Journal of International Money and Finance 42 (2014) 9–37

Contents lists available at ScienceDirect

Journal of International Money and Finance

journal homepage: www.elsevier.com/locate/jimf

Risk premia in crude oil futures prices☆

James D. Hamilton [a,*], Jing Cynthia Wu [b]

[a] Department of Economics, University of California, San Diego, United States
[b] Booth School of Business, University of Chicago, United States

Keywords:
Oil prices
Speculation
Futures risk premium
Affine term structure models

A B S T R A C T

If commercial producers or financial investors use futures contracts to hedge against commodity price risk, the arbitrageurs who take the other side of the contracts may receive compensation for their assumption of nondiversifiable risk in the form of positive expected returns from their positions. We show that this interaction can produce an affine factor structure to commodity futures prices, and develop new algorithms for estimation of such models using unbalanced data sets in which the duration of observed contracts changes with each observation. We document significant changes in oil futures risk premia since 2005, with the compensation to the long position smaller on average in more recent data. This observation is consistent with the claim that index-fund investing has become more important relative to commerical hedging in determining the structure of crude oil futures risk premia over time.

☆ Financial Support from the University of Chicago Booth School of Business is gratefully acknowledged. We thank Christiane Baumeister, Rob Engle, Lutz Kilian and anonymous reviewers for helpful comments on earlier drafts of this paper, as well as participants in the International Conference on Understanding International Commodity Price Fluctuations and seminar participants at the Bank of Canada, New York University, Econometric Society Annual Meeting at San Diego, Energy Policy Institute at Chicago workshop, Applied Time Series Econometrics Workshop at Federal Reserve Bank of St. Louis, Econometric Society North American Summer Meeting at Chicago, SNDE 21st Annual Symposium at Milan, Econometric Society Australasian Meeting at Melbourne, and Peking University Guanghua.

* Corresponding author.
E-mail addresses: jhamilton@ucsd.edu (J.D. Hamilton), Cynthia.Wu@chicagobooth.edu (J.C. Wu).

0261-5606/$ – see front matter
http://dx.doi.org/10.1016/j.jimonfin.2013.08.003

1. Introduction

Volatile oil prices have been drawing a lot of attention in recent years, with Hamilton (2009) for example suggesting that the oil price spike was a contributing factor in the recession of 2007–2009. There has been considerable interest in whether there is any connection between this volatility and the flow of dollars into commodity-index funds that take the long position in crude oil futures contracts. Recent empirical investigations of a possible link include Kilian and Murphy (2013), Tang and Xiong (2012), Buyuksahin and Robe (2011), Alquist and Gervais (2011), Mou (2010), Singleton (2011), Irwin and Sanders (2012), and Fattouh et al. (2013).

A separate question is the theoretical mechanism by which such an effect could arise in the first place. Keynes (1930) theory of normal backwardation proposed that if producers of the physical commodity want to hedge their price risk by selling futures contracts, then the arbitrageurs who take the other side of the contract may be compensated for assuming that risk in the form of a futures price below the expected future spot price. Empirical support for this view has come from Carter et al. (1983), Chang (1985), and De Roon et al. (2000), who interpreted the compensation as arising from the nondiversifiable component of commodity price risk, and from Bessembinder (1992), Etula (2013) and Acharya et al. (2013), who attributed the effect to capital limitations of potential arbitrageurs. In the modern era, buying pressure from commodity-index funds could exert a similar effect in the opposite direction, shifting the receipt of the risk premium from the long side to the short side of the contract.

In this paper we show that if arbitrageurs care about the mean and variance of their futures portfolio, then hedging pressure from commodity producers or index-fund investors can give rise to an affine factor structure to commodity futures prices. We do so by extending the models in Vayanos and Vila (2009) and Hamilton and Wu (2012a), which were originally used to describe how bond supplies affect relative yields, but are adapted in the current context to summarize how hedging demand would influence commodity futures prices. The result turns out to provide a motivation for specifications similar to the class of Gaussian affine term structure models originally developed by Vasicek (1977), Duffie and Kan (1996), Dai and Singleton (2002), Duffee (2002), and Ang and Piazzesi (2003) to characterize the relation between yields on bonds of different maturities. Related affine models have also been used to describe commodity futures prices by Schwartz (1997), Schwartz and Smith (2000), and Casassus and Collin-Dufresne (2006), among others.

In addition, this paper offers a number of methodological advances for use of this class of models to study commodity futures prices. First, we develop the basic relations directly for discrete-time observations, extending the contributions of Ang and Piazzesi (2003) to the setting of commodity futures prices. This allows a much more transparent mapping between model parameters and properties of observable OLS regressions. Second, we show how parameter estimates can be obtained directly from unbalanced data in which the remaining duration of observed contracts changes with each new observation, developing an alternative to the Kalman filter methodology used for this purpose by Cortazar and Naranjo (2006). Third, we show how the estimation method of Hamilton and Wu (2012b) provides diagnostic tools to reveal exactly where the model succeeds and where it fails to match the observed data.

We apply these methods to prices of crude oil futures contracts over 1990–2011. We document significant changes in risk premia in 2005 as the volume of futures trading began to grow significantly. While traders taking the long position in near contracts earned a positive return on average prior to 2005, that premium decreased substantially after 2005, becoming negative when the slope of the futures curve was high. This observation is consistent with the claim that historically commercial producers paid a premium to arbitrageurs for the privilege of hedging price risk, but in more recent periods financial investors have become natural counterparties for commercial hedgers. We also uncover seasonal variation of risk premia over the month, with changes as the nearest contract approaches expiry that cannot be explained from a shortening duration alone.

The plan of the paper is as follows. Section 2 develops the model, and Section 3 describes our approach to empirical estimation of parameters. Section 4 presents results for our baseline specification, while Section 5 presents results for a model allowing for more general variation as contracts near expiration. Conclusions are offered in Section 6.

2. Model

2.1. Role of arbitrageurs

Consider the incentives for a rational investor to become the counterparty to a commercial hedger or mechanical index-fund trader. We will refer to this rational investor as an arbitrageur, so named because the arbitrageur's participation guarantees that risk is priced consistently across all assets and futures contracts in equilibrium. Let F_{nt} denote the price of oil associated with an n-period futures contract entered into at date t. Let z_{nt} denote the arbitrageur's notional exposure (with $z_{nt} > 0$ denoting a long position and $z_{nt} < 0$ for short), so that z_{nt}/F_{nt} is the number of barrels purchased with n-period contracts. Following Duffie (1992, p. 39), we interpret a long position entered into at date t and closed at date $t+1$ as associated with a cash flow of zero at date t and $F_{n-1,t+1} - F_{nt}$ at date $t+1$. The arbitrageur's cash flow for period $t+1$ associated with the contemplated position z_{nt} is then $z_{nt}(F_{n-t,t+1} - F_{nt})/F_{nt}$. We assume the arbitrageur also takes positions q_{jt} in a set of other assets $j = 0, 1..., J$ with gross returns between t and $t+1$ denoted $\exp(r_{j,t+1})$ (so that the net return is approximately $r_{j,t+1}$) and where $r_{0,t+1}$ is assumed to be a risk-free yield. Then the arbitrageur's total wealth at $t+1$ will be

$$W_{t+1} = \sum_{j=0}^{J} q_{jt} \exp(r_{j,t+1}) + \sum_{n=1}^{N} z_{nt} \frac{F_{n-1,t+1} - F_{nt}}{F_{nt}}. \tag{1}$$

The arbitrageur is assumed to choose $\{q_{0t}, \ldots, q_{Jt}, z_{1t}, \ldots, z_{nt}\}$ so as to maximize[1]

$$E_t(W_{t+1}) - (\gamma/2)\mathrm{Var}_t(W_{t+1}) \tag{2}$$

subject to $\sum_{j=0}^{J} q_{jt} = W_t$.

We posit the existence of a vector of factors x_t that jointly determine all returns, which we assume follows a Gaussian vector autoregression (VAR)[2]:

$$x_{t+1} = c + \rho x_t + \Sigma u_{t+1} \quad u_t \sim \text{i.i.d. } N(0, I_m). \tag{3}$$

Log commodity prices and returns are assumed to be affine functions of these factors

$$f_{nt} = \log F_{nt} = \alpha_n + \beta_n' x_t \quad n = 1, \ldots, N \tag{4}$$

$$r_{jt} = \xi_j + \psi_j' x_t \quad j = 1, \ldots, J.$$

Using a similar approximation to that in Hamilton and Wu (2012a), we show in Appendix A that under these assumptions,

$$\begin{aligned} E_t(W_{t+1}) \approx\ & q_{0t}(1 + r_{0,t+1}) + \sum_{j=1}^{J} q_{jt}\left[1 + \xi_j + \psi_j'(c + \rho x_t) + (1/2)\psi_j'\Sigma\Sigma'\psi_j\right] \\ & + \sum_{n=1}^{N} z_{nt}\left[\alpha_{n-1} + \beta_{n-1}'(c + \rho x_t) - \alpha_n - \beta_n' x_t + (1/2)\beta_{n-1}'\Sigma\Sigma'\beta_{n-1}\right] \end{aligned} \tag{5}$$

$$\mathrm{Var}_t(W_{t+1}) \approx \left(\sum_{j=1}^{J} q_{jt}\psi_j' + \sum_{n=1}^{N} z_{nt}\beta_{n-1}'\right)\Sigma\Sigma'\left(\sum_{j=1}^{J} q_{jt}\psi_j + \sum_{\ell=1}^{N} z_{\ell t}\beta_{\ell-1}\right). \tag{6}$$

[1] It is trivial to extend this to adding positions in futures contracts for a number of alternative commodities. We discuss here the case of the single commodity oil for notational simplicity.

[2] The assumption of Gaussian homoskedastic errors greatly simplifies the estimation because it implies that parameters of the reduced-form representation of the model can be optimally estimated using simple OLS. For an extension of this approach to the case of non-Gaussian factors with time-varying variances, see Creal and Wu (2013).

The first-order conditions for the arbitrageur's positions satisfy

$$\frac{\partial E_t(W_{t+1})}{\partial q_{jt}} = 1 + r_{0,t+1} + (\gamma/2)\frac{\partial \mathrm{Var}_t(W_{t+1})}{\partial q_{jt}} \quad j = 1, \ldots, J$$
$$\frac{\partial E_t(W_{t+1})}{\partial z_{nt}} = (\gamma/2)\frac{\partial \mathrm{Var}_t(W_{t+1})}{\partial z_{nt}} \quad n = 1, \ldots, N.$$

Under (5) and (6) these become

$$\xi_j + \psi_j'(c + \rho x_t) + (1/2)\psi_j'\Sigma\Sigma'\psi_j = r_{0,t+1} + \psi_j'\lambda_t$$
$$\alpha_{n-1} + \beta_{n-1}'(c + \rho x_t) - \alpha_n - \beta_n' x_t + (1/2)\beta_{n-1}'\Sigma\Sigma'\beta_{n-1} = \beta_{n-1}'\lambda_t \tag{7}$$

for

$$\lambda_t = \gamma\Sigma\Sigma'\left(\sum_{j=1}^{J} q_{jt}\psi_j + \sum_{\ell=1}^{N} z_{\ell t}\beta_{\ell-1}\right). \tag{8}$$

Suppose we conjecture that in equilibrium the positions q_{jt}, z_{nt} selected by arbitrageurs are themselves affine functions of the vector of factors, so that

$$\lambda_t = \lambda + \Lambda x_t. \tag{9}$$

Then (7) requires

$$\beta_n' = \beta_{n-1}'\rho - \beta_{n-1}'\Lambda \tag{10}$$

$$\alpha_n = \alpha_{n-1} + \beta_{n-1}'c + (1/2)\beta_{n-1}'\Sigma\Sigma'\beta_{n-1} - \beta_{n-1}'\lambda. \tag{11}$$

From (5), the left side of (7) is the approximate expected return to a \$1 long position in an n-period contract entered at date t. Equation (7) thus characterizes equilibrium expected returns in terms of the price of risk λ_t:

$$E_t\left(\frac{F_{n-1,t+1} - F_{nt}}{F_{nt}}\right) \approx \beta_{n-1}'\lambda_t. \tag{12}$$

In the special case of risk-neutral arbitrageurs ($\gamma = 0$), from (8) we would have $\lambda = 0$ and $\Lambda = 0$ in (9).

Note that this framework allows for all kinds of factors (as embodied in the unobserved values of x_t) to influence commodity futures prices through Equation (4), including interest rates, fundamentals affecting supply and demand, and factors that might influence risk premia in other asset markets. If we consider physical inventory as another possible asset q_{jt}, this may help offset the risks associated with futures positions $z_{\ell t}$ as described in Equation (8) and could also be an element of the hypothesized factor vector x_t. We will demonstrate below that it is not necessary to have direct observations on the factor vector x_t itself in order to make use of the model's primary empirical implications (10) and (11). Instead, these restrictions can be represented solely in terms of implications for the dynamic behavior of the prices of different commodity-futures contracts that have to hold as a result of the factor structure itself and the behavior of the arbitrageurs. Moreover, we will see that it is possible to estimate the risk-pricing parameters λ and Λ solely on the basis of any predictabilities in the returns from positions in commodity-futures contracts.[3]

[3] Alternatively, one can try to make use of direct observations on the positions of commodity index-fund investors as we do in Hamilton and Wu (2012c).

