_{i,t} = (\chi^{\mathbb{Q}}_i + \omega_i(t) - \kappa_i \eta_{i,t})dt + \sigma_{n+i} dW^{\mathbb{Q}}_{n+i,t} \quad \text{for} \quad i = 1, \ldots, n \tag{19}$$

where $\chi^{\mathbb{Q}}_i$ is a constant and $\omega_i(t)$ is a periodical function on t to capture the seasonality of commodity futures prices (if any). Refer to Richter and Sørensen (2002) and Geman and Nguyen (2005) for a similar setup on the seasonality of the convenience yields. Following Harvey (1991) and Durbin and Koopman (2001), we specify $\omega_i(t)$ as:

$$\omega_i(t) = \sum_{l=1}^{L} \left(s^{c,l}_i \cos(2\pi\, l\, t) + s^{s,l}_i \sin(2\pi\, l\, t) \right) \tag{20}$$

Letting $Y_t = (x_{1,t}, \ldots, x_{n,t}, \eta_{1,t}, \ldots, \eta_{n,t})'$ denote the $2n$ factors driving the system of n commodity prices, the MCFA model can be rewritten in a vector form,

$$dY_t = (U^{\mathbb{Q}}(t) + \Psi^{\mathbb{Q}} Y_t)dt + dZ^{\mathbb{Q}}_t \tag{21}$$

where $U^{\mathbb{Q}}(t) = (\mu^{\mathbb{Q}}_1, \ldots, \mu^{\mathbb{Q}}_n, \chi^{\mathbb{Q}}_1 + \omega_1(t), \ldots, \chi^{\mathbb{Q}}_n + \omega_n(t))'$ and $\Psi^{\mathbb{Q}} = \begin{pmatrix} B^{\mathbb{Q}} & A \\ 0 & \mathcal{K} \end{pmatrix}$ with

$$B^{\mathbb{Q}} = \begin{pmatrix} b^{\mathbb{Q}}_{1,1} & b^{\mathbb{Q}}_{1,2} & \cdots & b^{\mathbb{Q}}_{1,n} \\ b^{\mathbb{Q}}_{2,1} & b^{\mathbb{Q}}_{2,2} & \ddots & b^{\mathbb{Q}}_{2,n} \\ \vdots & \ddots & \ddots & \vdots \\ b^{\mathbb{Q}}_{n,1} & b^{\mathbb{Q}}_{n,2} & \cdots & b^{\mathbb{Q}}_{n,n} \end{pmatrix}, \; A = \begin{pmatrix} -1 & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & -1 & \ddots & a_{2,n} \\ \vdots & \ddots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & -1 \end{pmatrix}, \; \mathcal{K} = \begin{pmatrix} -\kappa_1 & 0 & \cdots & 0 \\ 0 & -\kappa_2 & \ddots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & -\kappa_n \end{pmatrix}$$

In equation (21), $Z^{\mathbb{Q}}_t = \left(\sigma_1 W^{\mathbb{Q}}_{1,t}, \ldots, \sigma_{2n} W^{\mathbb{Q}}_{2n,t} \right)'$ is a scaled Brownian motion vector with covariance matrix $\Omega = \{\rho_{i,j}\sigma_i\sigma_j\}$ for $i,j = 1, \ldots, 2n$, where $\rho_{i,j}dt$ is the instantaneous correlation between the

Brownian motion increments $dW_{i,t}^{\mathbb{Q}}$ and $dW_{j,t}^{\mathbb{Q}}$.

We assume that the risk premium of the $x_{i,t}$ factor depends not only on itself but also on other associated commodity prices. Note that by making the risk premium of the $x_{i,t}$ factor depend on itself, the CCD model can capture the mean-reversion difference of the physical and risk-neutral measures. In our paper, by assuming that the $x_{i,t}$ factor depends on the relative scarcity, we are able to see the difference caused by the feedback effect under the risk-neutral and physical measures.[14] Also, we assume a constant risk premium for the η factors. We define the constants $\mu_i^{\mathbb{P}}$, $b_{i,j}^{\mathbb{P}}$ and $\chi_i^{\mathbb{P}}$ such that the physical processes can be expressed as follows:

$$dZ_t^{\mathbb{P}} = dZ_t^{\mathbb{Q}} - \Pi_t dt \tag{22}$$

where Π_t is a risk-premium vector with its ith element $\Pi_{i,t}$ specified as[15]

$$\Pi_{i,t} = \begin{cases} \mu_i^{\mathbb{P}} - \mu_i^{\mathbb{Q}} + \Sigma_{j=1}^n(b_{i,j}^{\mathbb{P}} - b_{i,j}^{\mathbb{Q}})x_{j,t} & \text{for} \quad 1 \leq i \leq n \\ \chi_{i-n}^{\mathbb{P}} - \chi_{i-n}^{\mathbb{Q}} & \text{for} \quad n+1 \leq i \leq 2n \end{cases}$$

Thus, under the physical measure, the stochastic behavior of the factors can be expressed as

$$dY_t = (U^{\mathbb{P}}(t) + \Psi^{\mathbb{P}}Y_t)dt + dZ_t^{\mathbb{P}} \tag{23}$$

where $Z_t^{\mathbb{P}} = (\sigma_1 W_{1,t}^{\mathbb{P}}, \ldots, \sigma_{2n} W_{2n,t}^{\mathbb{P}})'$, $U^{\mathbb{P}}(t) = (\mu_1^{\mathbb{P}}, \ldots, \mu_n^{\mathbb{P}}, \chi_1^{\mathbb{P}} + \omega_1(t), \ldots, \chi_n^{\mathbb{P}} + \omega_n(t))'$ and $\Psi^{\mathbb{P}} = $

$$\begin{pmatrix} B^{\mathbb{P}} & A \\ 0 & \mathcal{K} \end{pmatrix} \text{ with } B^{\mathbb{P}} = \begin{pmatrix} b_{1,1}^{\mathbb{P}} & b_{1,2}^{\mathbb{P}} & \cdots & b_{1,n}^{\mathbb{P}} \\ b_{2,1}^{\mathbb{P}} & b_{2,2}^{\mathbb{P}} & \ddots & b_{2,n}^{\mathbb{P}} \\ \vdots & \ddots & \ddots & \vdots \\ b_{n,1}^{\mathbb{P}} & b_{n,2}^{\mathbb{P}} & \cdots & b_{n,n}^{\mathbb{P}} \end{pmatrix}.$$

The non-diagonal terms in the $B^{\mathbb{Q}}$, $B^{\mathbb{P}}$ and A matrices connect the dynamics of the n commodity

---

[14]We thank our editor Pietro Veronesi for pointing this out.

[15] The risk-premium vector Π_t can also be expressed in terms of the same commodity log price $x_{i,t}$, and the relative scarcities $(x_{i,t} - x_{j,t})$:

$$\Pi_{i,t} = \begin{cases} \mu_i^{\mathbb{P}} - \mu_i^{\mathbb{Q}} + \Sigma_{j=1}^n(b_{i,j}^{\mathbb{P}} - b_{i,j}^{\mathbb{Q}})x_{i,t} + \Sigma_{j \neq i}^n(b_{i,j}^{\mathbb{Q}} - b_{i,j}^{\mathbb{P}})(x_{i,t} - x_{j,t}) & \text{for} \quad 1 \leq i \leq n \\ \chi_{i-n}^{\mathbb{P}} - \chi_{i-n}^{\mathbb{Q}} & \text{for} \quad n+1 \leq i \leq 2n \end{cases}$$

prices. The MCFA model presented above nests several classical models. In particular, it nests the correlated GS and CCD models that correspond to the GS and CCD models when the spot prices and convenience yields across commodities have correlated shocks (i.e., are instantaneously correlated). Indeed, if $b_{i,j}^{\mathbb{Q}} = b_{i,j}^{\mathbb{P}} = 0$ and $a_{i,j \neq i} = 0$, the MCFA model reduces to correlated GS models on commodities. Also, if $b_{i,j \neq i}^{\mathbb{Q}} = b_{i,j \neq i}^{\mathbb{P}} = 0$ and $a_{i,j \neq i} = 0$, the MCFA model reduces to correlated CCD models with a constant interest rate. The correlated versions of the these classical models are more flexible than the original ones and later will be considered as benchmarks for the MCFA model.

2.2 Correlation Term Structure of Futures Returns

This section first presents the closed-form solution for the futures prices in the MCFA model and then demonstrates that the correlation between the futures returns of two commodities depends on the maturity of the contracts.

The futures price $F_{i,t}(Y_t, T)$ at time t for the purchase of one unit of commodity at time T is the expected future spot price under the risk-neutral measure (e.g., Duffie 2001). Using standard results on pricing within the affine framework (e.g., Duffie and Kan 1996), we obtain the following closed-form expression for the futures prices:

Proposition 3 *Let $F_{i,t}(Y_t, T)$ be the ith commodity futures price maturing at time T. In the MCFA model, the futures prices are determined by*

$$\log(F_{i,t}(Y_t, T)) = m_i(T - t) + G_i(T - t)Y_t \quad for \quad i = 1, \ldots, n \tag{24}$$

where $m_i(\tau) = \int_0^\tau \left(G_i(u)U^{\mathbb{Q}} + \frac{1}{2} G_i(u)\Omega\, G_i(u)' \right) du$ and $G_i(\tau)$ denotes the ith row of $G(\tau) \equiv \exp(\Psi^{\mathbb{Q}}\tau)$.

Proof See Appendix B.1. $\quad\square$

Let us define $r_{i,t,t_1}(T)$ as the log futures return between t and t_1 on a contract that expires at time T

$$r_{i,t,t_1}(T) = \log\left(\frac{F_{i,t_1}(Y_{t_1},T)}{F_{i,t}(Y_t,T)}\right) \tag{25}$$

and $\Sigma_{t,t_1}(T)$ as the covariance of these futures returns

$$\Sigma_{t,t_1}(T) = \mathbb{E}_t^{\mathbb{P}}\left[\left(r_{t,t_1}(T) - \mathbb{E}_t^{\mathbb{P}}[r_{t,t_1}(T)]\right)\left(r_{t,t_1}(T) - \mathbb{E}_t^{\mathbb{P}}[r_{t,t_1}(T)]\right)'\right] \tag{26}$$

Here, $r_{t,t_1}(T)$ is the 1×n vector of commodity futures returns from equation (25). The next proposition presents closed-form expressions for the covariance matrix and for the correlation term structures of futures returns.

