![BANK OF ENGLAND]

Staff Working Paper No. 591
Risk premia and seasonality in commodity futures
Constantino Hevia, Ivan Petrella and Martin Sola

April 2016

Staff Working Paper No. 591
Risk premia and seasonality in commodity futures
Constantino Hevia,[1] Ivan Petrella[2] and Martin Sola[3]

Abstract

We develop and estimate a multifactor affine model of commodity futures that allows for stochastic variations in seasonality. We show conditions under which the yield curve and the cost-of-carry curve adopt augmented Nelson and Siegel functional forms. This restricted version of the model is parsimonious, does not suffer from identification problems, and matches well the yield curve and futures curve over time. We estimate the model using heating oil futures prices over the period 1984–2012. We find strong evidence of stochastic seasonality in the data. We analyse risk premia in futures markets and discuss two traditional theories of commodity futures: the theory of storage and the theory of normal backwardation. The data strongly supports the theory of storage.

Key words: Commodity futures, Nelson and Siegel, seasonality, risk premium, theory of storage.

JEL classification: G13, E43, Q02, Q40.

(1) Universidad Torcuato di Tella. Email: chevia@utdt.edu
(2) Bank of England, Birkbeck University of London and CEPR. Email: ivan.petrella@bankofengland.co.uk
(3) Universidad Torcuato di Tella.and Birkbeck University of London. Email: msola@utdt.edu

1 Introduction

Seasonal fluctuations in commodity prices are driven by systematic intra-year changes in supply and demand. Energy commodities, such as heating oil and natural gas, display a demand peak during the winter season in the northern hemisphere. Agricultural commodities, such as corn and soybeans, display periodic changes in supply: prices tend to be lower during the harvest season and higher during the planting season. Often, seasonal fluctuations are not perfectly predictable. For example, a mild winter lowers the demand for energy consumption, dampening the seasonal content of heating oil and natural gas prices. From the point of view of hedgers and speculators, stochastic seasonal fluctuations imply a source of risk that manifests itself in futures prices and risk premia. In contrast, perfectly predictable seasonal fluctuations would be reflected in prices but not on the risk faced by market participants.[1]

We develop and estimate a multifactor affine model of commodity futures that allows for stochastic variations in seasonality. Ours is a generalization of the conventional three-factor affine model of commodity futures.[2] In the usual model, futures prices are driven by three factors: one factor associated with the spot commodity price, a second factor describing the short interest rate, and a third factor representing an instantaneous convenience yield on inventory or cost-of-carry. Yet, it is known since Litterman and Scheinkman (1991) that one needs three factors to properly describe the yield curve for government bonds—often labeled level, slope, and curvature factors. Using this insight we propose a flexible yet parsimonious nine-factor model. Three of them determine the yield curve on government bonds; one factor is associated with the non-seasonal component of the spot commodity price, three factors determine the cost-of-carry (or convenience yield) curve, and two factors are associated with seasonal shocks. Stochastic seasonal fluctuations are driven by two unobserved factors, as in Hannan (1964). This specification allows us to price seasonal shocks by attaching market prices of risk to the seasonal factors, and to match futures prices and bond yields with great accuracy. We estimate the model using data on heating oil prices, which contain a clear seasonal pattern. The model, however, can be applied to analyze any term structure of commodity futures prices.

To solve the well known identification problems of affine models (see e.g. Hamilton and

---

[1]Existing seasonal models of commodity futures include deterministic seasonal fluctuations in prices (e.g. Sorensen, 2002) or in the convenience yield (e.g. Borovkova and Geman, 2006). These models cannot be used to measure the relevance and risks associated with variations in seasonal fluctuations.

[2]Gibson and Schwartz (1990) and Litzenberger and Rabinowitz (1995) are early contributions that include a stochastic instantaneous convenience yield factor. Schwartz (1997) extended the model to include stochastic interest rates. See also Casassus and Collin-Dufresne (2005) for a different interpretation of the three-factor model. Hilliard and Reis (1998), Miltersen and Schwartz (1998) and Chin and Liu (2015) use the affine framework to price derivatives on commodity futures, and Hamilton and Wu (2013) study risk premia in oil futures markets.

Wu, 2012) we propose a Nelson and Siegel representation of the yield and cost-of-carry curves, and find conditions under which this is arbitrage-free.[3] The Nelson and Siegel representation of the model imposes strong restrictions on the evolution of the state variables under the risk neutral measure. Those restrictions allow us to easily identify the market prices of the risk factors: identifying the 81 parameters of the 9×9 matrix that maps factors into market prices of risk amounts to estimating only 3 parameters. We view this substantial reduction in the number of parameters as a key advantage of the Nelson and Siegel representation over other unrestricted versions of the model.