The recursions (10) and (11) can equivalently be viewed as the equilibrium conditions that would result if risk-neutral arbitrageurs were to regard the factor dynamics as being governed not by (3) but instead by

$$x_{t+1} = c^Q + \rho^Q x_t + \Sigma u^Q_{t+1} \tag{13}$$

$$c^Q = c - \lambda \tag{14}$$

$$\rho^Q = \rho - \Lambda \tag{15}$$

$$u^Q_{t+1} \overset{Q}{\sim} N(0, I_m).$$

The recursions (10) and (11) that characterize the relation between the prices of futures contracts of different maturities will be recognized as similar to those that have been developed in the affine term structure literature[4] to characterize the relations that should hold in equilibrium between the interest rates on assets of different maturities. In addition to providing a derivation of how these relations can be obtained in the case of commodity futures contracts, the derivation above demonstrates how commercial hedging or commodity-index funds might be expected to influence commodity futures prices. An increase in the demand for long positions in contract n will require in equilibrium a price process in which arbitrageurs are persuaded to take a corresponding short position in exactly that amount. A larger absolute value of z_{nt} in turn will expose arbitrageurs to different levels of risk which would change the equilibrium compensation to risk λ_t according to equation (8). Again, from (8) and (9), these index traders could be responding through an affine function to interest rates or other economic fundamentals. What matters is that this behavior causes the net risk exposure of arbitrageurs λ_t to be an affine function of the factors in equilibrium. In the following subsection we illustrate this potential effect using a simple example.

2.2. *Example of the potential role of index-fund traders*

Suppose there are some investors who always want to have a long position in the 2-period contract, regardless of anything happening to fundamentals. At the start of each new period, these investors close out their previous position (which is now a 1-period contract) and replace it with a new long position in what is now the current 2-period contract.[5] Let the scalar K_t denote the notional value of 2-period contracts that investors want to buy in period t, and suppose this evolves exogenously according to

$$K_t = c^K + \rho^K K_{t-1} + \Sigma^K u^K_t. \tag{16}$$

If investors and arbitrageurs are the only participants in the market, then equilibrium futures prices must be such as to persuade arbitrageurs to take the opposite side of the investors. Thus arbitrageurs are always short the two-period contract, close that position when it becomes a 1-period contract, take the short side of the new 2-period contract, and have zero net exposure to any other contract in equilibrium. In other words, the process for $\{f_{nt}\}_{n=0}^N$ must be such that (7) and (8) are satisfied with

---

[4] Our recursions (10) and (11) are essentially the same as Equation (17) in Ang and Piazzesi (2003), with the important difference being that their recursion for the intercept adds a term δ_0 for each n, corresponding to the interest earned each period. No such term appears in our expression because there is no initial capital invested. Another minor notational difference is that our λ corresponds to their $\Sigma\lambda_0$ while our Λ corresponds to their $\Sigma\lambda_1$. An advantage of our notation in the current setting is that our λ_t is then measured in the same units as x_t and is immediately interpreted as the direct adjustment to c and ρ that results from risk aversion by arbitrageurs.

[5] In the case of crude oil contracts, what typically happens is that the commodity-index fund takes a long position from a swap dealer which in turn hedges its exposure by taking a long position in an organized exchange contract. We view the swap fund in such an arrangement as simply an intermediary, with the ultimate demand for the long position (K_t) coming from the commodity-index fund and the index-fund's ultimate counterparty being the short on the organized exchange contract (z_{2t}).

$$z_{nt} = \begin{cases} -K_t & \text{for } n = 2 \\ 0 & \text{otherwise} \end{cases}.$$

Suppose that arbitrageurs' only risk exposure comes from commodities ($q_{jt} = 0$ for $j = 1,\ldots,J$). Then from (8), in equilibrium we will have

$$\lambda_t = -\gamma\Sigma\Sigma'\beta_1 K_t. \tag{17}$$

Suppose that the spot price depends solely on a scalar "fundamentals" factor x_t^*:

$$f_{0t} = x_t^* \tag{18}$$

$$x_t^* = c^* + \rho^* x_{t-1}^* + \Sigma^* u_t^*.$$

We conjecture that in equilibrium, the factor x_t governing futures prices includes both fundamentals and the level of index-fund investment, $x_t = (x_t^*, K_t)'$, with (18) implying $\beta_0' = (1,0)$ and the factor evolving according to

$$x_t = c + \rho x_{t-1} + \Sigma u_t$$

or written out explicitly,

$$\begin{bmatrix} x_t^* \\ K_t \end{bmatrix} = \begin{bmatrix} c^* \\ c^K \end{bmatrix} + \begin{bmatrix} \rho^* & 0 \\ 0 & \rho^K \end{bmatrix}\begin{bmatrix} x_{t-1}^* \\ K_{t-1} \end{bmatrix} + \begin{bmatrix} \Sigma^* & 0 \\ 0 & \Sigma^K \end{bmatrix}\begin{bmatrix} u_t^* \\ u_t^K \end{bmatrix}.$$

We can then recognize (17) as a special case of (9) with

$$\lambda = 0$$

$$\underset{(2\times 2)}{\Lambda} = \begin{bmatrix} 0 & -\gamma\Sigma\Sigma'\beta_1 \end{bmatrix}.$$

Hence

$$\rho^Q = \begin{bmatrix} \rho^* & 0 \\ 0 & \rho^K \end{bmatrix} + [\,0 \quad \gamma\Sigma\Sigma'\beta_1\,] \tag{19}$$

$$\begin{aligned} \beta_1' &= \beta_0'\rho^Q \\ &= [\,1 \quad 0\,]\rho^Q \\ &= \begin{bmatrix} \rho^* & \gamma\begin{bmatrix} (\Sigma^*)^2 & 0 \end{bmatrix}\beta_1 \end{bmatrix} \end{aligned}$$

$$\beta_1 = \begin{bmatrix} \rho^* \\ \gamma\rho^*(\Sigma^*)^2 \end{bmatrix}. \tag{20}$$

Assuming $\rho^* > 0$ and $K_t > 0$, the effect of index-fund buying of the 2-period contract is also to increase the price of a 1-period contract. The reason is that the 2-period contract that the arbitrageurs are currently being induced to short exposes the arbitrageurs to risk associated with uncertainty about the value of x_{t+1}^*. The 1-period contract is also exposed to risk from x_{t+1}^*. If a 1-period contract purchased at t provided zero expected return, arbitrageurs would want to go long the 1-period contract in order to diversify their risk associated with being short the 2-period contract. But there is no counterparty who wants to short the 1-period contract, so equilibrium requires a price f_{1t} such that someone shorting the 1-period contract would also have a positive expected return, earned in the form of a higher price for f_{1t}.

Table 1
Weekly durations associated with monthly contracts at different points in time.

j	$k = 0$	$k = 1$	$k = 2$
1	3	7	11
2	2	6	10
3	1	5	9
4	0	4	8

For specified week of the month j and months until the contract expires k, table entry indicates weeks n remaining until expiry.

Substituting (20) into (19), we now know ρ^Q and can calculate $\beta_n = (\rho^{Q'})^n \beta_0$ for each n. Thus investment buying does not matter for f_{0t} but does affect every f_{nt} for $n > 0$, through the same mechanism as operates on the 1-period contract. In particular, from (12),

$$E_t\left(\frac{F_{n-1,t+1} - F_{nt}}{F_{nt}}\right) \approx -\gamma\beta_1'\left(\rho^Q\right)^{n-2}\Sigma\Sigma'\beta_1 K_t,$$

which in general has the opposite sign of K_t for all n; someone would earn a positive expected return by taking the short position in a contract of any duration.

2.3. Empirical implementation

There are two general strategies for empirical implementation of this framework. The first is to make direct use of data on the positions of different types of traders. Hamilton and Wu (2012c) use this approach to study agricultural futures prices. Unfortunately, the data publicly available on trader positions in crude oil futures contracts have some serious problems (see the discussion in Irwin and Sanders (2012) and Hamilton and Wu (2012c)). An alternative approach, which we adopt for purposes of modeling crude oil futures prices in this paper, is to infer the factors x_t based on the behavior of the futures prices themselves. In this case, risk premia are identified from differences between observed futures prices and a rational expectation of future prices. We will use the framework to characterize the dynamic behavior of risk premia and their changes over time.

For purposes of empirical estimation we interpret t as describing weekly intervals. This allows us to capture some key calendar regularities in the data without introducing an excessive number of parameters. NYMEX crude oil futures contracts expire on the third business day prior to the 25th calendar day of the month prior to the month on which the contract is written. To preserve the important calendar structure of the raw data, we divide the "month" leading up to a contract expiry into four "weeks", defined as follows:

week 1 ends on the last business day of the previous calendar month
week 2 ends on the 5th business day of the current calendar month
week 3 ends on the 10th business day of the current calendar month
week 4 ends on the day when the near contract expires

Associated with any week t is an indicator $j_t \in \{1, 2, 3, 4\}$ of where in the month week t falls.

Our estimation uses the nearest three contracts. If we interpret the price at expiry as an $n = 0$ week-ahead contract, the observation y_t for week t would be characterized using the notation of Section 2 as follows:

$$y_t = \begin{cases} \left(f_{3t}, f_{7t}, f_{11,t}\right)' & \text{if } j_t = 1 \\ \left(f_{2t}, f_{6t}, f_{10,t}\right)' & \text{if } j_t = 2 \\ \left(f_{1t}, f_{5t}, f_{9t}\right)' & \text{if } j_t = 3 \\ \left(f_{0t}, f_{4t}, f_{8t}\right)' & \text{if } j_t = 4 \end{cases}.$$

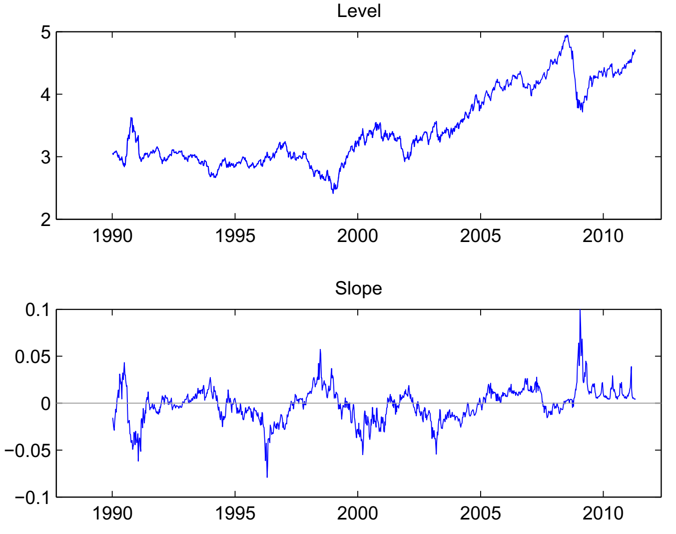

Fig. 1. Data used in the analysis. Weekly observations (with specification of weeks as given in text), January 1990 to June 2011. Top panel: first element of y_{1t} (the average of the log prices of second and third contracts). Bottom panel: second element of y_{1t} (the difference between the log price of third and second contracts).

Table 1 summarizes the relation between the weekly indicator (j), months until expiry is reached (k), and weeks remaining until expiry (n). This feature that the maturity of observed contracts changes with each observation t is one reason that much of the research with commodity futures contracts has used monthly data. However, in our application a key interest is in the higher-frequency movements and specific calendar effects. Fortunately, the framework developed in Section 2 gives us an exact description of the likelihood function for the data as actually observed, as we now describe.

We will assume that there are two underlying factors (that is, x_t is 2×1) Since (4) implies that each element of the (3×1) vector y_t could be written as an exact linear function of x_t, the system as written is stochastically singular – according to the model, the third element of y_t should be given by an exact linear combination of the first two. This issue also commonly arises in studies of the term structure of interest rates. A standard approach in that literature[6] is to assume that some elements or linear combinations of y_t differ from the magnitude predicted in (4) by a measurement or specification error. In the results reported below, we assume that the $k = 1$ - and 2- month contracts are priced exactly as the model predicts. It is helpful for purposes of interpreting parameter estimates to summarize the information in these contracts in terms of the average level of the two prices, which we will associate with the first factor in the system, and spread between them, which we will associate with the second factor:

$$y_{1t} = H_1 y_t$$

$$H_1 = \begin{bmatrix} 0 & (1/2) & (1/2) \\ 0 & -1 & 1 \end{bmatrix}. \tag{21}$$

The two elements of y_{1t} are plotted in Fig. 1.

We assume that the model correctly characterizes these two observed magnitudes. Since $y_t = (f_{4-j_t}, f_{8-j_t}, f_{12-j_t})'$, this implies that

[6] See for example Chen and Scott (1993), Ang and Piazzesi (2003) and Joslin et al. (2011). The observable implications of this assumption are explored in detail in Hamilton and Wu (2013).

$$y_{1t} = A_{1,j_t} + B_{1,j_t} x_t \tag{22}$$

$$\underset{(2\times1)}{A_{1,j_t}} = H_1 \begin{bmatrix} \alpha_{4-j_t} \\ \alpha_{8-j_t} \\ \alpha_{12-j_t} \end{bmatrix} \quad \text{for } j_t = 1,2,3, \text{ or } 4$$

$$\underset{(2\times2)}{B_{1,j_t}} = H_1 \begin{bmatrix} \beta'_{4-j_t} \\ \beta'_{8-j_t} \\ \beta'_{12-j_t} \end{bmatrix} \quad \text{for } j_t = 1,2,3, \text{ or } 4. \tag{23}$$

We will use the notational convention that if $j_t = 1$, then $A_{1,j_{t-1}} = A_{14}$.