Proposition 4 *The covariance matrix of the futures returns in the MCFA model is:*

$$\Sigma_{t,t_1}(T) = e^{\Psi^{\mathbb{Q}}(T-t_1)}\mathbb{V}ar_t^{\mathbb{P}}[Y_{t_1}]e^{\Psi^{\mathbb{Q}'}(T-t_1)} \tag{27}$$

where $\mathbb{V}ar_t^{\mathbb{P}}[Y_{t_1}] = \int_t^{t_1} e^{\Psi^{\mathbb{P}}(t_1-v)}\Omega\, e^{(\Psi^{\mathbb{P}})'(t_1-v)}dv$ is the conditional covariance of the state variables Y_t under the physical measure. Moreover, the instantaneous covariance of futures returns is

$$\Sigma_t(T) = \lim_{t_1 \to t+dt}\Sigma_{t,t_1}(T) = e^{\Psi^{\mathbb{Q}}(T-t)}\Omega\, e^{\Psi^{\mathbb{Q}'}(T-t)}dt \tag{28}$$

Finally, the instantaneous correlation term structure between futures returns of commodities i and j is defined as

$$\rho_{i,j,t}(T) = \frac{\Sigma_t(T)_{i,j}}{\sqrt{\Sigma_t(T)_{i,i}\Sigma_t(T)_{j,j}}} \quad for \quad i,j = 1,\ldots,n \tag{29}$$

Proof See Appendix B.2. $\square$

Equation (27) shows that not only the long-run matrix $\Psi^{\mathbb{Q}}$ but also $\Psi^{\mathbb{P}}$ (hence the risk premium) influence the correlation term structure of futures returns. From (28), we see that the instantaneous covariance of futures returns depends on the matrix $\Psi^{\mathbb{Q}}$, as does the instantaneous correlation of

18

futures returns. The longer the maturity of the futures, the stronger the role of $\Psi^{\mathbb{Q}}$ in the instantaneous correlation of futures returns. Intuitively, the elements $b_{i,j}^{\mathbb{Q}}$ in the matrix $\Psi^{\mathbb{Q}}$ have a significant influence on the *expected* spot price co-movement in the risk-neutral measure, since futures prices are *expected* spot prices under the risk-neutral measure. Specifically, in the risk-neutral measure the expected change in $x_{i,t}$ is

$$\mathbb{E}_t^{\mathbb{Q}}[dx_{i,t}] = \left(r - \frac{1}{2}\sigma_i^2 + \sum_{j=1}^{n} b_{i,j}^{\mathbb{Q}} x_{j,t} - \eta_{i,t} + \sum_{j=1, i\neq j}^{n} a_{i,j}\eta_{j,t} \right) dt \tag{30}$$

where the $b_{i,j}^{\mathbb{Q}}$'s (for $j \neq i$) relate the expected return of commodity i with the price and convenience yield of commodity j. Thus, these parameters represent the long-term source of co-movement. The correlated GS and CCD models set these parameters to zero; therefore, they completely ignore the cross-commodity feedback effect between distinct commodities.

We classify the co-movement between (log) commodity prices $x_{i,t}$ and $x_{j,t}$ $(j \neq i)$ into four classes according to the sign of the feedbacks, $b_{i,j}^{\mathbb{Q}}$. If both $b_{i,j}^{\mathbb{Q}} \geq 0$ and $b_{j,i}^{\mathbb{Q}} \geq 0$, a positive increment of $x_{i,t}$ tends to feed a positive increment back on $x_{i,t}$, which is in turn likely to strengthen $x_{i,t}$ by another positive feedback; hence $x_{i,t}$ and $x_{j,t}$ move together. Note that the positive feedback effect strengthens the co-movement of two commodities in addition to the correlation of the increments of the commodity prices. As shown in Engle and Granger (1987), this effect will become more influential with a longer time horizon. Similarly, if only one feedback is positive, say $b_{i,j}^{\mathbb{Q}} > 0$ and $b_{j,i}^{\mathbb{Q}} = 0$, then $x_{i,t}$ follows the dynamics of $x_{j,t}$ generating also long-term correlation between the commodities.[16] In an opposing manner, if $b_{i,j}^{\mathbb{Q}} \leq 0$ and $b_{j,i}^{\mathbb{Q}} \leq 0$, $x_{i,t}$ and $x_{j,t}$ move in opposite directions. Lastly, we have the other cases with where it is not easy to identify the co-movement between commodity prices by type.

In general, if there is a long-term economic relationship, it will appear in the feedbacks, which in

---

[16]If these analysis were made under the physical measure, we would say that $x_{j,t}$ Granger-causes $x_{i,t}$ as in the forecasting literature.

turn affects the long-run matrix $\Psi^{\mathbb{Q}}$. Therefore, the empirical model presented in this paper makes an important contribution regarding the long-term co-movement between distinct commodities. This long-term source of co-movement is a feedback effect that occurs mainly through the connection between the expected prices of distinct commodities. Note that this cross-commodity feedback effect corresponds to an error correction or cointegration between separate time series in the discrete-time econometric literature.

Figures 6, 8 and 10 demonstrate the term structures of the futures return correlations between distinct commodities. These plots show that the cross-commodity feedback effect due to the economic relationship plays an important role in explaining the co-movement of commodity prices. By neglecting the cross-commodity feedback parameters, the GS and CCD models impose strong restrictions on the pricing structure. Therefore, the cross-commodity feedback effect is important for matching the upward-sloping correlation structure in the data.

A similar analysis to the one made under the risk-neutral measure can be used to understand the impact of the the feedbacks $b_{i,j}^{\mathbb{P}}$ (for $j \neq i$) on the co-movement of prices in the time-series dimension. Our MCFA model is able to distinguish the feedback effect of the convenience yield from that of the risk premium in this type of co-movement. Indeed, the feedback parameters $b_{i,j}^{\mathbb{P}}$ can be decomposed into two terms: 1) $b_{i,j}^{\mathbb{Q}}$, which is associated to the convenience yield, and (2) $b_{i,j}^{\mathbb{P}} - b_{i,j}^{\mathbb{Q}}$, which is associated to the time-varying risk premium. The magnitude of these two terms is critical in understanding which component dominates in explaining the observed multi-commodity co-movement in the time series.

3 Estimation

We demonstrate the importance of long-term economic relationships in futures pricing using the heating oil and crude oil production pair from Section 1. Even though our model can be applied to price a system of n commodities jointly, two commodities are enough to highlight the main characteristics of our model and the intuition behind the results.[17] We also estimate the model for two commodities that are substitute goods (WTI crude oil and Brent crude oil) and for two commodities that are complementary goods (heating oil and gasoline).

3.1 The Data

Our data consist of weekly futures prices of three pairs of commodities: 1) the WTI crude oil and heating oil pair, 2) the WTI and Brent crude oil pair, and 3) the heating oil and unleaded gasoline pair. The weekly futures in the above pairs are obtained through NYMEX and the London International Petroleum Exchange for the period running from January of 1995 to September of 2010 (821 observations for each commodity). Time to maturity ranges from 1 month to 17 months for these commodities. We denote Fm as futures contracts with roughly m months to maturity; e.g., F0 denotes the cash spot prices and F12 denotes the futures prices with 12 months to maturity. We use five time series, F1, F5, F9, F13, F17, for the WTI crude and heating oil pair; and F1, F3, F6, F9, F11 for the WTI and Brent crude oil pair and the heating oil and unleaded gasoline pair.[18] Figure 2, 3 and 4 show the historical prices of the three commodity pairs.

---

[17]The computational loads increase exponentially for the case of more than two commodities. Furthermore, commodity pairs are building blocks of any commodity system. Any multi-commodity system can be decomposed into multiple commodity pairs, e.g., the system with three commodities can be priced using no more than three pairs of commodities.

[18]We use 4% risk-free interest rate in the estimation, which is the average interest rates during the sample periods.

3.2 Empirical Examination of Cross-Commodity Economic Linkages

In this section, we examine three commodity pairs, one for each type of economic linkage: the WTI crude oil and heating oil pair (a production relationship), the WTI and Brent crude oil pair (a substitution relationship), and the heating oil and unleaded gasoline pair (a complementary relationship). We estimate our model with Maximum Likelihood Estimation and use the Kalman filter methodology to estimate the latent variables. For details, please refer to Appendix C.

WTI Crude and Heating Oil Pair

Since WTI crude oil and heating oil are the input and output of an oil refinery process, this commodity pair has a production relationship. Following the theoretical model in Section 1, we define crude oil as commodity 1 and heating oil as commodity 2. From observation of crude and heating oil prices, we find that crude oil prices do not exhibit seasonality, which is consistent with the literature on oil futures, such as Schwartz (1997). However, heating oil prices exhibit strong seasonality, which is consistent with Richter and Sørensen (2002). This occurs because demand for heating oil is typically high in the winter, but there are usually not enough available facilities in which to store the heating oil; hence, in the winter, heating oil has relatively higher convenience yield. Therefore, winter-maturing futures prices tend to be higher than are those maturing in summer. Since the seasonality of heating oil is in an annual frequency, by setting $L = 1$, equation (20) reduces to the following:

$$\omega_1(t) \equiv 0 \quad \text{and} \quad \omega_2(t) = s_2^c \cos(2\pi t) + s_2^s \sin(2\pi t) \tag{31}$$

From the model estimation showed in Table 1 panel A, we see that most parameters are significant and consistent with the literature on single commodity dynamics. For example, $b_{1,1}^{\mathbb{Q}}$ and $b_{2,2}^{\mathbb{Q}}$ are negative, indicating the two commodity futures prices are mean-reverting because of the positive

effect of the commodity prices in their convenience yields. Also, as expected heating oil displays a profound seasonality.

Our particular interest is in the feedback effects between crude oil and heating oil. The coefficient $b_{2,1}^{\mathbb{Q}} = 1.921$ is highly significant, which shows that the convenience yield of heating oil does depend *positively* on the relative scarcity of heating oil to crude oil. This result is consistent with the main prediction of Proposition 1. An increase in the relative scarcity of heating oil, increases the demand for crude oil and the production of heating oil. In expected terms, more production today implies more stocks and lower prices of heating oil tomorrow, which is consistent with a higher convenience yield today.[19] Our theoretical model also indicates that $b_{1,2}^{\mathbb{Q}}$ should be zero; however, although $b_{1,2}^{\mathbb{Q}} = 0.609$ is significant, its magnitude is much smaller than $b_{2,1}^{\mathbb{Q}}$ (about one-fourth). The fact that the feedback effect from crude oil to heating oil is much stronger than the one in the other direction, gives empirical support to the theoretical model.

Figure 5 shows the convenience yields for both WTI crude oil and heating oil that are implied by the MCFA model. Figure 6 shows the correlation structure for the correlated GS model, the correlated CCD model, and the MCFA model. We bootstrap the model parameters by assuming that each parameter estimate has a normal distribution and obtain the 95% confidence level of the correlation. The plot shows that, for the MCFA model, the correlation curve is upward sloping and the errors of correlations diminish when the correlation approaches one. However, for the GS and CCD models, the correlations begin to go down when the futures time to maturity is longer than two years; in the meantime the error of the correlation also becomes larger. Hence, only the MCFA model is able to generate the upward-sloping correlation curve present in the data—mainly because of the significant *positive* value of $b_{1,2}^{\mathbb{Q}}$ and $b_{2,1}^{\mathbb{Q}}$, which, as mentioned in Section 2.2, links the two commodities by a positive feedback effect. It is easy to understand that the positive value of $b_{1,2}^{\mathbb{Q}}$

---

[19]Although lower expected prices could be explained also by a lower risk premium, our theoretical model shows that this decrease occurs because of the convenience yield. Moreover, our empirical model also suggests that the impact of the relative scarcity to expected prices are mainly through the convenience yield rather than through the risk premium.

and $b_{2,1}^{\mathbb{Q}}$ has a significant influence on the *expected* future spot prices in the risk-neutral measure; and, since futures prices are the *expected* spot prices in the risk-neutral measure, $b_{1,2}^{\mathbb{Q}}$ and $b_{2,1}^{\mathbb{Q}}$ both play an important role in determining the correlation term structure of futures returns on the crude and heating oil pair (see also equation (28)). Furthermore, the longer the futures time to maturity, the stronger the role of these positive feedbacks will play and hence the larger the correlation. These feedback effects thus result in an upward-sloping correlation term structure in the MCFA model.

In the short run, we see that the correlation is smaller in the MCFA model than it is in the correlated GS and CCD models. This occurs because the MCFA model is more flexible when capturing the co-movement between two futures prices, which allows us to disentangle the various sources of co-movement (i.e., the instantaneous correlation and the long-term feedback effects).[20] Indeed, the correlated versions of the GS and CCD models, which do not consider economic linkages, are forced to include some existing mid-term correlations in the short-term component of co-movement. In the long run, the MCFA model allows for a greater correlation than the other two models do, which is consistent with the significance of the cross-commodity relationship.