We estimate the model using monthly data on heating oil futures prices with maturities up to 24 months and U.S. zero coupon bond prices with maturities up to 5 years for the period 1984-2012. The model is able to match the cross-section of futures prices over time, including their seasonal pattern. We find strong evidence of stochastic seasonality: the amplitude of the seasonal fluctuations decreased considerably over time, particularly at the end of the sample. A model with deterministic seasonality is unable to capture this pattern. Also, consistent with the theory of storage, the moderation of the seasonal component coincides with a similar moderation of the seasonal component of heating oil inventories. The model is misspecified if we do not allow for time variation in the seasonal pattern. Among other problems, a model with deterministic seasonality erroneously attributes the time-variation in the seasonal component to the spot and cost-of-carry factors. And those spurious fluctuations in the commodity factors translate into spurious fluctuations in risk premia.

Since entering into a futures contract costs zero, any expected return is a risk premium. Expected returns of holding a futures contract fluctuate widely over time, and much of those fluctuations come from high frequency variations in the spot and cost-of-carry factors. In addition, the risk premium associated with holding a futures contract for several periods is correlated with medium frequency movements in the spot factor. Although non-negligible, the contribution of seasonal shocks to risk premia is relatively small. Seasonal shocks account for variations in expected returns of about 0.5 percentage points on an annualized basis. Therefore, correctly specifying seasonality as stochastic is important not so much because risk premia depend a lot on seasonal fluctuations, but to avoid erroneously assigning those fluctuations to other factors.

A common claim in the literature is that interest rate shocks have a minor impact on the time variation of risk premia. Schwartz (1997) assumes a constant interest rate because interest rate fluctuations are orders of magnitudes lower than those in futures returns. In their three factor model, Casassus and Collin-Dufresne (2005) argue that the market price of interest rate shocks is barely significant. We find, however, that yield curve factors do

---

[3]We extend the results of Christensen et al. (2011) to the pricing of commodity futures.

have a significant impact on risk premia, mostly at medium and lower frequencies. Interest rates declined from about 12 percentage points to roughly zero during our sample period. The contribution of interest rate factors to expected holding returns went from about -10 percentage points to zero over the same time frame. Fluctuations in the slope of the yield curve also affect expected holding returns. When the slope of the yield curve is positive, longer term contracts are relatively more expensive than shorter term contracts while the reverse holds when the yield curve is inverted. Changes in the slope of the yield curve over time thus affect futures prices and risk premia. Overall, we find that several measures of risk premia begun to drop by 2007. This drop is associated with a decline in the risk premia associated with the commodity factors and a decline in the (negative) risk premia associated with the yield curve factors. The contribution of the seasonal shocks to risk premia also declines, but this effect is much smaller than that of the other factors.

We also use the estimated model to evaluate the theory of storage and the theory of normal backwardation (see e.g. Gorton et al., 2012) and find evidence in favor of the former. Tests of those theories usually rely on short term contracts due to the lack of long time series of long dated contracts. For example, futures on heating oil that mature in 2 years begun trading in 2007. Instead, the structure imposed by the model, estimated using an unbalanced panel of futures prices, allows us to construct futures prices and risk premia for contracts of any maturity between 1 and 24 months ahead over the entire sample. We find a negative relation between the non-seasonal component of inventories and the net convenience yield for all futures contracts. The R-squares of the regressions increase with the maturity of the contracts, from about 0.25 for 1-month contracts to over 0.50 for 24-months contracts. We also find a nonlinear relation between inventories and the net convenience yield, which is stronger at shorter maturities. This result is in line with the hypothesis of decreasing returns from holding the commodity in storage (Deaton and Laroque, 1992). Also consistent with the theory of storage, we observe a negative and significant relation between expected holding returns and inventories. This relation is stronger when the holding period is 12 months or more. If we subtract the contribution of the interest rate factors from the overall risk premium, the relation between inventories and expected returns becomes tighter: the R-squares of the regressions of long term holding returns on inventories are over 0.50. Support for the theory of normal backwardation is instead weak: we only find a positive effect of hedging pressure on expected returns for holding 1 or 3 months a 24-month futures contract with R-squares of only about 2 percent. For other maturities and holding periods, the coefficients of the regression of risk premia on hedging pressure are zero or negative, which is inconsistent with the theory.

2 Affine model of commodity futures

In this section we describe an affine model of commodity futures with stochastic seasonality. The risk factors are represented by a vector of state variables $X_t \in \mathbb{R}^n$, where the time period t is measured in months. The state vector includes factors capturing the stochastic variation in seasonality, which we specify below. The state variables evolve as

$$X_{t+1} = \mu + \Phi X_t + \Gamma \eta_{t+1}, \tag{1}$$

where $\eta_{t+1}|X_t \sim N(0, I)$ and Γ is lower triangular.