If B_{1j} is invertible, the dynamics of the observed vector y_{1t} can be characterized by substituting (22) into (3):

$$y_{1t} = A_{1,j_t} + B_{1,j_t} c + B_{1,j_t} \rho \left[B^{-1}_{1,j_{t-1}} \left(y_{1,t-1} - A_{1,j_{t-1}} \right) \right] + B_{1,j_t} \Sigma u_t.$$

Since u_t is independent of $\{y_{t-1}, y_{t-2}, \ldots, y_0\}$, this means that the density of y_{1t} conditional on all previous observations is characterized by a VAR(1) with seasonally varying parameters:

$$y_{1t} | y_{t-1}, y_{t-2}, \ldots, y_0 \sim N\left(\phi_{j_t} + \Phi_{j_t} y_{1,t-1}, \Omega_{j_t} \right) \tag{24}$$

$$\Omega_{j_t} = B_{1,j_t} \Sigma \Sigma' B'_{1,j_t}$$

$$\Phi_{j_t} = B_{1,j_t} \rho B^{-1}_{1,j_{t-1}}$$

$$\phi_{j_t} = A_{1,j_t} + B_{1,j_t} c - \Phi_{j_t} A_{1,j_{t-1}}.$$

Note that the predicted seasonal parameter variation arises from the fact that the number of weeks remaining until expiry of the observed contracts changes with each new week.

We postulate that the nearest contract, which we write as

$$y_{2t} = H_2 y_t$$

$$H_2 = [1 \quad 0 \quad 0], \tag{25}$$

differs from the value predicted by the framework by a measurement or specification error with mean zero and variance σ^2_{e,j_t}:

$$y_{2t} = A_{2,j_t} + B_{2,j_t} x_t + \sigma_{e,j_t} u_{e,t}$$

$$\underset{(1\times1)}{A_{2,j_t}} = H_2 \begin{bmatrix} \alpha_{4-j_t} \\ \alpha_{8-j_t} \\ \alpha_{12-j_t} \end{bmatrix} \quad \text{for } j = 1,2,3,4$$

$$\underset{(1\times2)}{B_{2,j_t}} = H_2 \begin{bmatrix} \beta'_{4-j_t} \\ \beta'_{8-j_t} \\ \beta'_{12-j_t} \end{bmatrix} \quad \text{for } j_t = 1,2,3, \text{ or } 4.$$

If the measurement error u_{et} is independent of past observations, this gives the conditional distribution

$$y_{2t}|y_{1t}, y_{t-1}, y_{t-2}, \ldots, y_0 \sim N\left(\gamma_{j_t} + \Gamma_{j_t} y_{1t}, \sigma^2_{e,j_t}\right) \tag{26}$$

$$\Gamma_{j_t} = B_{2,j_t} B^{-1}_{1,j_t} \tag{27}$$

$$\gamma_{j_t} = A_{2,j_t} - \Gamma_{j_t} A_{1,j_t}.$$

The density of y_t conditional on its own past history is thus the product of (24) with (26), meaning that the log likelihood for the full sample of observations $(y'_T, y'_{T-1}, \ldots, y'_1)'$ conditional on the initial observation y_0 is given by

$$\mathcal{L} = \sum_{t=1}^{T} \left[\log g\left(y_{1t}; \phi_{j_t} + \Phi_{j_t} y_{1,t-1}, \Omega_{j_t}\right) + \log g\left(y_{2t}; \gamma_{j_t} + \Gamma_{j_t} y_{1t}, \sigma^2_{e,j_t}\right]\right. \tag{28}$$

where $g(y; \mu, \Omega)$ denotes the multivariate Normal density with mean μ and variance Ω evaluated at the point y.

3. Estimation

3.1. Unrestricted reduced form

The traditional approach to estimation of these kind of models would be to maximize the likelihood function with respect to the unknown structural parameters. However, Hamilton and Wu (2012b) demonstrate that there can be big benefits from using an estimator that turns out to be asymptotically equivalent to MLE but is derived from simple OLS regressions. To understand this estimator, consider first how we would maximize the likelihood if we thought of ϕ_j, Φ_j, Ω_j, γ_j, Γ_j, and σ_{ej} in the above representation as completely unrestricted parameters rather than the particular values implied by the structural model presented above. From this perspective, the log likelihood (28) could be written

$$\mathcal{L}(\phi_1, \Phi_1, \Omega_1, \gamma_1, \Gamma_1, \sigma_{e1}, \ldots, \phi_4, \Phi_4, \Omega_4, \gamma_4, \Gamma_4, \sigma_{e4}) = \sum_{j=1}^{4} \mathcal{L}_{1j}\left(\phi_j, \Phi_j, \Omega_j\right) + \sum_{j=1}^{4} \mathcal{L}_{2j}\left(\gamma_j, \Gamma_j, \sigma_{ej}\right) \tag{29}$$

$$\mathcal{L}_{1j}\left(\phi_j, \Phi_j, \Omega_j\right) = \sum_{t=1}^{T} \delta(j_t = j) \log g\left(y_{1t}; \phi_j + \Phi_j y_{1,t-1}, \Omega_j\right)$$

$$\log g\left(y_{1t}; \phi_j + \Phi_j y_{1,t-1}, \Omega_j\right) = -\log 2\pi - (1/2)\log|\Omega_j| - (1/2)\left(y_{1t} - \phi_j - \Phi_j y_{1,t-1}\right)' \Omega_j^{-1} \left(y_{1t} - \phi_j - \Phi_j y_{1,t-1}\right)$$

$$\mathcal{L}_{2j}\left(\gamma_j, \Gamma_j, \sigma_{ej}\right) = \sum_{t=1}^{T} \delta(j_t = j) \log g\left(y_{2t}; \gamma_j + \Gamma_j y_{1t}, \sigma^2_{ej}\right)$$

$$\log g\left(y_{2t}; \gamma_j + \Gamma_j y_{1t}, \sigma^2_{ej}\right) = -(1/2)\log 2\pi - (1/2)\log \sigma^2_{ej} - \frac{\left(y_{2t} - \gamma_j - \Gamma_j y_{1t}\right)^2}{2\sigma^2_{ej}}$$

where for example $\delta(j_t = 1)$ is 1 if t is in the first week of the month and is zero otherwise. It is clear that the unconstrained likelihood function is in fact maximized by a series of OLS regressions. To estimate the parameters in block j, we collect all observations whose left-hand variable is in the jth week of the month, and simply perform OLS regressions on what now looks like a monthly data set.

Specifically, to estimate $(\phi_j, \Phi_j, \Omega_j)$ for a particular j, we associate month τ with an observed monthly-frequency vector $y^{\dagger}_{1,j,\tau}$ defined as follows. For illustration, consider $j = 1$ and suppose that τ corresponds to the month spanned by the last week of December and first 3 weeks of January. The first element of $y^{\dagger}_{1,1,\tau|\mathrm{month}(\tau)=\mathrm{Jan}}$ is the average of the log prices of the March and April contracts as of the last business day of December. The second element of $y^{\dagger}_{1,1,\tau|\mathrm{month}(\tau)=\mathrm{Jan}}$ is based on the log price of the April contract on the last day of December minus the log price of March contract. For $j = 1$ and general τ,

$$y^{\dagger}_{1,1,\tau} = H_1 \begin{bmatrix} f_{3,t(\tau)} \\ f_{7,t(\tau)} \\ f_{11,t(\tau)} \end{bmatrix}$$

for H_1 given in (21) and where $t(\tau)$ denotes the week t associated with month τ. The explanatory variables in these $j = 1$ block regressions consist of a constant, the average log prices of the February and March contracts on the day in December when the January contract expired, and the spread between the March price and February price at the December expiry of the January contract:

$$x^{\dagger}_{1,1,\tau} = \begin{bmatrix} 1 \\ H_1 \begin{bmatrix} f_{0,t(\tau)-1} \\ f_{4,t(\tau)-1} \\ f_{8,t(\tau)-1} \end{bmatrix} \end{bmatrix}. \tag{30}$$

Consider the estimates from OLS regression of $y^{\dagger}_{1,1,\tau}$ on $x^{\dagger}_{1,1,\tau}$,

$$\begin{bmatrix} \hat{\phi}_1 & \hat{\Phi}_1 \end{bmatrix} = \left(\sum_{\tau=1}^{\mathcal{T}} y^{\dagger}_{1,1,\tau} x^{\dagger\prime}_{1,1,\tau} \right) \left(\sum_{\tau=1}^{\mathcal{T}} x^{\dagger}_{1,1,\tau} x^{\dagger\prime}_{1,1,\tau} \right)^{-1}$$

$$\hat{\Omega}_1 = \mathcal{T}^{-1} \sum_{\tau=1}^{\mathcal{T}} \left(y^{\dagger}_{1,1,\tau} - \begin{bmatrix} \hat{\phi}_1 & \hat{\Phi}_1 \end{bmatrix} x^{\dagger}_{1,1,\tau} \right) \left(y^{\dagger}_{1,1,\tau} - \begin{bmatrix} \hat{\phi}_1 & \hat{\Phi}_1 \end{bmatrix} x^{\dagger}_{1,1,\tau} \right)'$$

where $\mathcal{T}$ denotes the number of months in the sample. These estimates maximize the log likelihood (29) with respect to $\{\phi_1, \Phi_1, \Omega_1\}$

For $j = 2$ we regress $y^{\dagger}_{1,2,\tau}$ (whose first element, for example, would be the average of the March and April contracts as of the fifth business day in January) on $x^{\dagger}_{1,2,\tau}$ (e.g., a constant and the level and spread as of the last day of December),

$$\begin{bmatrix} \hat{\phi}_2 & \hat{\Phi}_2 \end{bmatrix} = \left(\sum_{\tau=1}^{\mathcal{T}} y^{\dagger}_{1,2,\tau} x^{\dagger\prime}_{1,2,\tau} \right) \left(\sum_{\tau=1}^{\mathcal{T}} x^{\dagger}_{1,2,\tau} x^{\dagger\prime}_{1,2,\tau} \right)^{-1}$$

$$\hat{\Omega}_2 = \mathcal{T}^{-1} \sum_{\tau=1}^{\mathcal{T}} \left(y^{\dagger}_{1,2,\tau} - \begin{bmatrix} \hat{\phi}_2 & \hat{\Phi}_2 \end{bmatrix} x^{\dagger}_{1,2,\tau} \right) \left(y^{\dagger}_{1,2,\tau} - \begin{bmatrix} \hat{\phi}_2 & \hat{\Phi}_2 \end{bmatrix} x^{\dagger}_{1,2,\tau} \right)',$$

to obtain $\hat{\phi}_2$, $\hat{\Phi}_2$, and $\hat{\Omega}_2$. Similar separate monthly regressions of the 1- and 2-month prices in the third or fourth week of each month on their values the week before produce $\{\hat{\phi}_j, \hat{\Phi}_j, \hat{\Omega}_j\}$ for $j = 3$ or 4.

Likewise, note that the components of $\sum_{j=1}^{4} \mathscr{L}_{2j}(\gamma_j, \Gamma_j, \sigma_e)$ take the form of regressions in which the residuals are uncorrelated across blocks, meaning full-information maximum likelihood estimates of γ_j

and Γ_j are obtained by OLS regressions for individual j. For example, for $j = 1$ and τ corresponding to December–January, $y^{\dagger}_{2,j,\tau}$ is the price of the February contract on the last day of December,

$$y^{\dagger}_{2,1,\tau} = H_2 \begin{bmatrix} f_{3,t(\tau)} \\ f_{7,t(\tau)} \\ f_{11,t(\tau)} \end{bmatrix},$$

for H_2 in (25) and explanatory variables the level and slope as of the last day of December:

$$x^{\dagger}_{2,1,\tau} = \begin{bmatrix} 1 \\ H_1 \begin{bmatrix} f_{3,t(\tau)} \\ f_{7,t(\tau)} \\ f_{11,t(\tau)} \end{bmatrix} \end{bmatrix}.$$

The maximum likelihood estimates are given by

$$\begin{aligned} \begin{bmatrix} \widehat{\gamma}_j & \hat{\Gamma}_j \end{bmatrix} &= \left(\sum_{\tau=1}^{\mathcal{T}} y^{\dagger}_{2,j,\tau} x^{\dagger\prime}_{2,j,\tau} \right) \left(\sum_{\tau=1}^{\mathcal{T}} x^{\dagger}_{2,j,\tau} x^{\dagger\prime}_{2,j,\tau} \right)^{-1} \quad \text{for } j = 1,2,3,4 \\ \widehat{\sigma}^2_{ej} &= \mathcal{T}^{-1} \sum_{\tau=1}^{\mathcal{T}} \left(y^{\dagger}_{2,j,\tau} - \begin{bmatrix} \widehat{\gamma}_j & \hat{\Gamma}_j \end{bmatrix} x^{\dagger}_{2,j,\tau} \right)^2. \end{aligned} \tag{31}$$

3.2. *Structural estimation of the baseline model*

Now consider estimation of the underlying structural parameters of the model presented in Section 2. The key point to note is that the above OLS estimates $\{\widehat{\phi}_1, \widehat{\Phi}_1, \widehat{\Omega}_1, \widehat{\gamma}_1, \hat{\Gamma}_1, \widehat{\sigma}_{e1}, \ldots, \widehat{\phi}_4, \widehat{\Phi}_4, \widehat{\Omega}_4, \widehat{\gamma}_4, \hat{\Gamma}_4, \widehat{\sigma}_{e4}\}$ are sufficient statistics for inference about these parameters – anything that the full sample of data is able to tell us about the model parameters can be summarized by the values of these OLS estimates. The idea behind the minimum-chi-square estimation proposed by Hamilton and Wu (2012b) is to choose structural parameters that would imply reduced-form coefficients as close as possible to the unrestricted estimates, an approach that turns out to be asymptotically equivalent to full MLE.