In order to test whether the MCFA model is better than the correlated versions of the GS and CCD models at fitting the futures prices, we run a likelihood ratio test on the three models. Table 2 shows that, in terms of fitting the futures curves, the MCFA model is significantly better than either the correlated GS model or the correlated CCD model. This result suggests that a maximal specification is indispensable when jointly modeling multiple commodities.

Finally, the empirical results suggests that the commodity risk premium also depends on the relative scarcity. This can be seen by decomposing the risk premium of crude oil into two parts: the log price of crude oil and the relative scarcity (log price of crude oil - log price of heating oil).[21]

---

[20]Note that the functional form in our MCFA model does not impose the upward-sloping correlation structure, it is a result. We thank the referee for pointing this out.

[21]See footnote 15.

The crude oil risk premium is positively associated with its own price and negatively associated with the relative scarcity of crude oil. Similarly, the heating oil risk premium is positively associated with its own price and negatively associated with relative scarcity of heating oil. The sensitivities of the risk premiums to the relative scarcity measures are $b_{1,2}^{\mathbb{Q}} - b_{1,2}^{\mathbb{P}} = -0.06$ for crude oil and $b_{2,1}^{\mathbb{Q}} - b_{2,1}^{\mathbb{P}} = -0.625$ for heating oil. The fact that for heating oil the effect of the relative scarcity on the risk premium is smaller than the one on the convenience yield ($\|-0.625\|$ vs. $\|1.921\|$) evidences that the mechanism proposed in this paper is the one driving the co-movement in prices. These findings show that the heating oil relative scarcity determines the expected heating oil prices mainly through the convenience yield. Furthermore, the observed co-movement in the time-series of crude oil and heating oil (represented in our model by the positive feedback $b_{1,2}^{\mathbb{P}}$ and $b_{2,1}^{\mathbb{P}}$) is mainly driven by the feedbacks in the convenience yields rather than the ones in the risk premiums.

WTI and Brent Crude Oil Pair

As mentioned before, since WTI and Brent crude oils have very similar quality and thus similar usage, the relationship between WTI and Brent crude oil belongs to the substitution relationship. We arbitrarily define WTI crude oil as commodity 1 and Brent crude oil as commodity 2. Neither WTI nor Brent crude oil exhibits seasonal behavior. We thus set

$$\omega_1(t) \equiv 0 \quad \text{and} \quad \omega_2(t) \equiv 0 \tag{32}$$

We use the Kalman filter to estimate the MCFA model. Table 1 panel B shows the results. From the model estimation, we see that most parameters are significant and consistent with the literature on single commodity dynamics. The table shows also that both $b_{1,2}^{\mathbb{Q}} = 0.764$ and $b_{2,1}^{\mathbb{Q}} = 0.456$ are highly significant and of similar magnitude. This indicates that the convenience yield of WTI crude oil depends positively on the relative scarcity of WTI crude oil with respect to Brent crude oil, and vice versa. This is easy to understand since these two commodities are substitutes for one another;

if one commodity has a higher price (larger relative scarcity), people tend to switch to the other commodity. The demand for the abundant oil increases and the expected future scarcity tends to decrease. This implies lower expected prices and a higher convenience yield for the scarce oil.

Figure 7 shows the convenience yield of WTI and Brent crude oils implied by the MCFA model. Figure 8 shows the correlation term-structure for the correlated GS model, the correlated CCD model, and the MCFA model. As is true regarding the heating and crude oil pair, the MCFA model indicates an upward-sloping returns correlation term structure, and the errors of correlations diminish when the correlation approaches one. However, for the GS and CCD models, the correlations begin to go down when the futures time to maturity is longer than one year; in the meantime the error of the correlation also becomes larger. Hence, only the MCFA model is able to generate the upward-sloping correlation curve present in the data. This is largely caused by the positive $b_{1,2}^{\mathbb{Q}}$ and $b_{2,1}^{\mathbb{Q}}$. From the likelihood ratio tests in Table 2, we again see that the MCFA model is significantly better than either the CCD model or the GS model in fitting the futures prices.

The estimation also shows that the risk premiums of the two commodities depend on their relative scarcity. The WTI crude oil risk premium is positively associated with its own price and positively associated with the relative scarcity of WTI crude oil. Similarly, the Brent crude oil risk premium is positively associated with its own price and positively associated with the relative scarcity of Brent crude oil. The sensitivities of WTI crude oil and Brent crude oil to the relative scarcity measures are $b_{1,2}^{\mathbb{Q}} - b_{1,2}^{\mathbb{P}} = 0.232$ and $b_{2,1}^{\mathbb{Q}} - b_{2,1}^{\mathbb{P}} = 0.001$, respectively. As in the previous pair, the magnitudes of the risk-neutral feedback parameters are greater than those related to the risk premiums, implying that the co-movement in the time series is mostly driven by the feedbacks of the convenience yields.

The heating oil and unleaded gasoline pair is a good example of commodities having a complementary relationship, because both share a balanced supply as products of crude oil. In the model estimation, we arbitrarily set heating oil as commodity 1 and gasoline as commodity 2.

From observations on the futures term structures, we see that both heating oil and unleaded gasoline exhibit seasonality; we hence set

$$\omega_1(t) \equiv s_1^c \cos(2\pi t) + s_1^s \sin(2\pi t) \quad \text{and} \quad \omega_2(t) = s_2^c \cos(2\pi t) + s_2^s \sin(2\pi t) \tag{33}$$

From the model estimation in Table 1 panel C, we see that most parameters are significant and consistent with the literature on single commodity dynamics. For example, $b_{1,1}^{\mathbb{P}}$ and $b_{2,2}^{\mathbb{P}}$ are negative, indicating the two commodity futures prices are mean-reverting; both the heating oil and unleaded gasoline display a profound seasonality. The table also shows that the feedback effects between heating oil and unleaded gasoline (i.e., $b_{1,2}^{\mathbb{Q}} = 0.264$ and $b_{2,1}^{\mathbb{Q}} = 1.910$) are both positive and significant. This is also true for the cross-commodity coefficients $b_{1,2}^{\mathbb{P}} = 0.286$ and $b_{2,1}^{\mathbb{P}} = 1.805$, therefore, the two commodities tend to move in the same direction in both physical and risk-neutral measures. Note that with the complementary relationship there are two scenarios regarding movements of the two commodities. First, if the demand and supply shocks are from one *output* commodity (e.g., heating oil), the two output commodities tend to move in opposite directions. For example, if heating oil is experiencing a high demand shock (but gasoline is not), then more crude oil will be refined to produce heating oil. However, since gasoline is the byproduct of this refinery process, the gasoline price will be suppressed. On the other hand, if the demand and supply shocks are from the input commodity (e.g., crude oil), the two output commodities tend to move in the same direction.

Figure 9 shows the convenience yield for both heating oil and unleaded gasoline that is implied by the MCFA model. Figure 10 shows the correlation term-structure for the correlated GS, correlated

CCD, and MCFA models. Again, the MCFA model shows an upward-sloping correlation term structure that the other two models do not have. From the likelihood ratio tests in Table 2, we again see that our model is better than either the CCD model or the GS model in fitting the futures prices.

Finally, the risk premiums of both commodities depend on their relative scarcity but with different signs. Heating oil risk premium is positively associated with its own price and negatively associated with its relative scarcity ($b_{1,2}^{\mathbb{Q}} - b_{1,2}^{\mathbb{P}} = -0.022$). The unleaded gasoline risk premium is positively associated with its own price and positively associated with relative scarcity of gasoline ($b_{2,1}^{\mathbb{Q}} - b_{2,1}^{\mathbb{P}} = 0.105$). Comparing these results with the risk-neutral feedbacks implies that the co-movement in prices is mainly driven by the effect of the relative scarcity on the convenience yields.

In the next section, we show that the MCFA model can guide investors in correctly pricing financial contingent claims.

4 Spread Options Valuation

Spread options are based on the difference between two commodity prices. This difference can be, for example, between the price of an input and the price of the output of a production process (processing spread). NYMEX offers one tradable option on substitute spread (or location spread) between the WTI and Brent crude oil (introduced in March 2008); and two options on the crack spread, the heating oil-crude oil and gasoline-crude oil spread options (introduced in 1994). Also, many firms face "real options" on spreads. For example, manufacturing firms possess an option of transferring the raw material to products at a certain cost, because they can choose not to produce. This option is based on the spread between input and output prices and the strike price corresponds to the production cost. The spread option is of great importance to both commodity market participants and real production firms.

Since the spread is determined by the difference between the two asset prices, it is natural to model the spread by modeling each asset separately. This is the main characteristic of the so-called two-price model, where the short-term correlation is the driver of most of the action in the spread (as in the correlated GS and CCD models). Nearly all researchers use the two-price model for spread option valuation (see Margrabe 1978 and Carmona and Durrleman 2003).[22] However, the two-price model ignores the long-term feedback effect implied by our model. Therefore, the two-price models might be flawed especially for long-maturity spread options. Mbanefo (1997) and Dempster, Medova, and Tang (2008), among others, have documented that the traditional two-price model suffers a problem of overpricing the spread option. Therefore, spread option pricing can be regarded as an out-of-sample test for our theoretical model.

At current time t, the pricing of call and put spread options, $c_t(T, M)$ and $p_t(T, M)$, with strike K on two commodities with futures prices $F_{1,t}(M)$ and $F_{2,t}(M)$, are specified as:

$$c_t(T, M) = e^{-r(T-t)}\mathbb{E}_t^{\mathbb{Q}}\left[\max(F_{2,t}(M) - F_{1,t}(M) - K, 0)\right] \tag{34}$$

$$p_t(T, M) = e^{-r(T-t)}\mathbb{E}_t^{\mathbb{Q}}\left[\max(K - F_{2,t}(M) + F_{1,t}(M), 0)\right] \tag{35}$$

where the time to maturity for the commodity futures and spread options is M and T respectively. To the best of our knowledge, the analytical solution for spread options is not available if $K \neq 0$. Thus, to price the options, we use Monte Carlo simulation. In this section, we simulate the futures prices using three models-the MCFA, the correlated CCD, and the correlated GS models. The futures price dynamics under the risk-neutral measure are specified as,

$$\frac{dF_{i,t}(M)}{F_{i,t}(M)} = G_i(M - t)dZ_t^{\mathbb{Q}} \quad \text{for} \quad i = 1, 2 \tag{36}$$

We choose two spread options: the crack spread option – the spread between heating oil and

---

[22] A notable exception is Duan and Theriault (2007) who price the crack spread option in a cointegration GARCH framework.

WTI crude oil – and the substitute spread option – the spread between WTI crude oil and Brent crude oil. For the crack spread, we assume the underlying futures prices as $F_{1,t}(M) = 100$ (crude oil) and $F_{2,t}(M) = 105$ (heating oil), respectively; and for Brent and WTI crude oil, we use $F_{1,t}(M) = 100$ (Brent crude) and $F_{2,t}(M) = 102$ (WTI crude), respectively.