Note that the model developed here specifies observed prices in terms of an unobserved factor vector x_t. There is an arbitrary normalization in any such system, in that if we were to multiply x_t by a nonsingular matrix and add a constant, the result would be observationally equivalent in terms of the implied likelihood for observed y_t.[7] Since we have treated the factors x_t as directly inferable from the values of y_{1t}, we normalize the factors so that they could be interpreted as the level and slope as of the date of expiry of the near-term contract:

$$x_t = H_1 y_t \quad \text{for } j_t = 4. \tag{32}$$

Recalling (22), this would be the case if

$$H_1 y_t = H_1 \begin{bmatrix} \alpha_0 \\ \alpha_4 \\ \alpha_8 \end{bmatrix} + H_1 \begin{bmatrix} \beta_0' \\ \beta_4' \\ \beta_8' \end{bmatrix} x_t \quad \text{for } j_t = 4. \tag{33}$$

Substituting (32) into (33), our chosen normalization thus calls for

$$x_t = H_1 \begin{bmatrix} \alpha_0 \\ \alpha_4 \\ \alpha_8 \end{bmatrix} + H_1 \begin{bmatrix} \beta_0' \\ \beta_4' \\ \beta_8' \end{bmatrix} x_t \quad \text{for } j_t = 4$$

[7] For further discussion of identification and normalization, see Hamilton and Wu (2012b).

$$H_1\begin{bmatrix}\beta_0'\\ \beta_4'\\ \beta_8'\end{bmatrix} = I_2 \tag{34}$$

$$H_1\begin{bmatrix}\alpha_0\\ \alpha_4\\ \alpha_8\end{bmatrix} = 0. \tag{35}$$

Since

$$\begin{bmatrix}f_{0t}\\ f_{4t}\\ f_{8t}\end{bmatrix} = \begin{bmatrix}\alpha_0\\ \alpha_4\\ \alpha_8\end{bmatrix} + \begin{bmatrix}\beta_0'\\ \beta_4'\\ \beta_8'\end{bmatrix}x_t,$$

our normalization could alternatively be described as $x_t = H_1(f_{0t}, f_{4t}, f_{8t})'$ for all t. Following Joslin et al. (2011) and Hamilton and Wu (2013), this can be implemented by defining ξ_1 and ξ_2 to be the eigenvalues of $\rho^Q = \rho - \Lambda$. Given this normalization and values for ξ_1, ξ_2, Σ, and α_0, we can then determine the values for ρ^Q, c^Q, and $\{\beta_n, \alpha_n\}_{n=0}^N$; details are provided in Appendix B. These along with ρ, c, and σ_e then provide everything we need to evaluate the likelihood function or to calculate what the predicted values for any of the unrestricted reduced-form coefficients ought to be.

Let θ denote the vector of unknown structural parameters, that is, the 16 elements of $\{\xi_1, \xi_2, \Sigma, \alpha_0, \rho, c, \sigma_{e1}, \sigma_{e2}, \sigma_{e3}, \sigma_{e4}\}$ for Σ lower triangular. Collect elements of the unrestricted OLS estimates in a vector $\widehat{\pi}$:

$$\widehat{\pi} = (\widehat{\pi}_\Phi', \widehat{\pi}_\Omega', \widehat{\pi}_\Gamma', \widehat{\pi}_\sigma)'$$

$$\widehat{\pi}_\Phi = \left(\left[\mathrm{vec}\left(\left[\widehat{\phi}_1 \quad \widehat{\Phi}_1\right]'\right)\right]', \ldots, \left[\mathrm{vec}\left(\left[\widehat{\phi}_4 \quad \widehat{\Phi}_4\right]'\right)\right]'\right)'$$

$$\widehat{\pi}_\Omega = \left(\left[\mathrm{vech}\left(\widehat{\Omega}_1\right)\right]', \ldots, \left[\mathrm{vech}\left(\widehat{\Omega}_4\right)\right]'\right)'$$

$$\widehat{\pi}_\Gamma = \left(\left[\mathrm{vec}\left(\left[\widehat{\gamma}_1 \quad \hat{\Gamma}_1\right]'\right)\right]', \ldots, \left[\mathrm{vec}\left(\left[\widehat{\gamma}_4 \quad \hat{\Gamma}_4\right]'\right)\right]'\right)'$$

$$\widehat{\pi}_\sigma = (\widehat{\sigma}_{e1}, \widehat{\sigma}_{e2}, \widehat{\sigma}_{e3}, \widehat{\sigma}_{e4})'.$$

Let $g(\theta)$ denote the corresponding predicted values for those coefficients from the model; specific values for the elements of $g(\theta)$ are summarized in Appendix C. The minimum-chi-square (MCS) estimate of θ is the value that minimizes

$$\mathcal{T}[\widehat{\pi} - g(\theta)]'\widehat{R}[\widehat{\pi} - g(\theta)] \tag{36}$$

for $\widehat{R}$ the information matrix associated with the OLS estimates $\widehat{\pi}$, which is also detailed in Appendix C. The MCS estimator has the same asymptotic distribution as the maximum likelihood estimator, but has a number of computational and interpretive advantages over MLE discussed in Hamilton and Wu (2012b). Because $\widehat{R}$ is block-diagonal with respect to π_σ, the MCS estimates of these parameters are given immediately by the OLS estimates (31). Hamilton and Wu (2012b) show that asymptotic standard errors can be estimated using

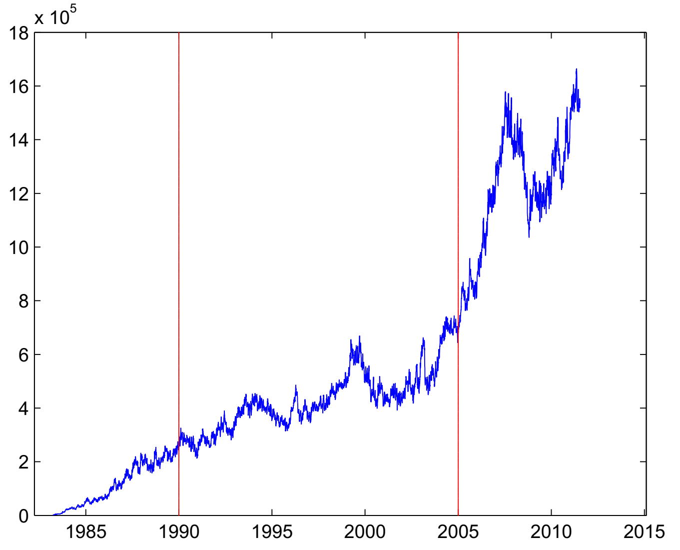

Fig. 2. Open Interest. Total number of outstanding crude oil futures contracts of all maturities, daily observations, March 3, 1983 to July 17, 2011. Vertical lines drawn at January 1, 1990 and January 1, 2005.

$$E\left(\widehat{\theta}-\theta_0\right)\left(\widehat{\theta}-\theta_0\right)'\simeq \mathcal{T}^{-1}\left(\widehat{G}'\widehat{R}\widehat{G}\right)^{-1}$$

$$\widehat{G} = \left.\frac{\partial g(\theta)}{\partial \theta'}\right|_{\theta=\widehat{\theta},}$$

which are identical to the usual asymptotic errors that would be obtained by taking second derivatives of the log likelihood function (28) with respect to θ.

4. Empirical results for the baseline model

Crude oil futures contracts were first traded on the New York Mercantile Exchange (NYMEX) in 1983. In the first few years, volume was much lighter than the more recent data, and we choose to begin our empirical analysis in January, 1990. Fig. 2 plots the total open interest on all NYMEX light sweet crude contracts. Volume expanded very quickly after 2004, in part in response to the increased purchases of futures contracts as a vehicle for financial diversification. Some researchers have suggested that participation in the markets by this new class of traders resulted in significant changes in the dynamic behavior of crude oil futures prices.[8] A likelihood ratio test (e.g., Hamilton (1994, p. 296)) of the null hypothesis that the coefficients of the unrestricted reduced form are constant over time against the alternative that all 52 parameters changed in January 2005 produces a $\chi^2(52)$ statistic of 181.96, which calls for dramatic rejection of the null hypothesis (p value of 2.2×10^{-16}). Since one of our interests in this paper is to document how futures price dynamics have changed over time, we conduct our analysis on two subsamples, the first covering January 1990 through December 2004, and the second January 2005 through June 2011.

The left panel of Table 2 reports minimum-chi-square estimates of the 16 elements of θ based on the first subsample. The eigenvalues of ρ, the matrix summarizing the objective P-measure persistence of

[8] See for example Alquist and Kilian (2007), Singleton (2011), Tang and Xiong (2012), Mou (2010), and Buyuksahin and Robe (2011).

Table 2
Pre-2005 parameter estimates for baseline model.

Estimated parameters			Implied parameters		
c	0.0102	0.0016	λ	0.0104	−0.0061*
	(0.0209)	(0.0029)		(0.0209)	(0.0029)
ρ	0.9974*	0.1839*	Λ	−0.0023	−0.0591
	(0.0067)	(0.0918)		(0.0067)	(0.0918)
	−0.0006	0.9301*		0.0020*	−0.0050
	(0.0009)	(0.0134)		(0.0009)	(0.0136)
ξ	0.9876*	0.9473*	$\lambda + \Lambda\bar{x}$	0.0037*	3.53e-005
	(0.0010)	(0.0044)		(0.0018)	(2.63e-004)
Σ	0.0449*	0			
	(0.0012)				
	−0.0038*	0.0047*			
	(0.0002)	(0.0001)			
α_0	0.0357*				
	(0.0007)				
π_σ	0.0099*	0.0081*			
	(0.0005)	(0.0004)			
	0.0105*	0.0201*			
	(0.0006)	(0.0011)			

Left panel: MCS estimates of elements of θ for data from January 1990 through December 2004 (asymptotic standard errors in parentheses). Right panel: assorted magnitudes of interest implied by value of $\widehat{\theta}$ (asymptotic standard errors in parentheses). * denotes statistically significant at the 5% level.

factors, are 0.9956 and 0.9319, implying that both level and slope are highly persistent, with similar estimates for their Q-measure counterparts (ξ_1 and ξ_2).

The differences between the P- and Q-measures, or implied characterization of λ_t, are reported in the right panel of Table 2.[9] The individual elements of λ and Λ are generally small and statistically insignificant. The last two entries of Table 2 report the elements of $\lambda + \Lambda\bar{x}$, where $\bar{x}$ is the average value for the level and spread over the sample. The positive value of 0.0037 for the first element of this vector suggests that an investor who was always long in the two contracts would on average have come out ahead over this period, an estimate that is just statistically significant at the 5% level.[10]

Table 3 reports parameter estimates for the later subsample, in which there appear to be significant differences in risk pricing from the earlier data. Most noteworthy is the large negative value for Λ_{12}. This signifies that when the spread (the second element of x_t) gets sufficiently high, a long position in the 1- and 2-month contracts would on average lose money. We also see from the last entry of Table 3 that the first element of $\lambda + \Lambda\bar{x}$ is smaller in the second subsample than in the first, and is no longer statistically significant. The average reward for taking long positions in the second subsample is not as evident as in the first subsample.

Fig. 3 plots our estimated values for $\lambda_t = \lambda + \Lambda x_t$ for each week t in our sample, along with 95% confidence intervals. The price of level risk (top panel) was uniformly positive up until 2006, but has often been negative since 2008. By contrast, slope risk (bottom panel) was typically not priced before 2004, whereas going long the 2-month contract and short the 1-month has frequently been associated with positive expected returns since then.[11]

[9] Standard errors for λ and Λ were obtained by reparameterizing the MCS estimation in terms of λ and Λ instead of c and ρ. The values of λ and Λ can be obtained analytically from $\lambda = c - c^Q$ and $\Lambda = \rho - \rho^Q$ with c^Q given by equation (47) and ρ^Q given by equation (42).

[10] Note from (12) and (34) that the expected return on a portfolio with equal weights on the second and third contracts is given by $(1/2)(\beta_4' + \beta_8')\lambda_t = [1 \quad 0]\lambda_t$ whose average value is the first element of $\lambda + \Lambda\bar{x}$.

[11] From (12) and (34), the expected return on a portfolio that is long the third contract and short the second is given by $(\beta_8' - \beta_4')\lambda_t = [0 \quad 1]\lambda_t$ whose average value is the second element of $\lambda + \Lambda\bar{x}$.

Table 3
Post-2005 parameter estimates for baseline model.

Estimated parameters			Implied parameters		
c	0.1802*	0.0164*	λ	0.1813*	0.0179*
	(0.0574)	(0.0070)		(0.0574)	(0.0070)
ρ	0.9600*	−0.3487	Λ	−0.0400*	−0.5892*
	(0.0131)	(0.2018)		(0.0131)	(0.2018)
	−0.0035*	0.8629*		−0.0039*	−0.0311
	(0.0016)	(0.0241)		(0.0016)	(0.0243)
ξ	1.0010*	0.8931*	$\lambda + \Lambda\bar{x}$	0.0028	0.0009*
	(0.0001)	(0.0047)		(0.0026)	(0.0003)
Σ	0.0439*	0			
	(0.0019)				
	−0.0021*	0.0049*			
	(0.0003)	(0.0002)			
α_0	−0.0086*				
	(0.0009)				
π_σ	0.0059*	0.0086*			
	(0.0005)	(0.0007)			
	0.0087*	0.0223*			
	(0.0007)	(0.0018)			

Left panel: MCS estimates of elements of θ for data from January 2005 through June 2011 (asymptotic standard errors in parentheses). Right panel: assorted magnitudes of interest implied by value of $\widehat{\theta}$ (asymptotic standard errors in parentheses). * denotes statistically significant at the 5% level.

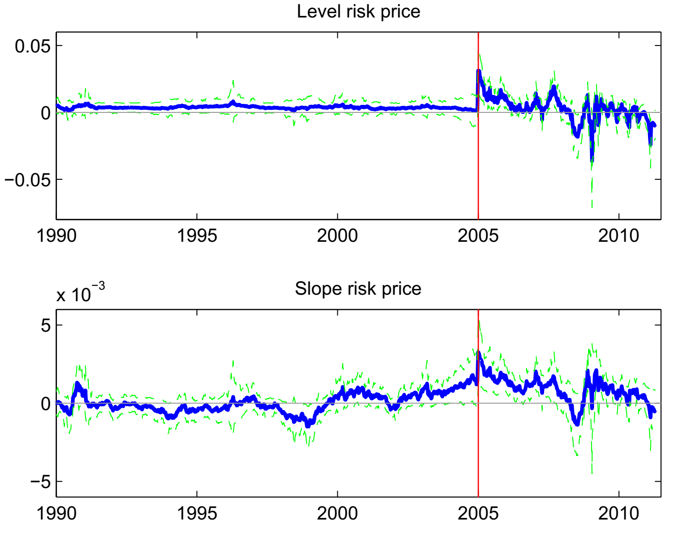

Fig. 3. Prices of factor risk. Top panel: first element of $\lambda + \Lambda x_t$ as estimated from baseline model, with sample split in 2005. Bottom panel: second element. Dashed lines indicate 95% confidence intervals.

Following Cochrane and Piazzesi (2009) and Bauer et al. (2012), another way to summarize the implications of these results is to calculate how different the log price of a given contract would be if there was no compensation for risk. To get this number, we calculate $\tilde{f}_{nt} = \tilde{\alpha}_n + \tilde{\beta}_n x_t$, where $\tilde{\beta}_n$ and $\tilde{\alpha}_n$ denote the values that would be obtained from the recursions (10) and (11) if Λ and λ were both set to zero. The value for the difference $\tilde{f}_{nt} - f_{nt}$ for $n = 8$ weeks is plotted in Fig. 4. In the absence of risk

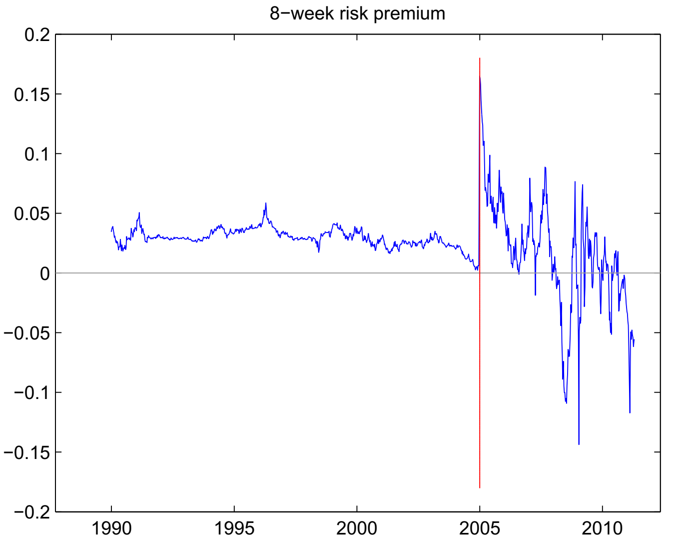

Fig. 4. Risk premium on 8-week futures contract. Plot of $\tilde{f}_{8t} - f_{8t}$ as estimated from the baseline model with sample split in 2005.

effects, an 8-week contract price would have been a few percent higher on average over the 1990–2004 subsample.[12] Since 2005, risk aversion has made a more volatile contribution, though the average effect is significantly smaller.

In terms of the framework proposed in Section 2 for interpreting these results, the positive average value for the first element of λ_t in the first subsample suggests that arbitrageurs were on average long in crude oil futures contracts over this period, accepting the positive expected earnings from their positions as compensation for providing insurance to sellers, who were presumably commercial producers who wanted to hedge their price risks by selling futures contracts. From that perspective, an increase in index fund buying could have been one explanation for why a long position in futures contracts no longer has a statistically significant positive return. In effect, index-fund buyers are serving as counterparty for commercial hedgers, and are willing to do so without the risk compensation that the position earned on average in the first subsample. The emerging positive return to a spreading position (positive average second element of λ_t in the second subsample) would be consistent with the view that arbitrageurs are buying and holding 2-month futures from oil producers, but then selling these positions and going short 1-month futures as they sell to index-fund investors.

As noted by Hamilton and Wu (2012b), another benefit of estimation by minimum chi square is that the optimized value for the objective function provides an immediate test of the overall framework. Under the null hypothesis that the model is correctly specified, the minimum value achieved for (36) has an asymptotic χ^2 distribution with degrees of freedom given by the number of overidentifying restrictions. The first column of Table 4 reports the value of this statistic for each of the two subsamples. The model is overwhelmingly rejected in either subsample.

Because the weighting matrix $\hat{R}$ in (49) is block-diagonal, it is easy to decompose these test statistics into components coming from the respective elements of π, as is done in subsequent columns of Table

[12] The average size of this risk premium is 2.9% for the first sample. This compares with an average realized 2-month ex post return over this period of 2.0% for the long position on a 3-month contract (that is, the average log value of the first contract minus the average log value of the third contract two months earlier), and an average difference between the first contract and third contract at the same date of 1.2% (that is, the futures curve sloped down on average with a slope of −1.2%). The last number is similar to the value reported by Alquist and Kilian (2010), who noted that the 3-month futures price was 1.1% below the spot price on average over 1987–2007. The difference between the average ex post return to the long position and the negative of the average slope results from the significantly higher price of oil at the end of the sample than at the beginning.

Table 4
χ^2 specification test and breakdown by individual components.

	χ^2	$d.f.$	p-value	π_Φ	π_Ω	π_Γ
Before 2005	86.57	36	4.73e-6	25.61	43.98	16.99
Since 2005	151.87	36	3.33e-16	120.01	17.22	14.64

χ^2: minimum value achieved for MCSE objective function. $d.f.$: degrees of freedom. p-value: probability of observing $\chi^2(d.f.)$ value this large. Last 3 columns: contribution to χ^2 of individual parameter blocks.

4. In the first subsample, about half of the value of the test statistic comes from the π_Ω block – the differences in the variability of the level and slope across different weeks of the month is more than can be explained by the fact that the maturities of observed contracts are changing week to week. The biggest problem in the second subsample come from the π_Φ block – unrestricted forecasts of the level and slope vary more week-to-week than is readily explained by differences in the maturities of the contracts.

It is also possible to look one parameter at a time at where the structural model misses. For each of the unrestricted reduced-form parameters π there is a corresponding prediction from the model $g(\theta)$ for what that value is supposed to be if the model is correct. Figs. 5 and 6 plot the unrestricted OLS estimates of the various elements of π along with their 95% confidence intervals for the first subsample. The thick red lines indicate the value the coefficient is predicted to have according to the structural parameters reported in Table 2. The biggest problems come from the fact that the model underpredicts the difficulty of forecasting the spread in weeks 1 and 3 (the lower left panel of Fig. 6). Figs. 7 and 8 provide the analogous plots for the second subsample. Here the biggest problems come from the fact that the equations one would want to use to forecast the spread in weeks 3 and 4 are quite different

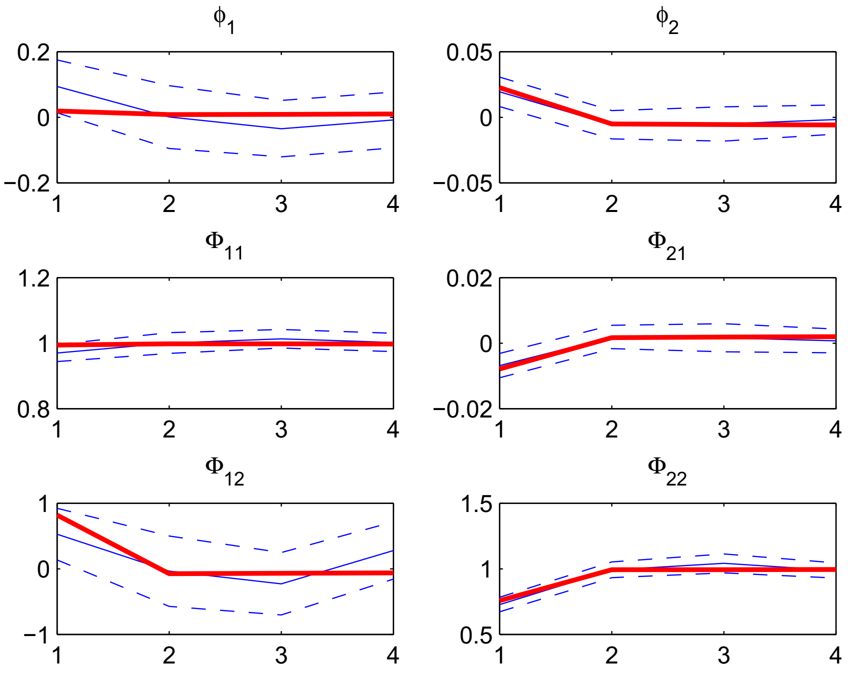

Fig. 5. $\pi_\Phi$before 2005. Light blue line: Unrestricted OLS estimates of coefficients for regression in which y_{1t} is the dependent variable, plotted as a function of week of the month. Dashed blue lines: 95% confidence intervals for unrestricted OLS estimates. Bold red line: predicted values for coefficients derived from baseline model. All estimates based on data January 1990 to December 2004. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

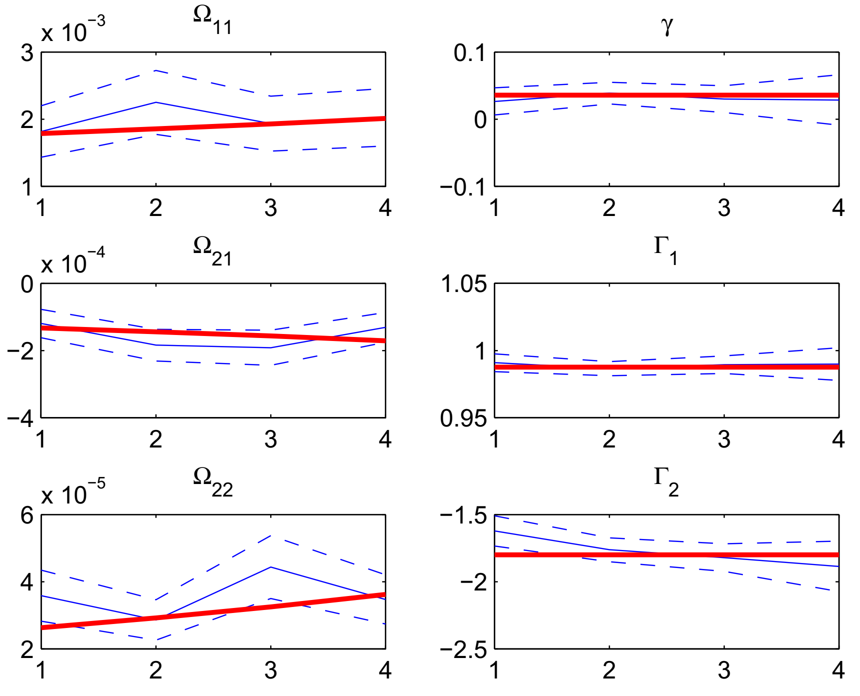

Fig. 6. π_Ω and π_Γ before 2005. First column: estimated elements of variance-covariance matrix for regression in which y_{1t} is the dependent variable, plotted as a function of week of the month. Second column: Estimated values of coefficients for regression in which y_{2t} is the dependent variable, plotted as a function of week of the month. In each panel, light blue lines are unrestricted OLS estimates, dashed blue lines are 95% confidence intervals for unrestricted OLS estimates, and bold red lines are predicted values for coefficients derived from baseline model. All estimates based on data January 1990 to December 2004. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

from those found for weeks 1 and 2 (see the right-hand column of Fig. 7). These considerations suggest looking at models that allow for more general seasonal variation than our baseline specification, which we explore in the next section.