We focus on spread options of varying maturities to understand the effect of the correlation structure implied by the models. We choose $T = 3$ months for short-maturity options and $T = 5$ years for long-maturity options. Also, for both crack and substitution spreads we choose the same maturity on futures and options (i.e., $M = T$), which is the convention of the spread option specification on NYMEX. We use the estimates from the crude-heating oil and WTI-Brent oil pairs to conduct our simulations, where 2000 paths are simulated for the three models. In order to make the simulation accurate, we use anti-variate techniques in generating random variables and use the same random seed for all three models. Table 3 shows the option values with various strikes for both call and put options of the crack spread and the substitutive spread, respectively. The tables show that both short-term and long-term effects are important determinants of spread option prices. The results indicate that, for long-maturity options ($T = 5$ years), the MCFA model implies lower call and put spread option prices than do the correlated GS and CCD models. Our finding is consistent with the evidence of Mbanefo (1997) that the two-price models tend to overprice the spread option by ignoring the equilibrium relationship, especially for long-maturity options. This is a consequence of the higher long-term correlations implied by the MCFA model. Intuitively, the feedback-effect (positive $b_{1,2}^{\mathbb{Q}}$ and $b_{2,1}^{\mathbb{Q}}$) restricts commodity prices from large deviations from their equilibrium, and thus makes the spread of the prices relatively smaller and less volatile than it is in models without this feature. The lower volatility of the spread leads to lower option values.

The opposite occurs for short-maturity options ($T = 3$ month). The results suggest that the two-price model may under-price short-maturity option values. The short-term correlation in the

CCD and GS models is contaminated because these models are misspecified.[23] Indeed, these models cannot capture the long-term source of co-movement; they tend, therefore, to accommodate long-term effects in the short end of the correlation structure. This creates important biases in option prices.

We perform an out-of-sample test using crack spread option data from NYMEX. We use 2,594 calls and 2,786 puts from January 2000 to December 2006 with maturities between 3 and 12 months and moneyness between 0.6 to 1.4 (strike/spot). Table 4 presents the results of short-maturity heating oil-crude oil (1:1) crack spread options for the MCFA model and for the correlated GS and CCD models. It shows that the MCFA model does considerably better than the others do in matching real data. The other two models tend to under-price both the call and put options. The lower option values are consistent with higher short-term correlation estimates as predicted by our previous analysis. However, the MCFA model reduces the mean pricing error to approximately one-third the size of the error in the CCD model. The root mean square error columns also show that the MCFA model outperforms the benchmark models. Long-maturity options data are not available so we are unable to test the long-term predictions implied by the MCFA model.

5   Conclusion

This paper shows that the economic linkages among commodities, such as production, substitution, or complementary relationships, create a source of correlation between the futures returns of related commodities. The Theory of Storage predicts a negative relation between the convenience yield of a commodity and its own inventories. We extend this result by showing that for the production relationship between crude oil and heating oil, the heating oil convenience yield is also increasing on

---

[23]Figures 6 and 8 show that the cross-commodity feedback effect in our model implies a lower short-term correlation and a larger long-term correlation than is found in the correlated GS and CCD models.

the relative scarcity of this commodity with respect to the crude oil stocks. This result implies a positive feedback effect from the crude oil price to the heating oil price dynamics that is necessary to replicate the observed upward-sloping correlation term structure of futures returns.

In the empirical section, we propose a MCFA reduced-form model that nests the GS and CCD models. We explicitly consider the interdependence of the convenience yield of a commodity and the (log) price difference between this and the other commodities in the economy. Our model allows us to disentangle the two sources of co-movement and implies a flexible correlation term structure. We find that traditional commodity pricing models, such as the GS and CCD models, impose strong restrictions on the correlation structure. We estimate the model for three commodity pairs: heating oil-crude oil, WTI-Brent crude oil, and heating oil-gasoline. The empirical results are consistent with our economic model. Likelihood-ratio tests show that our model is significantly better at fitting futures prices than the correlated versions of the GS and CCD models, which proves the importance of modeling cross-commodity relationships. The empirical results also suggest that the observed co-movement in the time series of the commodity pairs is mainly driven by the feedbacks in the convenience yields rather than in the risk premiums.

We then price the spread options using our MCFA model, because the underlying spreads of these options crucially depend on the cross-commodity economic linkages. For long-maturity options, the MCFA model predicts lower prices than those predicted by the correlated GS and CCD models. This occurs because the MCFA model correctly accounts for higher long-term correlations, and therefore, lower long-term volatilities. The MCFA model also imply higher prices for short-maturity spread options, because of the lower short-term futures correlation. An out-of-sample test using short-maturity crack spread options data shows that our MCFA model considerably reduces the pricing errors generated by the benchmark GS and CCD models.

References

Ahn, Daniel, and Leonid Kogan, 2011, Crude or refined: Identifying oil price dynamics through the crack spread, Working paper, Sloan School, MIT.

Ai, Chunrong, Arjun Chatrath, and Frank Song, 2006, On the comovement of commodity prices, *American Journal of Agricultural Economics* 88, 574–588.

Benveniste, Lawrence M., and Jose A. Scheinkman, 1979, On the differentiability of the value function in dynamic models of economics, *Econometrica* 47, 727–732.

Bessembinder, Hendrik, Jay F. Coughenour, Paul J. Seguin, and Margaret Monroe Smoller, 1995, Mean reversion in equilibrium asset prices: Evidence from the futures term structure, *Journal of Finance* 50, 361–375.

Brennan, Michael J., 1958, The supply of storage, *American Economic Review* 48, 50–72.

——— , 1991, The price of convenience and the valuation of commodity contingent claims, in D. Lund, and B. Oksendal, ed.: *Stochastic Models and Option Values* (North Holland).

Carmona, René, and Valdo Durrleman, 2003, Pricing and hedging spread options, *SIAM Review* 45, 627–685.

Casassus, Jaime, and Pierre Collin-Dufresne, 2005, Stochastic convenience yield implied from commodity futures and interest rates, *Journal of Finance* 60, 2283–2332.

——— , and Bryan Routledge, 2008, Equilibrium commodity prices with irreversible investment and non-linear technology, Working Paper, Columbia University.

Cortazar, Gonzalo, Carlos Milla, and Felipe Severino, 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.

Dempster, M.A.H., Elena Medova, and Ke Tang, 2008, Long term spread option valuation and hedging, *Journal of Banking and Finance* 32, 2530–2540.

Duan, Jin-Chuan, and Annie Theriault, 2007, Co-integration in crude oil components and the pricing of crack spread options, Working paper University of Toronto.

Duffie, Darrell, 2001, *Dynamic Asset Pricing Theory* (Princeton University Press).

——— , and Rui Kan, 1996, A yield-factor model of interest rates, *Mathematical Finance* 6, 379–406.

Durbin, James, and Siem Jan Koopman, 2001, *Time Series Analysis by State Space Methods* (Oxford University Press).

Energy Information Administration, 2002, Manufacturing energy consumption survey, *U.S. Department of Energy*.

Engle, Robert F., and Clive W.J. Granger, 1987, Co-integration and error correction: Representation, estimation, and testing, *Econometrica* 55, 251–276.

Geman, Helyette, and Vu-Nhat Nguyen, 2005, Soybean inventory and forward curve dynamics, *Management Science* 51, 1076–1091.

Gibson, Rajna, and Eduardo S. Schwartz, 1990, Stochastic convenience yield and the pricing of oil contingent claims, *Journal of Finance* 45, 959–976.

Girma, Paul B., and Albert S. Paulson, 1999, Risk arbitrage opportunities in petroleum futures spreads, *Journal of Futures Markets* 19, 931–955.

Hamilton, James D., 1994, *Time Series Analysis* (Princeton University Press).

Harvey, Andrew C., 1991, *Forecasting, Structural Time Series Models and the Kalman Filter* (Cambridge University Press).

Kaldor, Nicholas, 1939, Speculation and economic stability, *Review of Economic Studies* 7, 1–27.

Keynes, John M., 1923, Some aspects of commodity markets, *Manchester Guardian Commercial: European Reconstruction Series*.

Malliaris, A. G., and Jorge L. Urrutia, 1996, Linkages between agricultural commodity futures contracts, *Journal of Futures Markets* 16, 595–609.

Margrabe, William, 1978, The value of an option to exchange one asset for another, *Journal of Finance* 33, 177–186.

Mbanefo, Art, 1997, Co-movement term structure and the valuation of energy spread options, in Michael A. H. Dempster, and Stanley R. Pliska, ed.: *Mathematics of Derivative Securities*No. 15 in Publications of the Newton Institute . pp. 88–102 (Cambridge University Press).

Paschke, Raphael, and Marcel Prokopczuk, 2007, Integrating multiple commodities in a model of stochastic price dynamics, Working Paper University of Mannheim.

Pindyck, Robert S., 2001, The dynamics of commodity spot and futures markets: A primer, *Energy Journal* 22, 1–29.

Richter, Martin C., and Carsten Sørensen, 2002, Stochastic volatility and seasonality in commodity futures and options: The case of soybeans, Working Paper, Copenhagen Business School.

Routledge, Bryan R., Duane J. Seppi, and Chester S. Spatt, 2000, Equilibrium forward curves for commodities, *Journal of Finance* 55, 1297–1338.

——— , 2001, The spark spread: An equilibrium model of cross-commodity price relationships in electricity, Working Paper, Carnegie Mellon University.

Schwartz, Eduardo S., 1997, The stochastic behavior of commodity prices: Implications for valuation and hedging, *Journal of Finance* 52, 923–973.

Telser, Lester G., 1958, Futures trading and the storage of cotton and wheat, *Journal of Political Economy* 66, 233–255.

Trolle, Anders B., and Eduardo S. Schwartz, 2009, Unspanned stochastic volatility and the pricing of commodity derivatives, *Review of Financial Studies* 22, 4423–4461.

Working, Holbrook, 1948, Theory of the inverse carrying charge in futures markets, *Journal of Farm Economics* 30, 1–28.

Appendix

A   Proofs for the Economic Model

A.1   Proof of Proposition 1

Let us denote by $J[K_t, Q_{1,t}, Q_{2,t}, t]$ the value function associated with the representative agent's problem in equations (2)-(5), and by $j[K_t, Q_{1,t}, Q_{2,t}] = e^{\theta t} J[K_t, Q_{1,t}, Q_{2,t}, t]$ the *current* value function for the same problem. Therefore,

$$j[K_t, Q_{1,t}, Q_{2,t}] = \sup_{\{C_{K,v}, C_{1,v}, C_{2,v}, q_v, I_{1,v}, I_{2,v}\} \in \mathcal{A}} \mathbb{E}_t^{\mathbb{P}} \left[\int_t^{\infty} e^{-\theta(v-t)} u[C_{K,v}, C_{1,v}, C_{2,v}] dv \right] \tag{A1}$$

Note that given the set-up of the model, the value function $j[\cdot]$ is not a function of time. The solution of the our problem is determined by the following Hamilton-Jacobi-Bellman (HJB) equation:[24]

$$\sup_{\{C_K, C_1, C_2, q, I_1, I_2\} \in \mathcal{A}} \{u[C_K, C_1, C_2] + \mathcal{D}j - \theta j\} = 0 \tag{A2}$$

where $\mathcal{D}$ is the Itô operator

$$\mathcal{D}j = (\alpha_K K - C_K - I_1 - I_2) \frac{\partial j}{\partial K} + (f_1[I_1; Q_1] - q - C_1) \frac{\partial j}{\partial Q_1} + (f_2[I_2, q; Q_2] - C_2) \frac{\partial j}{\partial Q_2} + \frac{1}{2} \sigma_K^2 K^2 \frac{\partial^2 j}{\partial K^2} \tag{A3}$$

with $\frac{\partial j}{\partial K}$, $\frac{\partial j}{\partial Q_1}$ and $\frac{\partial j}{\partial Q_2}$ representing the marginal value of an additional unit of numeraire good, crude oil and heating oil, respectively. $\frac{\partial^2 j}{\partial K^2}$ is the second derivative of the current value function with respect to K. The first-order conditions with respect to the consumption of capital, heating oil and crude oil are:

$$u_K = \frac{\partial j}{\partial K}, \quad u_1 = \frac{\partial j}{\partial Q_1} \quad \text{and} \quad u_2 = \frac{\partial j}{\partial Q_2} \tag{A4}$$

where u_i for $i \in \{K, Q_1, Q_2\}$ are the marginal utilities of consumption of capital, heating oil and crude oil, respectively. The first-order conditions with respect to the demand of crude oil and the investment in the commodity sectors in terms of the marginal utilities are:

$$\frac{\partial f_2}{\partial q} = \frac{u_1}{u_2}, \quad \frac{\partial f_1}{\partial I_1} = \frac{u_K}{u_1} \quad \text{and} \quad \frac{\partial f_2}{\partial I_2} = \frac{u_K}{u_2} \tag{A5}$$