5. Less restrictive seasonal models

5.1. Structural estimation

Here we consider a system in which the dynamic process followed by the factors is itself dependent on which week of the month we are looking at:

$$x_{t+1} = c_{j_t} + \rho_{j_t} x_t + \Sigma_{j_t} u_{t+1}.$$

If we hypothesize that the risk-pricing parameters also vary with the season,

$$\lambda_t = \lambda_{j_t} + \Lambda_{j_t} x_t$$

then the no-arbitrage conditions (10) and (11) generalize to

$$\beta_n' = \beta_{n-1}' \rho^Q_{j(n)} \tag{37}$$

$$\alpha_n = \alpha_{n-1} + \beta_{n-1}' c^Q_{j(n)} + (1/2)\beta_{n-1}' \Sigma_{j(n)} \Sigma_{j(n)}' \beta_{n-1}.$$

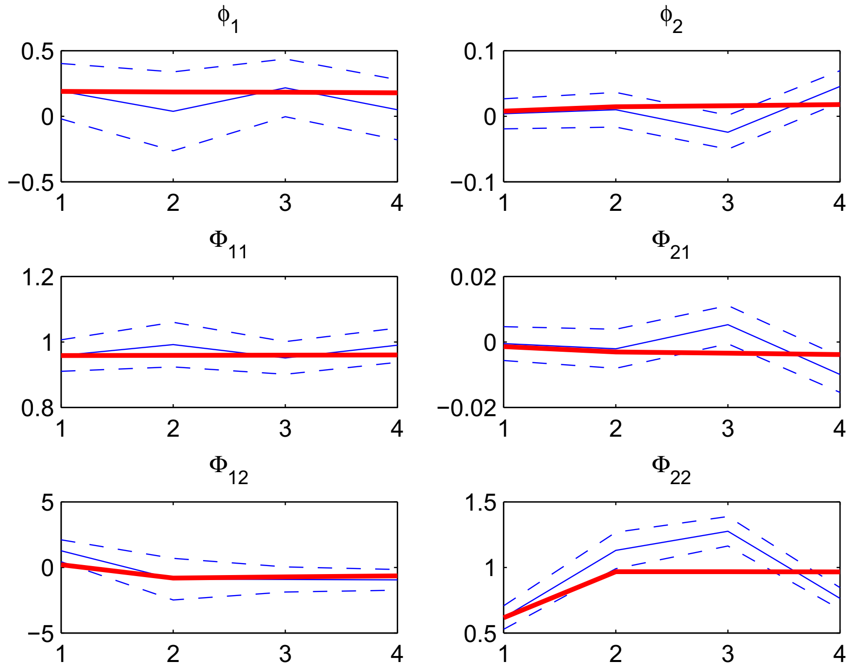

Fig. 7. π_Φ since 2005. Light blue line: Unrestricted OLS estimates of coefficients for regression in which y_{1t} is the dependent variable, plotted as a function of week of the month. Dashed blue lines: 95% confidence intervals for unrestricted OLS estimates. Bold red line: predicted values for coefficients derived from baseline model. All estimates based on data January 2005 to June 2011. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

where which observation week j is associated with a given maturity n can be read off of Table 1 and where we have defined $\rho_j^Q = \rho_j - \Lambda_j$ and $c_j^Q = c_j - \lambda_j$. Unfortunately, if all the parameters were allowed to vary with the week j in this way, the model would be unidentified. The reason is that even if one hypothesizes different values of ρ_j^Q for different j, a generalization of the algebra in (48) still implies that Γ_j should be the same for all j:

$$\Gamma_j = H_2 \begin{bmatrix} \beta_0' \\ \beta_4' \\ \beta_8' \end{bmatrix} \quad \text{for } j = 1, 2, 3, 4.$$

Since $\beta_4' = \beta_0' \rho_3^Q \rho_2^Q \rho_1^Q \rho_4^Q$ and $\beta_8' = \beta_4' \rho_3^Q \rho_2^Q \rho_1^Q \rho_4^Q$, the only information available from the regressions in which y_{2t} is the dependent variable (26) is about the product $\rho_3^Q \rho_2^Q \rho_1^Q \rho_4^Q$, which does not allow identification of the individual terms. In the next subsection we report estimates for a system in which although c_j, ρ_j, λ_j, and Λ_j all vary with j, the differences $c^Q = c_j - \lambda_j$ and $\rho^Q = \rho_j - \Lambda_j$ do not. For this system, the flexibility of the c_j and ρ_j parameters allows us to fit the unrestricted OLS values for ϕ_j and Φ_j perfectly. Details of the normalization and estimation for this less restrictive specification are reported in Appendices B and C.

5.2. Empirical results for the less restrictive seasonal model

Empirical estimates for the parameters of the above system for each of the two subsamples are reported in Tables 5 and 6. In the first subsample, the main differences are that the specification allows

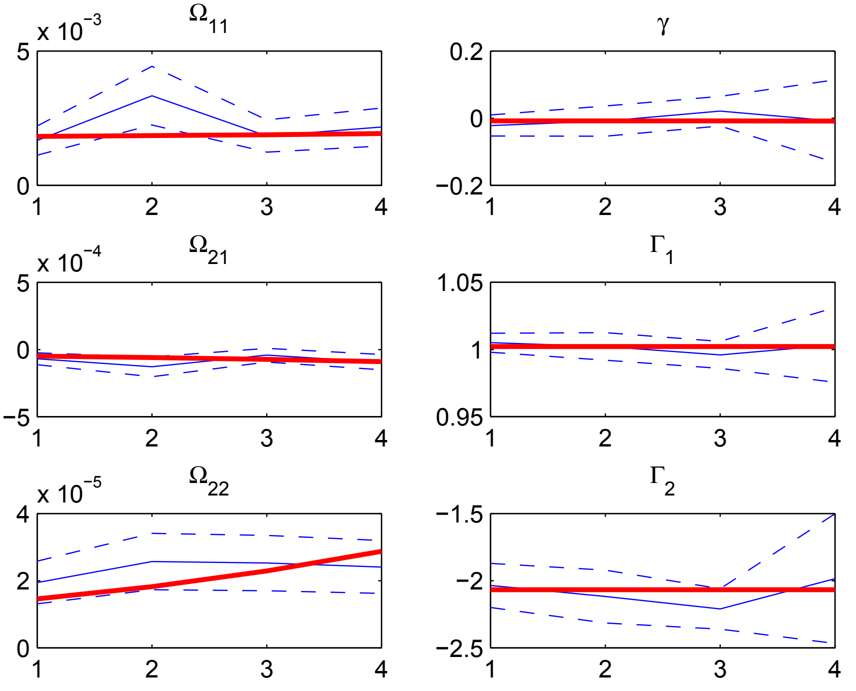

Fig. 8. π_Ω and π_Γ since 2005. First column: estimated elements of variance-covariance matrix for regression in which y_{1t} is the dependent variable, plotted as a function of week of the month. Second column: Estimated values of coefficients for regression in which y_{2t} is the dependent variable, plotted as a function of week of the month. In each panel, light blue lines are unrestricted OLS estimates, dashed blue lines are 95% confidence intervals for unrestricted OLS estimates, and bold red lines are predicted values for coefficients derived from baseline model. All estimates based on data January 2005 to June 2011. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

the spread to become harder to forecast as the near contract approaches expiry (that is, the (2,2) element of Σ_j increases in j) and the level and slope at the end of the month are less related to their values at expiry than is typical of the relation between y_{1t} and $y_{1,t-1}$ at other times (that is, diagonal elements of ρ_j are smaller for $j = 4$). Although implied values for λ and Λ are estimated with much less precision, the overall conclusion that individual elements are small and statistically insignificant applies across individual weeks as well.

For the second subsample (Table 6), the dependence of $\Lambda_{12,j}$ on week j is very dramatic, with an average value of -0.78 for $j = 1, 2$, or 3 but an estimated value of $+0.46$ for $j = 4$. A high spread signals lower returns to the long position during weeks 1–3, but this effect completely disappears, and may even take on the opposite sign, during expiry week 4. This may be related to the strong weekly pattern to index-fund strategies. For example, to replicate the crude oil holdings of the Goldman Sachs Commodity Index, an index fund would be selling the $k = 0$ contract and buying the $k = 1$ contract during week $j = 3$. It is interesting that we also find strong weekly patterns in the pricing of risk in data since 2005, though trying to interpret those changes in detail is beyond the scope of this paper.

Although our more general specification can fit the unrestricted OLS estimates $\widehat{\phi}_j$ and $\widehat{\Phi}_j$ perfectly, it still imposes testable overidentifying restrictions on other parameters, essentially using the 3 parameters in $\{\alpha_0, \xi_1, \xi_2\}$ to fit the 12 values for $\{\widehat{\gamma}_j, \widehat{\Gamma}_j\}_{j=1}^{4}$. The resulting $\chi^2(9)$ MCS test statistic for the first subsample is 13.86, which with a p-value of 0.13 is consistent with the null hypothesis that the

Table 5
Pre-2005 parameter estimates for seasonal model.

Estimated parameters					Implied parameters				
c^{base}	0.0102	0.0016			λ^{base}	0.0104	−0.0061*		
c_1	0.0042	0.0005			λ_1	0.0044	−0.0068		
	(0.0467)	(0.0056)				(0.0467)	(0.0055)		
c_2	−0.0359	0.0023			λ_2	−0.0356	−0.0050		
	(0.0422)	(0.0066)				(0.0422)	(0.0065)		
c_3	−0.0054	0.0053			λ_3	−0.0051	−0.0020		
	(0.0414)	(0.0055)				(0.0414)	(0.0054)		
c_4	0.0871*	−0.0003			λ_4	0.0874*	−0.0077		
	(0.0422)	(0.0073)				(0.0421)	(0.0079)		
ρ^{base}	0.9974*	0.1839*	−0.0006	0.9301*	Λ^{base}	−0.0023	−0.0591	0.0020*	−0.0050
ρ_1	0.9992*	0.2166	−0.0003	0.9369*	Λ_1	−0.0005	−0.0271	0.0022	−0.0061
	(0.0152)	(0.2386)	(0.0018)	(0.0285)		(0.0152)	(0.2386)	(0.0018)	(0.0281)
ρ_2	1.0137*	0.0371	−0.0010	0.9825*	Λ_2	0.0139	−0.2067	0.0015	0.0395
	(0.0138)	(0.2190)	(0.0022)	(0.0343)		(0.0138)	(0.2190)	(0.0021)	(0.0339)
ρ_3	1.0011*	0.5076*	−0.0018	0.9310*	Λ_3	0.0014	0.2639	0.0007	−0.0120
	(0.0135)	(0.2061)	(0.0018)	(0.0274)		(0.0135)	(0.2061)	(0.0018)	(0.0270)
ρ_4	0.9726*	−0.0718	0.0000	0.8690*	Λ_4	−0.0272*	−0.3155	0.0025	−0.0740
	(0.0138)	(0.2063)	(0.0024)	(0.0350)		(0.0138)	(0.2063)	(0.0026)	(0.0375)
Σ^{base}	0.0449*	−0.0038*	0.0047*						
Σ_1	0.0491*	−0.0042*	0.0039*						
	(0.0026)	(0.0004)	(0.0002)						
Σ_2	0.0447*	−0.0046*	0.0052*						
	(0.0024)	(0.0005)	(0.0003)						
Σ_3	0.0447*	−0.0029*	0.0051*						
	(0.0024)	(0.0004)	(0.0003)						
Σ_4	0.0446*	−0.0035*	0.0059*						
	(0.0024)	(0.0005)	(0.0003)						
ξ^{base}	0.9876*	0.9473*							
ξ	0.9854*	0.9574*							
	(0.0029)	(0.0070)							
α_0^{base}	0.0357*								
α_0	0.0331*								
	(0.0051)								
π_σ	0.0099*	0.0081*	0.0105*	0.0201*					

Left panel: MCS estimates for elements of θ (asymptotic standard errors in parentheses) for the unrestricted seasonal model, with estimates from baseline model also reported for comparison, with all estimates based on data from January 1990 through December 2004. Elements of matrices reported as first row, then second row. Right panel: values implied by the reported estimates of θ. * denotes statistically significant at the 5% level.

model has adequately captured all the week-to-week variations in parameters. The second subsample ($\chi^2(9) = 13.25$, $p = 0.15$) also passes this specification test.

6. Conclusions

In this paper, we studied the interaction between hedging demands from commercial producers or financial investors and risk aversion on the part of the arbitrageurs who are persuaded to be the hedgers' counterparties. We demonstrated that this interaction can produce an affine factor structure for the log prices of futures contracts in which expected returns depend on the arbitrageurs' net exposure to nondiversifiable risk. We developed new algorithms for estimation and diagnostic tools for testing this class of models appropriate for an unbalanced data set in which the duration of observed contracts changes with each observation.

Prior to 2005, we found that someone who consistently took the long side of nearby oil futures contracts received positive compensation on average, with relatively modest variation of this risk premium over time, consistent with the interpretation that the primary source of this premium was

Table 6
Post-2005 parameter estimates for seasonal model.

Estimated parameters					Implied parameters				
c^{base}	0.1802*	0.0164*			λ^{base}	0.1813*	0.0179*		
c_1	0.0361	0.0106			λ_1	0.0374	0.0119		
	(0.1528)	(0.0169)				(0.1528)	(0.0167)		
c_2	0.2252*	−0.0299*			λ_2	0.2264*	−0.0286		
	(0.1094)	(0.0147)				(0.1094)	(0.0144)		
c_3	0.0508	0.0443*			λ_3	0.0520	0.0456*		
	(0.1127)	(0.0120)				(0.1127)	(0.0119)		
c_4	0.1852	0.0102			λ_4	0.1865	0.0115		
	(0.1081)	(0.0175)				(0.1081)	(0.0184)		
ρ^{base}	0.9600*	−0.3487	−0.0035*	0.8629*	Λ^{base}	−0.0400*	−0.5892*	−0.0039*	−0.0311
ρ_1	0.9918*	−0.4445	−0.0021	1.0026*	Λ_1	−0.0082	−0.6847	−0.0025	0.1141
	(0.0346)	(0.5708)	(0.0038)	(0.0627)		(0.0346)	(0.5708)	(0.0038)	(0.0622)
ρ_2	0.9488*	−0.5659	0.0066*	1.1369*	Λ_2	−0.0512*	−0.8061*	0.0062	0.2484*
	(0.0247)	(0.3810)	(0.0033)	(0.0507)		(0.0247)	(0.3811)	(0.0033)	(0.0500)
ρ_3	0.9895*	−0.6038	−0.0097*	0.6759*	Λ_3	−0.0105	−0.8440*	−0.0100*	−0.2126*
	(0.0257)	(0.3454)	(0.0027)	(0.0367)		(0.0257)	(0.3454)	(0.0027)	(0.0364)
ρ_4	0.9592*	0.7020	−0.0021	0.8796*	Λ_4	−0.0409	0.4618	−0.0025	−0.0089
	(0.0245)	(0.4244)	(0.0040)	(0.0661)		(0.0245)	(0.4244)	(0.0042)	(0.0684)
Σ^{base}	0.0439*	−0.0021*	0.0049*						
Σ_1	0.0579*	−0.0031*	0.0055*						
	(0.0047)	(0.0007)	(0.0005)						
Σ_2	0.0423*	−0.0013*	0.0054*						
	(0.0035)	(0.0006)	(0.0004)						
Σ_3	0.0456*	−0.0020*	0.0044*						
	(0.0037)	(0.0005)	(0.0004)						
Σ_4	0.0417*	−0.0029*	0.0055*						
	(0.0034)	(0.0007)	(0.0005)						
ξ^{base}	1.0010*	0.8931*							
ξ	1.0008*	0.8877*							
	(0.0010)	(0.0059)							
α_0^{base}	−0.0086*								
α_0	−0.0077								
	(0.0106)								
π_σ	0.0059*	0.0086*	0.0087*	0.0223*					

Left panel: MCS estimates for elements of θ (asymptotic standard errors in parentheses) for the unrestricted seasonal model, with estimates from baseline model also reported for comparison, with all estimates based on data from January 2005 through June 2011. Elements of matrices reported as first row, then second row. Right panel: values implied by the reported estimates of θ. * denotes statistically significant at the 5% level.

hedging by commercial producers. However, we uncovered significant changes in the pricing of risk after the volume of trading in these contracts increased significantly in 2005. The expected compensation from a long position is lower on average in the recent data, often significantly negative when the futures curve slopes upward. We suggest that increased participation by financial investors in oil futures markets may have been a factor in changing the nature of risk premia in crude oil futures contracts.

Appendix A. Approximations to portfolio mean and variance

We first note that if $\log X \sim N(\mu, \sigma^2)$ then $E(X) = \exp(\mu + \sigma^2/2)$. Taking a first-order Taylor approximation around $\mu = \sigma^2 = 0$ we have $E(X) \approx 1 + \mu + \sigma^2/2$. Thus in particular since

$$\log(F_{n-1,t+1}/F_{nt}) \sim N\left(\mu_{n-1,t}, \sigma^2_{n-1}\right)$$

$$\mu_{n-1,t} = \alpha_{n-1} + \beta'_{n-1}(c + \rho x_t) - \alpha_n - \beta'_n x_t \qquad (38)$$

$$\sigma_{n-1}^2 = \beta_{n-1}'\Sigma\Sigma'\beta_{n-1} \tag{39}$$

we have the approximations

$$E_t\left[\left(F_{n-1,t+1} - F_{nt}\right)/F_{nt}\right] \approx \mu_{n-1,t} + \sigma_{n-1}^2/2$$

$$E_t\left[\exp\left(r_{j,t+1}\right)\right] \approx 1 + \xi_j + \psi_j'(c + \rho x_t) + \psi_j'\Sigma\Sigma'\psi_j/2.$$

Substituting these into (1) gives (5).

Likewise, if

$$\begin{bmatrix} \log X_i \\ \log X_j \end{bmatrix} \sim N\left(\begin{bmatrix} \mu_i \\ \mu_j \end{bmatrix}, \begin{bmatrix} \sigma_{ii} & \sigma_{ij} \\ \sigma_{ji} & \sigma_{jj} \end{bmatrix}\right),$$

then

$$\mathrm{Cov}\left(X_i, X_j\right) = \exp\left[\left(\mu_i + \mu_j\right) + (\sigma_{ii} + \sigma_{jj})/2\right]\left[\exp\left(\sigma_{ij}\right) - 1\right].$$

A first-order Taylor expansion around $\mu_i = \mu_j = \sigma_{ii} = \sigma_{jj} = \sigma_{ij} = 0$ gives

$$\mathrm{Cov}\left(X_i, X_j\right) \approx \sigma_{ij}.$$

To use this result we define

$$\underset{(L\times 1)}{y_{t+1}} = \left(r_{1,t+1}, \ldots, r_{J,t+1}, f_{0,t+1} - f_{1t}, f_{1,t+1} - f_{2t}, \ldots, f_{N-1,t+1} - f_{Nt}\right)'$$

for $L = J + N$. Notice that conditional on information at date t, $y_{t+1} \sim N(\mu_t, H'\Sigma\Sigma' H)$ for[13]

$$H = \begin{bmatrix} \psi_1 & \cdots & \psi_J & \beta_0 & \cdots & \beta_{N-1} \end{bmatrix}.$$

Notice further that (1) can be written $W_{t+1} = k_t + \sum_{\ell=1}^{L} h_{\ell t} \exp(y_{\ell,t+1})$ for $h_{\ell t} = q_{\ell t}$ for $\ell = 1, \ldots, J$ and $h_{\ell t} = z_{\ell-J,t}$ for $\ell = J+1, \ldots, L$. Thus for $h_t = (h_{1t}, \ldots, h_{Lt})'$,

$$\begin{aligned} &\mathrm{Var}_t(W_{t+1}) \approx h_t' H'\Sigma\Sigma' H h_t \\ &= \left(\sum_{j=1}^{J} q_{jt}\psi_j' + \sum_{n=1}^{N} z_{nt}\beta_{n-1}'\right)\Sigma\Sigma'\left(\sum_{j=1}^{J} q_{jt}\psi_j + \sum_{\ell=1}^{N} z_{\ell t}\beta_{\ell-1}\right). \end{aligned}$$

Appendix B. Normalization

Baseline model. Let $\rho^Q = \rho - \Lambda$, and notice from (10) that

$$\beta_n = \left(\rho^{Q\prime}\right)^n \beta_0. \tag{40}$$

We will parameterize this in terms of $\xi = (\xi_1, \xi_2)'$ where ξ_i denotes an eigenvalue of $\rho^{Q\prime}$. We could calculate the following matrix as a function of those eigenvalues:

$$\underset{(2\times 3)}{K(\xi)} = \begin{bmatrix} \xi_1^0 & \xi_1^4 & \xi_1^8 \\ \xi_2^0 & \xi_2^4 & \xi_2^8 \end{bmatrix}. \tag{41}$$

The claim is that if we specify

---

[13] Here $\mu_t = (\mu_{1t}, \ldots, \mu_{Lt})'$ for $\mu_{\ell t} = \xi_\ell + \psi_\ell'(c + \rho x_t)$ for $\ell = 1, \ldots, J$ and $\mu_{\ell t} = \alpha_{\ell-J-1} + \beta_{\ell-J-1}'(c + \rho x_t) - \alpha_{\ell-J} - \beta_{\ell-J}' x_t$ for $\ell = J+1, \ldots, L$.

$$\underset{(2\times2)}{\rho^{Q\prime}} = \begin{bmatrix} \underset{(2\times3)}{K(\xi)} & \underset{(3\times2)}{H_1'} \end{bmatrix}^{-1} \begin{bmatrix} \xi_1 & 0 \\ 0 & \xi_2 \end{bmatrix} \begin{bmatrix} \underset{(2\times3)}{K(\xi)} & \underset{(3\times2)}{H_1'} \end{bmatrix} \tag{42}$$

$$\underset{(2\times1)}{\beta_0} = \begin{bmatrix} \underset{(2\times3)}{K(\xi)} & \underset{(3\times2)}{H_1'} \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \tag{43}$$

then (34), the desired condition for β_n, would be satisfied. To prove this, observe from (40) that

$$\begin{aligned} \beta_n &= \left[K(\xi)H_1'\right]^{-1} \begin{bmatrix} \xi_1^n & 0 \\ 0 & \xi_2^n \end{bmatrix} \left[K(\xi)H_1'\right] \left[K(\xi)H_1'\right]^{-1} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \\ &= \left[K(\xi)H_1'\right]^{-1} \begin{bmatrix} \xi_1^n \\ \xi_2^n \end{bmatrix} \end{aligned} \tag{44}$$

so that

$$H_1 \begin{bmatrix} \beta_0' \\ \beta_4' \\ \beta_8' \end{bmatrix} = H_1 \begin{bmatrix} \xi_1^0 & \xi_2^0 \\ \xi_1^4 & \xi_2^4 \\ \xi_1^8 & \xi_2^8 \end{bmatrix} \left[H_1 K(\xi)'\right]^{-1}. \tag{45}$$

Substituting (41) into (45) produces (34), as claimed. Thus if we know ξ, we can use (41) and (44) to calculate the value of β_n for any n as well as ρ^Q from (42).

To achieve the separate condition (35) on α_n, notice from (11) that

$$\alpha_n = \alpha_0 + \left(\beta_{n-1}' + \beta_{n-2}' + \cdots + \beta_0'\right)c^Q + (1/2)\left(\beta_{n-1}'\Sigma\Sigma'\beta_{n-1} + \beta_{n-2}'\Sigma\Sigma'\beta_{n-2} + \cdots + \beta_0'\Sigma\Sigma'\beta_0\right). \tag{46}$$

Define

$$\begin{aligned} \underset{(1\times2)}{\zeta_n(\xi)} &= \beta_{n-1}' + \beta_{n-2}' + \cdots + \beta_0' \\ \underset{(1\times1)}{\psi_n(\xi, \alpha_0, \Sigma)} &= \alpha_0 + (1/2)\left(\beta_{n-1}'\Sigma\Sigma'\beta_{n-1} + \beta_{n-2}'\Sigma\Sigma'\beta_{n-2} + \cdots + \beta_0'\Sigma\Sigma'\beta_0\right) \end{aligned}$$

so that (46) can be written

$$\alpha_n = \zeta_n(\xi)c^Q + \psi_n(\xi, \alpha_0, \Sigma)$$

where for $n = 0$ we have $\zeta_0(\xi) = 0$ and $\psi_0(\xi, \alpha_0, \Sigma) = \alpha_0$. We claim that if we choose

$$c^Q = -\left(H_1 \begin{bmatrix} \zeta_0(\xi) \\ \zeta_4(\xi) \\ \zeta_8(\xi) \end{bmatrix}\right)^{-1} \left(H_1 \begin{bmatrix} \psi_0(\xi, \alpha_0, \Sigma) \\ \psi_4(\xi, \alpha_0, \Sigma) \\ \psi_8(\xi, \alpha_0, \Sigma) \end{bmatrix}\right), \tag{47}$$

then (35) would be satisfied. This is demonstrated as follows:

$$\begin{aligned} H_1 \begin{bmatrix} \alpha_0 \\ \alpha_4 \\ \alpha_8 \end{bmatrix} &= H_1 \begin{bmatrix} \zeta_0(\xi)c^Q + \psi_0(\xi, \alpha_0, \Sigma) \\ \zeta_4(\xi)c^Q + \psi_4(\xi, \alpha_0, \Sigma) \\ \zeta_8(\xi)c^Q + \psi_8(\xi, \alpha_0, \Sigma) \end{bmatrix} \\ &= -H_1 \begin{bmatrix} \psi_0(\xi, \alpha_0, \Sigma) \\ \psi_4(\xi, \alpha_0, \Sigma) \\ \psi_8(\xi, \alpha_0, \Sigma) \end{bmatrix} + H_1 \begin{bmatrix} \psi_0(\xi, \alpha_0, \Sigma) \\ \psi_4(\xi, \alpha_0, \Sigma) \\ \psi_8(\xi, \alpha_0, \Sigma) \end{bmatrix} \\ &= 0. \end{aligned}$$

Thus if we know ξ, α_0, and Σ, we can use (46) and (47) to calculate the value of α_n for any n. Also ξ, α_0, and Σ allow calculation of c^Q from (47).

Seasonal model. Just as in the baseline model, we let (ξ_1, ξ_2) denote the ordered eigenvalues of ρ^Q and write

$$\beta_n' = \begin{bmatrix} \xi_1^n & \xi_2^n \end{bmatrix} \left[H_1 K(\xi)' \right]^{-1}$$

which achieves the normalization (34). Likewise for α_n we again use (47) where now

$$\begin{aligned} \zeta_n(\xi) &= \beta_{n-1}' + \beta_{n-2}' + \cdots + \beta_0' \\ \psi_n(\xi, \alpha_0, \Sigma) &= \alpha_0 + (1/2)\left(\beta_{n-1}' \Sigma_{j(n)} \Sigma_{j(n)}' \beta_{n-1} + \beta_{n-2}' \Sigma_{j(n-1)} \Sigma_{j(n-1)}' \beta_{n-2} + \cdots + \beta_0' \Sigma_{j(1)} \Sigma_{j(1)}' \beta_0 \right). \end{aligned}$$

Appendix C. Mapping from structural to reduced-form parameters

Baseline model. Expressions involving B_{1j} in (23) can be simplified by noting from (40) and (34) that

$$B_{1j} = H_1 \begin{bmatrix} \beta_{4-j}' \\ \beta_{8-j}' \\ \beta_{12-j}' \end{bmatrix} = H_1 \begin{bmatrix} \beta_0' \\ \beta_4' \\ \beta_8' \end{bmatrix} \left(\rho^Q \right)^{4-j} = \left(\rho^Q \right)^{4-j}.$$

Thus for example (27) simplifies to

$$\begin{aligned} \Gamma_j &= H_2 \begin{bmatrix} \beta_{4-j}' \\ \beta_{8-j}' \\ \beta_{12-j}' \end{bmatrix} \left(H_1 \begin{bmatrix} \beta_{4-j}' \\ \beta_{8-j}' \\ \beta_{12-j}' \end{bmatrix} \right)^{-1} \\ &= H_2 \begin{bmatrix} \beta_0' \\ \beta_4' \\ \beta_8' \end{bmatrix} \left(\rho^Q \right)^{4-j} \left[\left(\rho^Q \right)^{4-j} \right]^{-1} \\ &= H_2 \begin{bmatrix} \beta_0' \\ \beta_4' \\ \beta_8' \end{bmatrix} \quad \text{for } j = 1, 2, 3, 4. \end{aligned} \tag{48}$$

The population magnitudes corresponding to the other reduced-form OLS coefficients are as follows:

$$\Omega_j = \left(\rho^Q \right)^{4-j} \Sigma \Sigma' \left(\rho^{Q\prime} \right)^{4-j} \quad \text{for } j = 1, 2, 3, 4$$

$$\Phi_1 = \left(\rho^Q \right)^3 \rho$$

$$\Phi_j = \left(\rho^Q \right)^{4-j} \rho \left[\left(\rho^Q \right)^{-1} \right]^{4-j+1} \quad \text{for } j = 2, 3, 4$$

$$\phi_1 = H_1 \begin{bmatrix} \alpha_3 \\ \alpha_7 \\ \alpha_{11} \end{bmatrix} + H_1 \begin{bmatrix} \beta_3' \\ \beta_7' \\ \beta_{11}' \end{bmatrix} c - \Phi_1 H_1 \begin{bmatrix} \alpha_0 \\ \alpha_4 \\ \alpha_8 \end{bmatrix}$$

$$\phi_j = H_1 \begin{bmatrix} \alpha_{4-j} \\ \alpha_{8-j} \\ \alpha_{12-j} \end{bmatrix} + H_1 \begin{bmatrix} \beta'_{4-j} \\ \beta'_{8-j} \\ \beta'_{12-j} \end{bmatrix} c - \Phi_j H_1 \begin{bmatrix} \alpha_{5-j} \\ \alpha_{9-j} \\ \alpha_{13-j} \end{bmatrix} \quad \text{for } j = 2,3,4$$

$$\gamma_j = H_2 \begin{bmatrix} \alpha_{4-j} \\ \alpha_{8-j} \\ \alpha_{12-j} \end{bmatrix} - \Gamma_j H_1 \begin{bmatrix} \alpha_{4-j} \\ \alpha_{8-j} \\ \alpha_{12-j} \end{bmatrix} \quad \text{for } j = 1,2,3,4.$$

Given the scalars $\{\xi_1, \xi_2\}$ (corresponding to the eigenvalues of $\rho^Q = \rho - \Lambda$), we can calculate β_n from (44) and (41). These β_n give us predicted values for $\{\Gamma_j\}_{j=1}^4$, and the β_n along with Σ give predicted values for $\{\Omega_j\}_{j=1}^4$. Note $\{\xi_1, \xi_2\}$ also gives us ρ^Q, and this plus ρ gives predicted values for $\{\Phi_j\}_{j=1}^4$. From β_n, Σ, and α_0 we can calculate c^Q from (47) and α_n from (46). Using these along with c we then obtain the predicted values for $\{\phi_j\}_{j=1}^4$ and $\{\gamma_j\}_{j=1}^4$.