The convenience yield of commodity i, $\delta_{i,t}$, is defined as a benefit of holding inventories of that asset; therefore, the price of commodity i at time t, $S_{i,t}$, must satisfy the following equilibrium condition:

$$e^{-\theta t} u_{K,t} S_{i,t} = \mathbb{E}_t^{\mathbb{P}} \left[\int_t^T e^{-\theta v} u_{K,v} \delta_{i,v} S_{i,v} dv + e^{-\theta T} u_{K,T} S_{i,T} \right] \quad \text{for} \quad i = 1, 2 \tag{A6}$$

where S_i is the equilibrium commodity price. Since both commodities are also consumption goods, the commodity price is the marginal rate of substitution of that commodity for the numeraire, i.e., $S_i = \frac{u_i}{u_K}$. Replacing the commodity price in the equation above yields a simple Euler equation for $\delta_{i,t}$:

$$e^{-\theta t} u_{i,t} = \mathbb{E}_t^{\mathbb{P}} \left[\int_t^T e^{-\theta v} u_{i,v} \delta_{i,v} dv + e^{-\theta T} u_{i,T} \right] \quad \text{for} \quad i = 1, 2 \tag{A7}$$

---

[24]The following variables are all time dependent. Hereafter, we drop this dependance to simplify the notation.

To obtain the convenience yield $\delta_{i,t}$ we build the following $\mathbb{P}$-Martingale:

$$M_{i,t} = \int_0^t e^{-\theta v} u_{i,v} \delta_{i,v} dv + e^{-\theta t} u_{i,t} \quad \text{for} \quad i = 1, 2 \tag{A8}$$

and use that $\mathbb{E}_t^{\mathbb{P}}[dM_{i,t}] = 0$ to get

$$\delta_{i,t} dt = -\mathbb{E}_t^{\mathbb{P}} \left[\frac{d(e^{-\theta t} u_{i,t})}{e^{-\theta t} u_{i,t}} \right] \quad \text{for} \quad i = 1, 2 \tag{A9}$$

To obtain the result in Proposition 1, we first apply Itô's Lemma to $e^{-\theta t} u_{i,t}$. Then we differentiate the HJB equation in (A2) with respect to Q_i and replace the high-order partial derivatives in equation (A9). This yields equation (6).

For the proof of second part of the proposition, we first note that equation (8) and the concavity of $f_2[.]$ imply that an increase in the relative scarcity of heating oil, $\frac{S_{2,t}}{S_{1,t}}$, increases the demand for crude oil, q_t. Using that $\frac{\partial^2 f_2}{\partial q \partial Q_2} > 0$ (see footnote 8) and the definition of the heating oil convenience yield in (6), we obtain equation (7).

A.2 Proof of Proposition 2

The convenience yields for crude and heating oil in equations (11) and (12) are obtained directly by replacing the Cobb-Douglas production functions from equation (9) and (10) into (6). To get equations (13) and (14) we obtain the optimal crude oil demand and investment rates from the first order condition in (A5) and use the relation between prices and marginal utilities to express the convenience yields in term of the commodity prices.

B Proofs for the Empirical Model

B.1 Proof of Proposition 3

Under the risk-neutral measure, the ith futures prices $F_{i,t}(Y_t, T)$ need to satisfy,

$$F_{i,t}(Y_t, T) = \mathbb{E}_t^{\mathbb{Q}}[S_{i,T}] \quad \text{for} \quad i = 1, \ldots, n \tag{B1}$$

Let $\tau = T - t$. The futures price $F_{i,t}(Y_t, t + \tau)$ should satisfy the following vector-based Feynman-Kac equation,

$$-\frac{\partial F_i}{\partial \tau} + \frac{\partial F_i}{\partial Y}' (U^{\mathbb{Q}} + \Psi^{\mathbb{Q}} Y) + \frac{1}{2} \text{Tr} \left(\frac{\partial^2 F_i}{\partial Y^2} \Omega \right) = 0 \tag{B2}$$

with boundary condition $F_{i,t}(Y_t, t) = \exp(x_{i,t})$.

Assume that

$$\log(F_{i,t}(Y_t, t + \tau)) = m_i(\tau) + G_i(\tau) Y_t \tag{B3}$$

where $m_i(\tau)$ is the ith element of the $m(\tau)$ vector, and $G_i(\tau)$ is the ith row of the $G(\tau)$ matrix. By plugging (B3) into (B2), we have two ordinary differential equations

$$-\frac{\partial m_i}{\partial \tau} + G_i U^{\mathbb{Q}} + \frac{1}{2} G_i(\tau) \Omega G_i(\tau)' = 0$$
$$\frac{\partial G_i}{\partial \tau} - G_i(\tau) \Psi^{\mathbb{Q}} = 0 \tag{B4}$$

36

with boundary condition

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

Thus, the solution for (B2) is

$$
\begin{aligned}
m_i(\tau) &= \int_0^\tau \left(G_i(u)\, U^{\mathbb{Q}} + \frac{1}{2} G_i(u)\, \Omega\, G_i(u)' \right) du && \text{(B5)} \\
G(\tau) &= \exp(\Psi^{\mathbb{Q}}\tau)
\end{aligned}
$$

$G_i(\tau)$ denotes the ith row of the $G(\tau)$ matrix. When $\Psi^{\mathbb{Q}}$ is diagnosable,

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

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

Grouping the elements m_i from equation (B5) yields the solution in Proposition 3.

B.2   Proof of Proposition 4

First we obtain the first two conditional moments for the $2n$ state variables. The solution of the Gaussian system of SDEs in equation (23) that drives the dynamics of the state vector Y_t is given by:

$$
Y_T = e^{\Psi^{\mathbb{P}}(T-t)}Y_t + \int_t^T e^{\Psi^{\mathbb{P}}(T-v)}U^{\mathbb{P}}(v)dv + \int_t^T e^{\Psi^{\mathbb{P}}(T-v)}dZ_v^{\mathbb{P}} \tag{B6}
$$

Hence, the conditional moments for Y_T are:

$$
\mathbb{E}_t^{\mathbb{P}}[Y_T] = e^{\Psi^{\mathbb{P}}(T-t)}Y_t + \int_t^T e^{\Psi^{\mathbb{P}}(T-v)}U^{\mathbb{P}}(v)dv \tag{B7}
$$

and

$$
\begin{aligned}
\mathbb{V}ar_t^{\mathbb{P}}[Y_T] &= \mathbb{E}_t^{\mathbb{P}}[(Y_T - \mathbb{E}_t^{\mathbb{P}}[Y_T])(Y_T - \mathbb{E}_t^{\mathbb{P}}[Y_T])'] && \text{(B8)} \\
&= \mathbb{E}_t^{\mathbb{P}}\left[\left(\int_t^T e^{\Psi^{\mathbb{P}}(T-v)}dZ_v^{\mathbb{P}}\right) \left(\int_t^T e^{\Psi^{\mathbb{P}}(T-v)}dZ_v^{\mathbb{P}}\right)' \right] && \text{(B9)} \\
&= \int_t^T e^{\Psi^{\mathbb{P}}(T-v)}\Omega\, e^{(\Psi^{\mathbb{P}})'(T-v)}dv && \text{(B10)}
\end{aligned}
$$

To obtain equation (27) in Proposition 4, we replace equation (25) in the covariance between the futures returns of commodities i and j and use the closed-form expression for the futures from Proposition 3:

$$
\begin{aligned}
\Sigma_{i,j,t,t_1}(T) &= \mathbb{E}_t^{\mathbb{P}}\left[\left(r_{i,t,t_1}(T) - \mathbb{E}_t^{\mathbb{P}}[r_{i,t,t_1}(T)]\right)\left(r_{j,t,t_1}(T) - \mathbb{E}_t^{\mathbb{P}}[r_{j,t,t_1}(T)]\right)' \right] && \text{(B11)} \\
&= \mathbb{E}_t^{\mathbb{P}}[(G_i(T-t_1)Y_{t_1} - \mathbb{E}_t^{\mathbb{P}}[G_i(T-t_1)Y_{t_1}])(G_j(T-t_1)Y_{t_1} - \mathbb{E}_t^{\mathbb{P}}[G_j(T-t_1)Y_{t_1}])'] && \text{(B12)} \\
&= G_i(T-t_1)\mathbb{V}ar_t^{\mathbb{P}}[Y_{t_1}]G_j(T-t_1)' && \text{(B13)}
\end{aligned}
$$

The elements of the covariance matrix of futures returns in equation (27), $\Sigma_{t,t_1}(T)$, are given by the above equation.

C   Empirical Method

One of the difficulties of estimating the model is that the state variables are not directly observable. A useful method for Maximum Likelihood Estimation of the model is addressing the model in a state-space form and using the Kalman filter methodology to estimate the latent variables.[25] The state-space form consists of a transition equation and a measurement equation. The transition equation shows the data-generating process. The measurement equation relates a multivariate time series of observable variables (in our case, futures prices at varying maturities) to an unobservable vector of state variables (in our case, the (log) spot prices $x_{i,t}$ and $\eta_{i,t}$ $(i = 1, \ldots, n)$). The measurement equation is obtained using the (log) futures prices in equation (24) by adding uncorrelated noises to take account of the pricing errors.

Suppose that data are sampled in equally separated times t_k, $k = 1, \ldots, K$. Denote $\Delta t = t_{k+1} - t_k$ as the time interval between two subsequent observations. Let Y_k represent the vector of state variables at time t_k. Thus, we can obtain the transition equation,

$$Y_{k+1} = (\Psi^{\mathbb{P}} \Delta t + I) Y_k + U^{\mathbb{P}}(t) \Delta t + w_k \tag{C14}$$

where w_k is a $2n \times 1$ random noise vector following zero-mean normal distributions.

For the measurement equation at time t_k, we consider the vector of the log of futures prices $F_k = (F_{1,k}(\tau_1), \ldots, F_{n,k}(\tau_1), \ldots, F_{1,k}(\tau_M), \ldots, F_{n,k}(\tau_M))'$ where τ_j denotes the times to maturity.[26] The $(nM) \times 1$ vector $\log(F_k)$ can be written as,

$$\log(F_k) = \overline{m} + \overline{G} Y_k + \epsilon_k \tag{C15}$$

where

$$\begin{aligned}
\overline{m} &= (m_1(\tau_1), \ldots, m_n(\tau_1), \ldots, m_1(\tau_M), \ldots, m_n(\tau_M))' \\
\overline{G} &= (G_1(\tau_1), \ldots, G_n(\tau_1), \ldots, G_1(\tau_M), \ldots, G_n(\tau_M))'
\end{aligned}$$

and ε_k is a $(nM) \times 1$ vector representing the model errors with its variance covariance matrix Υ. In order to reduce the number of parameters to estimate, we assume that the standard errors for all contracts are the same. This also reflects the notion that we want our model to price the n commodities and M contracts equally well. Therefore, we define $\Upsilon = e^2 I_{nM}$, where e is the pricing error of the log of the futures prices and I_{nM} is the $(nM) \times (nM)$ identity matrix.