The information matrix for the OLS estimates $\widehat{\pi} = (\widehat{\pi}'_\Phi, \widehat{\pi}'_\Omega, \widehat{\pi}'_\Gamma \widehat{\pi}_\sigma)'$ is given by

$$\widehat{R}_\pi = \begin{bmatrix} \widehat{R}_\Phi & 0 & 0 & 0 \\ 0 & \widehat{R}_\Omega & 0 & 0 \\ 0 & 0 & \widehat{R}_\Gamma & 0 \\ 0 & 0 & 0 & \widehat{R}_\sigma \end{bmatrix} \tag{49}$$

$$\widehat{R}_\Phi = \begin{bmatrix} \widehat{R}_{\Phi 1} & 0 & 0 & 0 \\ 0 & \widehat{R}_{\Phi 2} & 0 & 0 \\ 0 & 0 & \widehat{R}_{\Phi 3} & 0 \\ 0 & 0 & 0 & \widehat{R}_{\Phi 4} \end{bmatrix}$$

$$\widehat{R}_{\Phi j} = \widehat{\Omega}_j^{-1} \otimes \mathcal{T}^{-1} \sum_{\tau=1}^{\mathcal{T}} x^{\dagger}_{1,j,\tau} x^{\dagger\prime}_{1,j,\tau}$$

$$\widehat{R}_\Omega = \begin{bmatrix} \widehat{R}_{\Omega 1} & 0 & 0 & 0 \\ 0 & \widehat{R}_{\Omega 2} & 0 & 0 \\ 0 & 0 & \widehat{R}_{\Omega 3} & 0 \\ 0 & 0 & 0 & \widehat{R}_{\Omega 4} \end{bmatrix}$$

$$\widehat{R}_{\Omega j} = (1/2) D'_2 \left(\widehat{\Omega}_j^{-1} \otimes \widehat{\Omega}_j^{-1} \right) D_2$$

$$D_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\widehat{R}_\Gamma = \begin{bmatrix} \widehat{R}_{\Gamma 1} & 0 & 0 & 0 \\ 0 & \widehat{R}_{\Gamma 2} & 0 & 0 \\ 0 & 0 & \widehat{R}_{\Gamma 3} & 0 \\ 0 & 0 & 0 & \widehat{R}_{\Gamma 4} \end{bmatrix}$$

$$\widehat{R}_{\Gamma j} = \widehat{\sigma}_{ej}^{-2}\left(\mathcal{T}^{-1}\sum_{\tau=1}^{\mathcal{T}} x_{2,j,\tau}^{\dagger} x_{2,j,\tau}^{\dagger\prime}\right)$$

$$\widehat{R}_{\sigma} = \begin{bmatrix} (1/2)\widehat{\sigma}_{e1}^{-4} & 0 & 0 & 0 \\ 0 & (1/2)\widehat{\sigma}_{e2}^{-4} & 0 & 0 \\ 0 & 0 & (1/2)\widehat{\sigma}_{e3}^{-4} & 0 \\ 0 & 0 & 0 & (1/2)\widehat{\sigma}_{e4}^{-4} \end{bmatrix}$$

where D_2 is the duplication matrix satisfying $D_2\,\mathrm{vech}(\Omega) = \mathrm{vec}(\Omega)$ for 2×2 symmetric matrix Ω. Note that for all models considered the MCS estimate of σ_{ej} is always equal to the unconstrained MLE $\widehat{\sigma}_{ej}$ and so contributes 0 to the weighted objective function.

Unrestricted seasonal model. For this model the specifications for γ_j and Γ_j are the same as in the baseline model, while the expressions for the other parameters become

$$\Omega_1 = \left(\rho^{Q}\right)^{3}\Sigma_4\Sigma_4'\left(\rho^{Q'}\right)^{3}$$

$$\Omega_j = \left(\rho^{Q}\right)^{4-j}\Sigma_{j-1}\Sigma_{j-1}'\left(\rho^{Q'}\right)^{4-j} \quad \text{for } j = 2,3,4$$

$$\Phi_1 = \left(\rho^{Q}\right)^{3}\rho_4$$

$$\Phi_j = \left(\rho^{Q}\right)^{4-j}\rho_{j-1}\left[\left(\rho^{Q}\right)^{-1}\right]^{4-j+1} \quad \text{for } j = 2,3,4 \tag{50}$$

$$\phi_1 = H_1\begin{bmatrix}\alpha_3\\ \alpha_7\\ \alpha_{11}\end{bmatrix} + H_1\begin{bmatrix}\beta_3'\\ \beta_7'\\ \beta_{11}'\end{bmatrix}c_4 - \Phi_1 H_1\begin{bmatrix}\alpha_0\\ \alpha_4\\ \alpha_8\end{bmatrix}$$

$$\phi_j = H_1\begin{bmatrix}\alpha_{4-j}\\ \alpha_{8-j}\\ \alpha_{12-j}\end{bmatrix} + H_1\begin{bmatrix}\beta_{4-j}'\\ \beta_{8-j}'\\ \beta_{12-j}'\end{bmatrix}c_{j-1} - \Phi_j H_1\begin{bmatrix}\alpha_{5-j}\\ \alpha_{9-j}\\ \alpha_{13-j}\end{bmatrix} \quad \text{for } j = 2,3,4.$$

Minimum-chi-square estimation in this case is achieved by first choosing $\{\xi, \alpha_0, \Sigma_1, \Sigma_2, \Sigma_3, \Sigma_4\}$ so as to minimize the distance from the OLS estimates $\{\widehat{\gamma}_j, \widehat{\Gamma}_j, \widehat{\Omega}_j\}_{j=1}^{4}$. From ξ we can then calculate ρ^{Q}, with which we can obtain ρ_j analytically from (50) in order to fit these OLS coefficients perfectly. The values c_{j-1} are likewise obtained analytically from $\widehat{\phi}_j$.

References

Acharya, Viral V., Lochstoer, Lars A., Ramadorai, Tarun, 2013. Limits to arbitrage and hedging: evidence from commodity markets (forthcoming). Journal of Financial Economics.

Alquist, Ron, Gervais, Olivier, 2011. The Role of Financial Speculation in Driving the Price of Crude Oil. Working paper. Bank of Canada.
Alquist, Ron, Kilian, Lutz, 2007. What Do We Learn from the Price of Crude Oil Futures? CEPR. Working paper No. 6548.
Alquist, Ron, Kilian, Lutz, 2010. What do we learn from the price of crude oil futures? Journal of Applied Econometrics 25, 539–573.
Ang, Andrew, Piazzesi, Monika, 2003. A no-arbitrage vector autoregression of term structure dynamics with macroeconomic and latent variables. Journal of Monetary Economics 50, 745–787.
Bauer, Michael D., Rudebusch, Glenn D., Wu, Jing Cynthia, 2012. Correcting estimation bias in dynamic term structure models. Journal of Business & Economic Statistics 30, 454–467.
Bessembinder, Hendrik, 1992. Systemic risk, hedging pressure, and risk premiums in futures markets. Review of Financial Studies 5, 637–667.
Buyuksahin, Bahattin, Robe, Michel A., 2011. Does 'Paper Oil' Matter?. Working paper American University.
Carter, Colin A., Rausser, Gordon C., Schmitz, Andrew, 1983. Efficient asset portfolios and the theory of normal backwardation. Journal of Political Economy 91, 319–331.
Casassus, Jaime, Collin-Dufresne, Pierre, 2006. Stochastic convenience yield implied from commodity futures and interest rates. Journal of Finance 60, 2283–2331.
Chang, Eric C., 1985. Returns to speculators and the theory of normal backwardation. Journal of Finance 40, 193–208.
Chen, Ren-Raw, Scott, Louis, 1993. Maximum likelihood estimation for a multifactor equilibrium model of the term structure of interest rates. The Journal of Fixed Income 3, 14–31.
Cochrane, John H., Piazzesi, Monika, 2009. Decomposing the yield curve. In: AFA 2010 Atlanta Meetings Paper.
Cortazar, Gonzalo, Naranjo, Lorenzo, 2006. An N-factor Gaussian model of oil futures prices. Journal of Futures Markets 26, 243–268.
Creal, Drew D., Wu, Jing Cynthia, 2013. Estimation of Non-Gaussian Affine Term Structure Models. Chicago Booth Research Paper No. 13–72.
Dai, Qiang, Singleton, Kenneth J., 2002. Expectation puzzles, time-varying risk premia, and affine models of the term structure. Journal of Financial Economics 63, 415–441.
De Roon, Frans, Nijman, Theo, Veld, Chris, 2000. Hedging pressure effects in futures markets. Journal of Finance 55, 1437–1456.
Duffee, Gregory R., 2002. Term premia and interest rate forecasts in affine models. The Journal of Finance 57, 405–443.
Duffie, Darrell, 1992. Dynamic Asset Pricing Theory. Princeton University Press, Princeton, NJ.
Duffie, Darrell, Kan, Rui, 1996. A yield-factor model of interest rates. Mathematical Finance 6, 379–406.
Etula, Erkko, 2013. Broker-dealer risk appetite and commodity returns. Journal of Financial Econometrics 11, 486–521.
Fattouh, Bassam, Kilian, Lutz, Mahadeva, Lavan, 2013. The role of speculation in oil markets: what have we learned so far? Energy Journal 34, 7–33.
Hamilton, James D., 1994. Time Series Analysis. Princeton University Press, Princeton, New Jersey.
Hamilton, James D., 2009. Causes and Consequences of the Oil Shock of 2007–08. Brookings Papers on Economic Activity Spring, pp. 215–259.
Hamilton, James D., Wu, Jing Cynthia, 2012a. The effectiveness of alternative monetary policy tools in a zero lower bound environment. Journal of Money, Credit & Banking 44 (s1), 3–46.
Hamilton, James D., Wu, Jing Cynthia, 2012b. Identification and estimation of Gaussian affine term structure models. Journal of Econometrics 168, 315–331.
Hamilton, James D., Wu, Jing Cynthia, 2012c. Effects of Index-fund Investing on Commodity Futures Prices. Working paper. UCSD.
Hamilton, James D., Wu, Jing Cynthia, 2013. Testable implications of affine term structure models (forthcoming). Journal of Econometrics.
Irwin, Scott H., Sanders, Dwight R., 2012. Testing the masters hypothesis in commodity futures markets. Energy Economics 34, 256–269.
Joslin, Scott, Singleton, Kenneth J., Zhu, Haoxiang, 2011. A new perspective on Gaussian dynamic term structure models. Review of Financial Studies 24, 926–970.
Keynes, John M., 1930. A Treatise on Money, vol. 2. Macmillan, London.
Kilian, Lutz, Murphy, Daniel P., 2013. The role of inventories and speculative trading in the global market for crude oil (forthcoming). Journal of Applied Econometrics.
Mou, Yiqun, 2010. Limits to Arbitrage and Commodity Index Investment: Front-running the Goldman Roll. Working paper. Columbia University.
Schwartz, Eduardo S., 1997. The stochastic behavior of commodity prices: implications for valuation and hedging. Journal of Finance 52, 923–973.
Schwartz, Eduardo S., Smith, James E., 2000. Short-term variations and long-term dynamics in commodity prices. Management Science 46, 893–911.
Singleton, Kenneth J., 2011. Investor Flows and the 2008 Boom/Bust in Oil Prices. Working paper. Stanford University.
Tang, Ke, Xiong, Wei, 2012. Index investment and the financialization of commodities. Financial Analysts Journal 68, 54–74.
Vasicek, Oldrich, 1977. An equilibrium characterization of the term structure. Journal of Financial Economics 5, 177–188.
Vayanos, Dimitri, Vila, Jean-Luc, 2009. A Preferred-habitat Model of the Term Structure of Interest Rates. Working paper. London School of Economics.