---

[25]Hamilton (1994) and Harvey (1991) give a good description of estimation, testing, and model selection of state-space models.

[26]Since our model has $2n$ factors we need $M \geq 2$.

Table 1: Parameter estimation for three pairs

The data consist of weekly futures prices of three pairs from 1995.01 to 2010.09 (821 observations). The estimates correspond to the 4-factor multi-commodity feedback affine model.

Panel A: The WTI crude and heating oil pair

Parameter	Estimate	Std. Err.	Parameter	Estimate	Std. Err.
$b_{1,1}^{\mathbb{P}}$	-0.472	(0.031)	$\rho_{2,4}$	-0.179	(0.056)
$b_{1,1}^{\mathbb{Q}}$	-0.663	(0.020)	$\rho_{3,4}$	0.004	(0.083)
$b_{1,2}^{\mathbb{P}}$	0.669	(0.021)	σ_1	0.367	(0.010)
$b_{1,2}^{\mathbb{Q}}$	0.609	(0.035)	σ_2	0.352	(0.010)
$b_{2,1}^{\mathbb{P}}$	2.546	(0.111)	σ_3	0.372	(0.014)
$b_{2,1}^{\mathbb{Q}}$	1.921	(0.335)	σ_4	0.175	(0.013)
$b_{2,2}^{\mathbb{P}}$	-2.188	(0.317)	$\chi_1^{\mathbb{Q}}$	0.134	(0.012)
$b_{2,2}^{\mathbb{Q}}$	-2.888	(0.092)	$\chi_2^{\mathbb{Q}}$	-1.811	(0.123)
$a_{1,2}$	0.010	(0.009)	$\chi_1^{\mathbb{P}}$	0.288	(0.081)
$a_{2,1}$	-0.327	(0.033)	$\chi_2^{\mathbb{P}}$	-2.798	(0.242)
κ_1	1.324	(0.026)	$\mu_1^{\mathbb{P}}$	-0.212	(0.326)
κ_2	0.240	(0.012)	$\mu_2^{\mathbb{P}}$	-0.072	(0.406)
$\rho_{1,2}$	0.772	(0.032)	s_1^c	0	–
$\rho_{1,3}$	0.837	(0.027)	s_1^s	0	–
$\rho_{1,4}$	0.092	(0.070)	s_2^c	4.896	(0.235)
$\rho_{2,3}$	0.680	(0.039)	s_2^s	2.998	(0.194)
ϵ	0.013	(0.002)			
Log-likelihood	20,687				

Panel B: The WTI and Brent oil pair

Parameter	Estimate	Std. Err.	Parameter	Estimate	Std. Err.
$b_{1,1}^{\mathbb{P}}$	-0.555	(0.031)	$\rho_{2,4}$	0.564	(0.058)
$b_{1,1}^{\mathbb{Q}}$	-0.851	(0.041)	$\rho_{3,4}$	-0.049	(0.067)
$b_{1,2}^{\mathbb{P}}$	0.532	(0.138)	σ_1	0.376	(0.010)
$b_{1,2}^{\mathbb{Q}}$	0.764	(0.035)	σ_2	0.338	(0.008)
$b_{2,1}^{\mathbb{P}}$	0.455	(0.120)	σ_3	0.129	(0.001)
$b_{2,1}^{\mathbb{Q}}$	0.456	(0.034)	σ_4	0.369	(0.012)
$b_{2,2}^{\mathbb{P}}$	-0.382	(0.032)	$\chi_1^{\mathbb{Q}}$	-0.503	(0.062)
$b_{2,2}^{\mathbb{Q}}$	-0.450	(0.032)	$\chi_2^{\mathbb{Q}}$	0.228	(0.036)
$a_{1,2}$	-0.964	(0.041)	$\chi_1^{\mathbb{P}}$	-0.548	(0.062)
$a_{2,1}$	-0.253	(0.051)	$\chi_2^{\mathbb{P}}$	0.203	(0.063)
κ_1	1.074	(0.095)	$\mu_1^{\mathbb{P}}$	-0.153	(0.449)
κ_2	1.225	(0.020)	$\mu_2^{\mathbb{P}}$	-0.117	(0.408)
$\rho_{1,2}$	0.929	(0.017)	s_1^c	0	–
$\rho_{1,3}$	0.225	(0.091)	s_1^s	0	–
$\rho_{1,4}$	0.594	(0.046)	s_2^c	0	–
$\rho_{2,3}$	-0.057	(0.090)	s_2^s	0	–
ϵ	0.008	(0.001)			
Log-likelihood	25,663				

Panel C: The heating oil and gasoline pair

Parameter	Estimate	Std. Err.	Parameter	Estimate	Std. Err.
$b_{1,1}^{\mathbb{P}}$	-0.154	(0.031)	$\rho_{2,4}$	0.657	(0.045)
$b_{1,1}^{\mathbb{Q}}$	-0.245	(0.049)	$\rho_{3,4}$	-0.028	(0.079)
$b_{1,2}^{\mathbb{P}}$	0.286	(0.054)	σ_1	0.394	(0.012)
$b_{1,2}^{\mathbb{Q}}$	0.264	(0.059)	σ_2	0.423	(0.013)
$b_{2,1}^{\mathbb{P}}$	1.805	(0.360)	σ_3	0.089	(0.006)
$b_{2,1}^{\mathbb{Q}}$	1.910	(0.037)	σ_4	0.483	(0.033)
$b_{2,2}^{\mathbb{P}}$	-1.919	(0.144)	$\chi_1^{\mathbb{Q}}$	0.377	(0.114)
$b_{2,2}^{\mathbb{Q}}$	-2.065	(0.042)	$\chi_2^{\mathbb{Q}}$	-0.168	(0.019)
$a_{1,2}$	-1.585	(0.113)	$\chi_1^{\mathbb{P}}$	0.344	(0.153)
$a_{2,1}$	1.193	(0.062)	$\chi_2^{\mathbb{P}}$	-0.144	(0.048)
κ_1	0.279	(0.064)	$\mu_1^{\mathbb{P}}$	-0.288	(0.546)
κ_2	2.729	(0.066)	$\mu_2^{\mathbb{P}}$	0.031	(0.539)
$\rho_{1,2}$	0.820	(0.031)	s_1^c	4.647	(1.088)
$\rho_{1,3}$	0.225	(0.091)	s_1^s	-0.917	(0.273)
$\rho_{1,4}$	0.851	(0.028)	s_2^c	-0.195	(0.011)
$\rho_{2,3}$	0.073	(0.092)	s_2^s	-0.310	(0.016)
ϵ	0.019	(0.003)			
Log-likelihood	18,166				

Table 2: Likelihood ratio tests for three pairs

This table compares the MCFA model with the correlated CCD and GS models. The parameters used in the calculation are from Table 1. Correlated CCD and GS models correspond to the cases $b_{1,2}^{\mathbb{P}} = b_{1,2}^{\mathbb{Q}} = b_{2,1}^{\mathbb{P}} = b_{2,1}^{\mathbb{Q}} = a_{1,2} = a_{2,1} = 0$ and $b_{1,1}^{\mathbb{P}} = b_{1,1}^{\mathbb{Q}} = b_{1,2}^{\mathbb{P}} = b_{1,2}^{\mathbb{Q}} = b_{2,1}^{\mathbb{P}} = b_{2,1}^{\mathbb{Q}} = b_{2,2}^{\mathbb{P}} = b_{2,2}^{\mathbb{Q}} = a_{1,2} = a_{2,1} = 0$ respectively. The 1% significant levels are 16.81, 23.2 and 13.28, respectively for MCFA vs. correlated CCD, MCFA vs. correlated GS, and correlated CCD vs. correlated GS models. The statistics that are significant at the 1% level are marked with an asterisk.

Panel A: WTI crude and heating oil pair

	Log-likelihood		LR statistic
MCFA	20,687	MCFA vs. CCD	438 (*)
CCD	20,518	MCFA vs. GS	550 (*)
GS	20,462	CCD vs. GS	112 (*)

Panel B: WTI and Brent oil pair

	Log-likelihood		LR statistic
MCFA	25,663	MCFA vs. CCD	3,532 (*)
CCD	23,897	MCFA vs. GS	3,760 (*)
GS	23,783	CCD vs. GS	228 (*)

Panel C: Heating oil and Gasoline pair

	Log-likelihood		LR statistic
MCFA	18,166	MCFA vs. CCD	200 (*)
CCD	18,066	MCFA vs. GS	330 (*)
GS	18,001	CCD vs. GS	130 (*)

Table 3: Values for two spread options

The table shows the crack spread option prices between heating and WTI crude oil prices and between WTI and Brent oil prices for different strikes. Panel A presents the call option values, while Panel B presents the put option values. The options and the underlying futures have the same maturity. The parameters used in the calculation are from Table 1.

Panel A: The heating oil-crude oil crack spread option

Call Options

Strike	Time to maturity = 3 months			Time to maturity = 5 years		
	MCFA	CCD	GS	MCFA	CCD	GS
1	6.097	5.560	5.609	7.222	8.635	8.336
3	4.856	4.268	4.327	6.003	7.424	7.162
5	3.776	3.162	3.228	4.957	6.341	6.120
7	2.864	2.261	2.335	4.068	5.391	5.208
9	2.120	1.551	1.625	3.324	4.567	4.419

Put Options

Strike	Time to maturity = 3 months			Time to maturity = 5 years		
	MCFA	CCD	GS	MCFA	CCD	GS
1	2.090	1.551	1.601	3.055	4.542	4.239
3	2.849	2.259	2.319	3.837	5.331	5.065
5	3.769	3.152	3.220	4.791	6.248	6.023
7	4.857	4.252	4.327	5.901	7.298	7.111
9	6.113	5.542	5.617	7.157	8.473	8.322

Panel B: The WTI - Brent oil substitution (or location) spread option

Call Options

Strike	Time to maturity = 3 months			Time to maturity = 5 years		
	MCFA	CCD	GS	MCFA	CCD	GS
0	4.058	3.592	3.656	4.786	5.718	6.078
1	3.465	2.991	3.056	4.229	5.163	5.500
2	2.928	2.449	2.513	3.724	4.653	4.964
3	2.442	1.968	2.030	3.278	4.187	4.469
4	2.012	1.553	1.611	2.880	3.766	4.017

Put Options

Strike	Time to maturity = 3 months			Time to maturity = 5 years		
	MCFA	CCD	GS	MCFA	CCD	GS
0	2.055	1.586	1.649	2.740	3.728	4.103
1	2.462	1.985	2.050	3.183	4.173	4.525
2	2.925	2.442	2.506	3.678	4.663	4.989
3	3.439	2.962	3.023	4.232	5.197	5.495
4	4.009	3.547	3.604	4.834	5.776	6.042

Table 4: Out-of-sample comparison of heating-crude oil crack spread options

The table shows the results of the out-of-sample tests using short-maturity heating oil-crude oil (1:1) crack spread options data. The market data consists of 2,594 calls and 2,786 puts from January 2000 to December 2006 with maturities between 3 and 12 months and moneyness between 0.6 to 1.4 (strike/spot). The parameters used in the calculation are from Table 1.

	MCFA Model							
	Call Options				Put Options			
Time to Maturity	3 - 7 months		8 - 12 months		3 - 7 months		8 - 12 months	
Moneyness	MPE	RMSE	MPE	RMSE	MPE	RMSE	MPE	RMSE
0.6-0.9	-0.0992	0.2006	-0.1238	0.2616	0.0858	0.2358	0.0274	0.2408
0.9-1.1	-0.0472	0.2348	-0.0746	0.2598	0.0145	0.2373	0.0154	0.2831
1.1-1.4	-0.0174	0.2439	-0.0174	0.2439	-0.0278	0.2365	0.0465	0.2480

	Correlated CC Model							
	Call Options				Put Options			
Time to Maturity	3 - 7 months		8 - 12 months		3 - 7 months		8 - 12 months	
Moneyness	MPE	RMSE	MPE	RMSE	MPE	RMSE	MPE	RMSE
0.6-0.9	-0.321	0.504	0.052	0.577	-0.172	0.360	0.115	0.325
0.9-1.1	-0.127	0.565	0.135	0.473	-0.171	0.400	0.001	0.424
1.1-1.4	-0.272	0.525	-0.272	0.525	-0.316	0.496	0.390	0.408

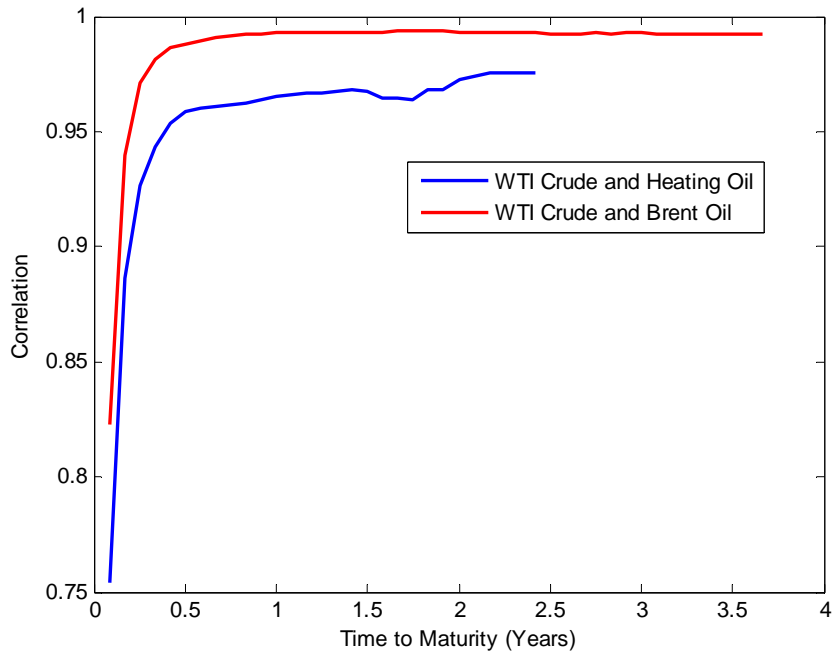|	Correlated Schwartz Model							
	Call Options				Put Options			
Time to Maturity	3 - 7 months		8 - 12 months		3 - 7 months		8 - 12 months	
Moneyness	MPE	RMSE	MPE	RMSE	MPE	RMSE	MPE	RMSE
0.6-0.9	-0.287	0.479	0.084	0.563	-0.133	0.341	0.143	0.332
0.9-1.1	-0.080	0.544	0.162	0.474	-0.122	0.375	0.030	0.420
1.1-1.4	-0.225	0.496	-0.225	0.496	-0.251	0.447	0.405	0.422

Figure 1: Correlation term structure for the WTI crude oil-heating oil and WTI-Brent crude oil pairs. The figure plots the correlation between weekly futures returns for various maturity futures.
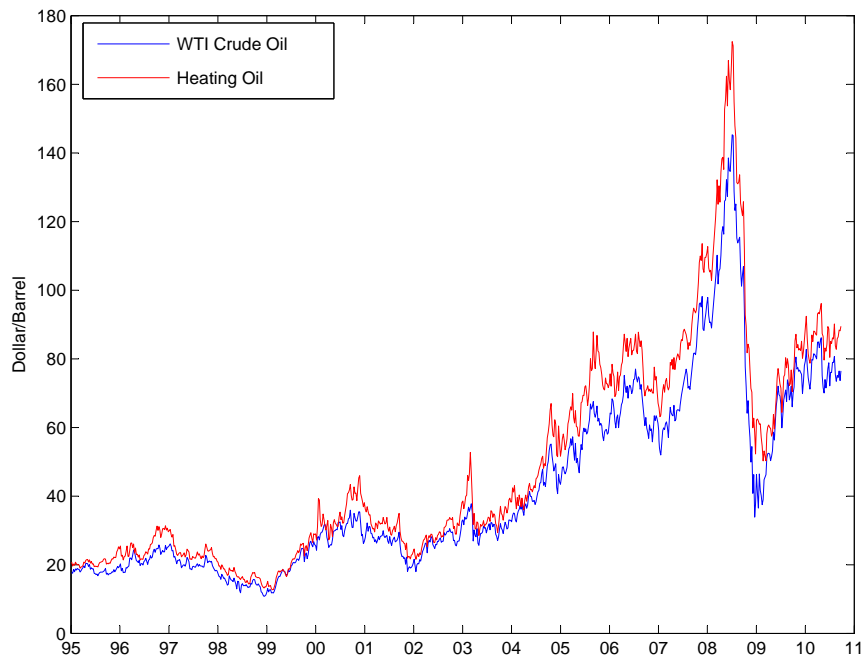

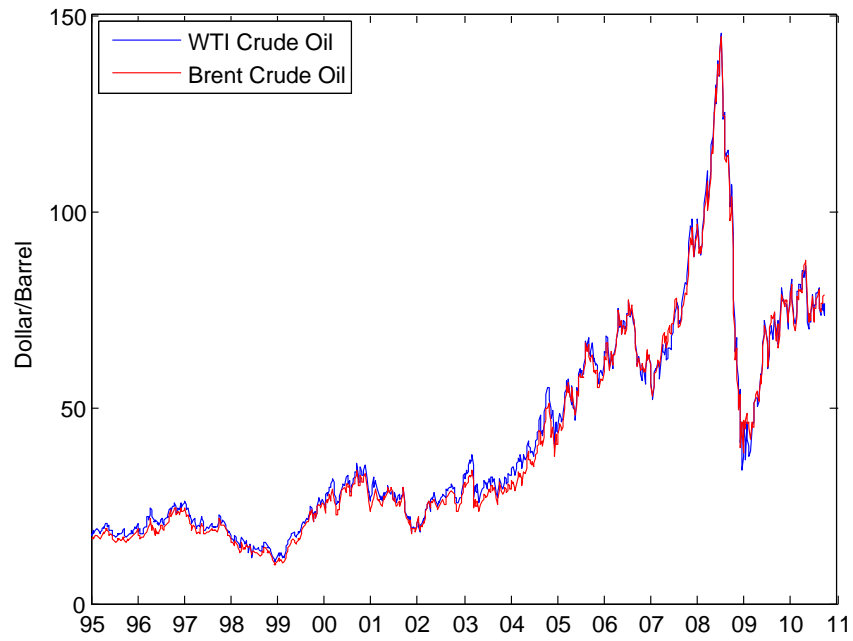Figure 2: Nearest-to-maturity futures prices of WTI crude oil and Heating oil.

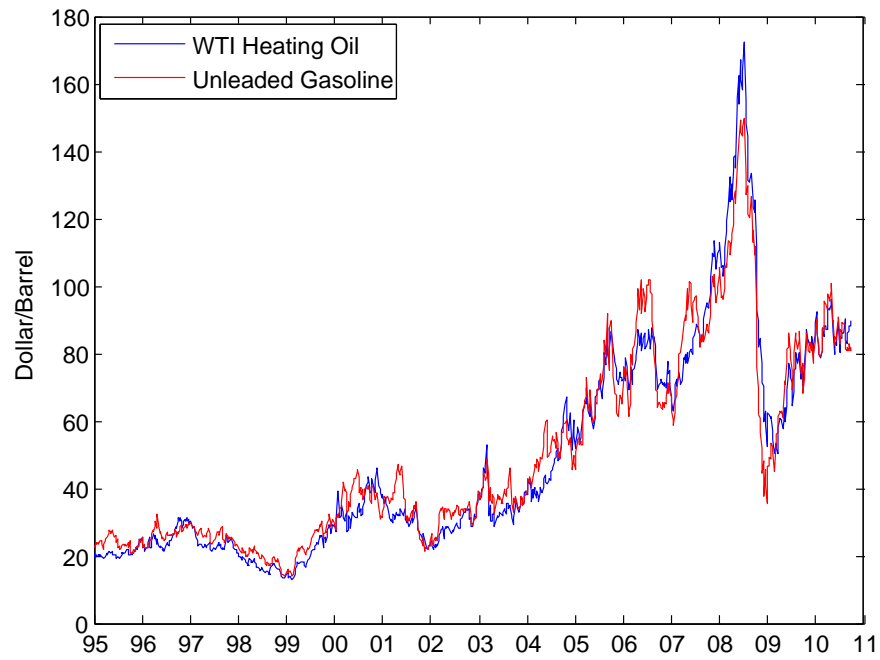Figure 3: Nearest-to-maturity futures prices of WTI and Brent crude oil.

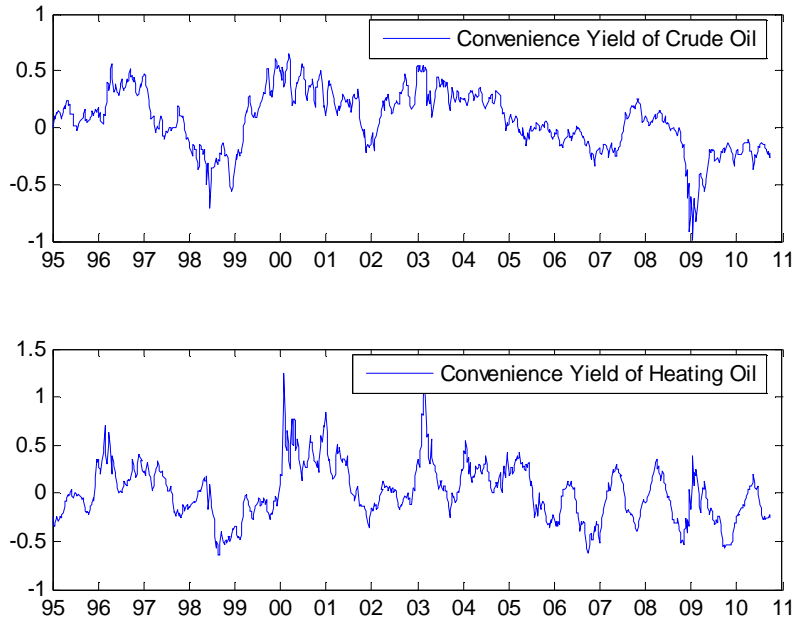Figure 4: Nearest-to-maturity futures prices of WTI heating oil and unleaded gasoline.
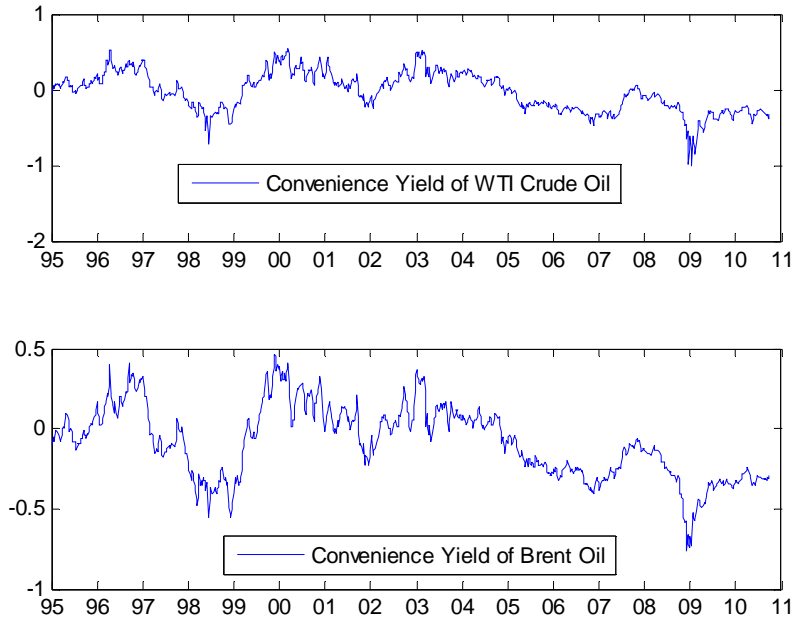
Figure 5: The implied convenience yield for the crude and heating oil. The implied convenience yields are from the MCFA model with the parameters obtained from Table 1.
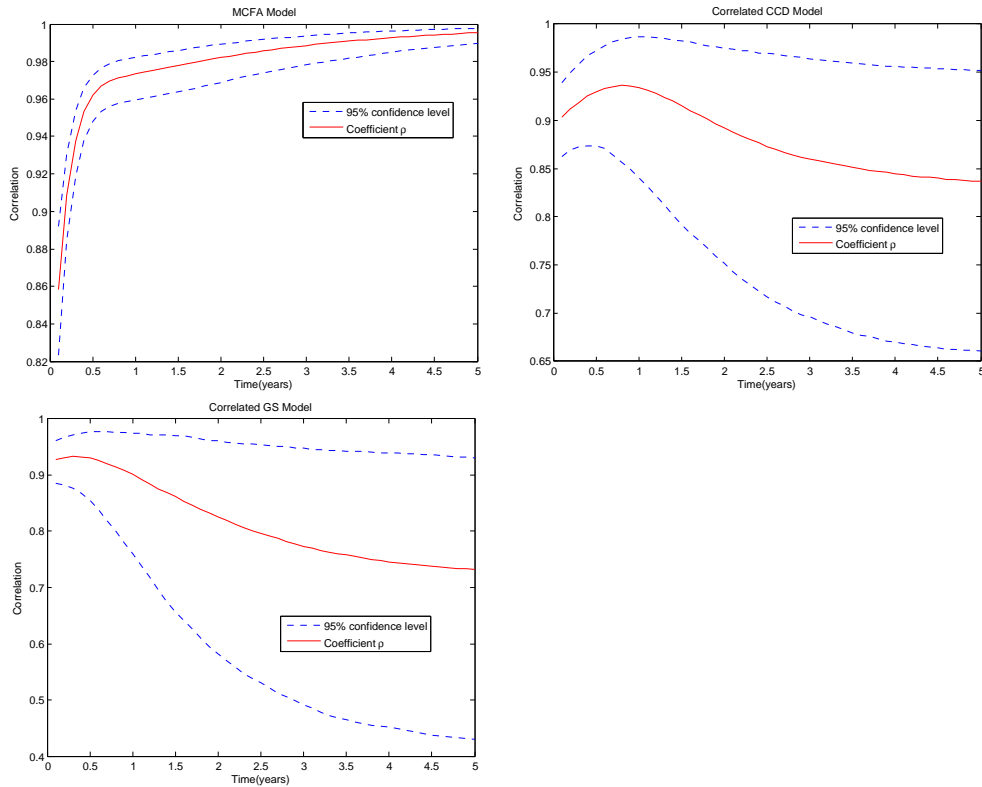
Figure 6: Correlation term structures (instantaneous futures correlation vs. time to maturity) of the WTI crude and heating oil pair for the MCFA, correlated CCD, and correlated GS models. The 95% confidence levels are obtained by bootstrapping the model parameters.

Figure 7: The implied convenience yield for the WTI crude oil and Brent crude oil. The implied convenience yields are from the MCFA model with the parameters obtained from Table 1.
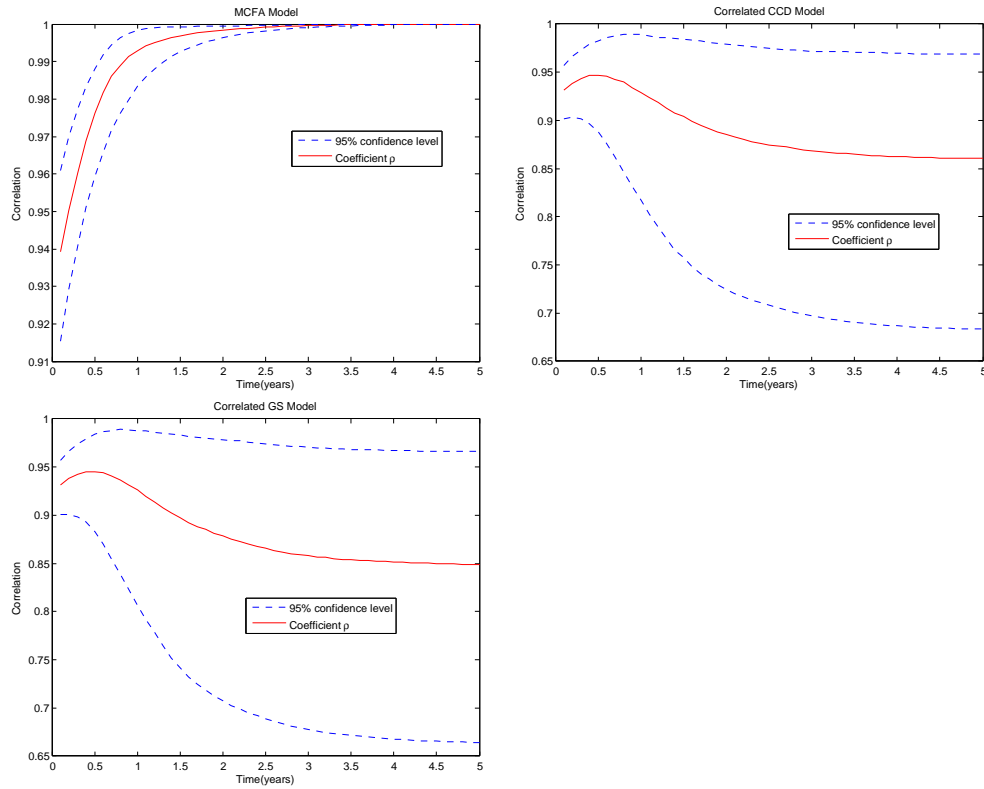

Figure 8: Correlation term structures (instantaneous futures correlation vs. time to maturity) of the WTI and Brent crude oil pair for the MCFA, correlated CCD, and correlated GS models. The 95% confidence levels are obtained by bootstrapping the model parameters.
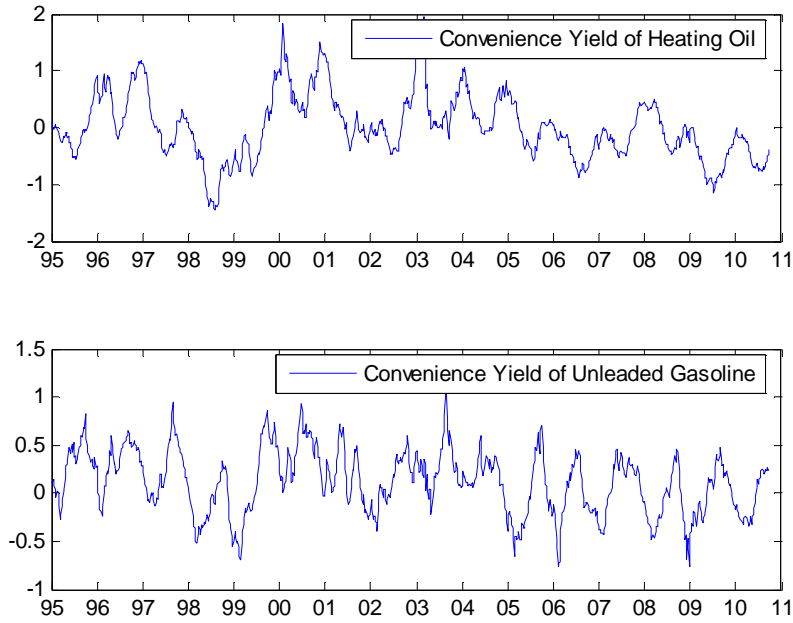
Figure 9: The implied convenience yield for the heating oil and unleaded gasoline. The implied convenience yields are from the MCFA model with the parameters obtained from Table 1.
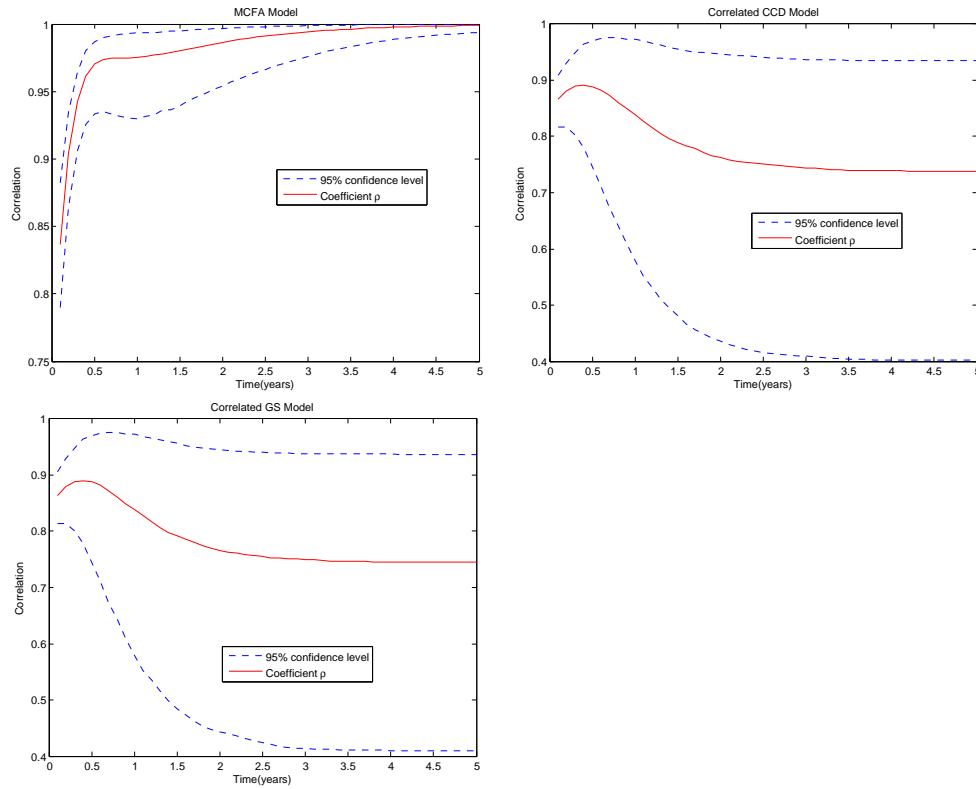

Figure 10: Correlation term structures (instantaneous futures correlation vs. time to maturity) of the heating oil and unleaded gasoline pair for the MCFA, correlated CCD, and correlated GS models. The 95% confidence levels are obtained by bootstrapping the model parameters.