Nominal cash-flows are priced using the stochastic discount factor

$$
\begin{aligned}
M_{t,t+1} &= e^{-\left(r_t + \frac{1}{2}\Lambda_t' \Lambda_t + \Lambda_t' \eta_{t+1}\right)}, \\
\Lambda_t &= \lambda_0 + \lambda_1 X_t,
\end{aligned}
\tag{2}
$$

where r_t is the one period spot interest rate and $\Lambda_t \in \mathbb{R}^n$ is the compensation for risk to shocks to the state vector η_{t+1}.

The spot interest rate r_t is an affine function of the state variables

$$r_t = \rho_0 + \rho_1' X_t. \tag{3}$$

where ρ_0 is a scalar and ρ_1 is an n-dimensional vector. Since there is no evidence of seasonality in interest rates, we set to zero the loadings of ρ_1 associated with the seasonal factors.

2.1 Pricing government bonds

Let $P_t^{(\tau)}$ and $y_t^{(\tau)}$ be the price and yield of a τ period zero-coupon bond. The absence of arbitrage implies that bond prices satisfy the pricing condition

$$P_t^{(\tau)} = E_t \left[M_{t,t+1} P_{t+1}^{(\tau-1)} \right].$$

Using standard results (Ang and Piazzesi, 2003) one can show that the logarithm of bond prices are affine functions of the risk factors

$$\log P_t^{(\tau)} = A_\tau + B_\tau' X_t, \tag{4}$$

where the scalar A_τ and the vector of loadings B_τ satisfy the recursions

$$A_\tau = A_{\tau-1} - \rho_0 + (\mu - \Gamma\lambda_0)' B_{\tau-1} + \frac{1}{2} B'_{\tau-1} \Gamma\Gamma' B_{\tau-1}, \tag{5}$$

$$B_\tau = (\Phi - \Gamma\lambda_1)' B_{\tau-1} - \rho_1, \tag{6}$$

with initial conditions $A_0 = 0$ and $B_0 = 0$.

The yield on a τ period zero-coupon bond at date t is thus

$$y_t^{(\tau)} = -\log(P_t^{(\tau)})/\tau = a_\tau + b'_\tau X_t, \tag{7}$$

where $a_\tau = -A_\tau/\tau$ and $b_\tau = -B_\tau/\tau$.

2.2   Spot price and implied cost-of-carry-convenience yield

Consider a storable commodity with spot price S_t and with a per-period net cost-of-carry of c_t, expressed as a continuously compounded rate of the spot commodity price. The net cost-of-carry of a storable commodity (cost-of-carry hereafter) represents the storage and insurance costs of physically holding the commodity net of any benefit or convenience yield on inventory during time t. It is the analog of the negative of the dividend yield of a stock and can be derived from equilibrium models under different assumptions about investment and storage costs (e.g. Routledge, Seppi, and Spatt, 2000).

We model seasonality in the spot price and the cost-of-carry in terms of loadings on X_t whose elements are periodic functions of time. This structure resembles the periodic linear model discussed in Hansen and Sargent (2014, Ch. 14). Yet, by specifying seasonal risk factors in X_t, our framework allows us to model stochastic changes in seasonality.

To capture time variation in the factor loadings, we index objects by the month (season) m_t associated with time t. Thus, $\{m_t\}$ is a periodic sequence mapping t into the set of months $\{1, 2, ..., 12\}$. We initialize the sequence by setting $m_t = t$ for $t = 1, 2, ..., 12$, and let $m_{t+12k} = m_t$ for every t and natural number k. We often use $\tilde{m}$ when referring to a generic month and impose the convention that $\tilde{m} + 1 = 1$ when $\tilde{m} = 12$.

The principle of no-arbitrage relates the time t spot price S_t with the time $t+1$ cost-of-carry and spot price, c_{t+1} and S_{t+1}. One could impose a process for c_{t+1} and use the conventional asset pricing formula to obtain the spot commodity price. Alternatively, one could use the insight of Casassus and Collin-Dufresne (2005) and obtain an implied cost-of-carry consistent with an arbitrary stochastic process for the spot commodity price. We follow the latter approach.

Assume that the log of the spot commodity price, $s_t = \log S_t$, is given by

$$s_t = \gamma_0 + \gamma_1^{m_t\prime} X_t, \tag{8}$$

where γ_0 is a scalar and $\gamma_1^{m_t}$ is a vector of factor loadings whose elements depend on the season m_t. The payoff from holding the commodity between periods t and $t+1$ is $\exp\left(-c_{t+1}\right) S_{t+1}$. Therefore, the principle of no arbitrage implies that the current spot price S_t is equal to the expected discounted value of its future payoff,

$$S_t = E_t\left[M_{t,t+1} e^{-c_{t+1}} S_{t+1}\right]. \tag{9}$$

The next proposition states that there is an affine and seasonal cost-of-carry process such that the pricing condition (9) is satisfied given the evolution for the spot price (8).

Proposition 1: *The cost-of-carry consistent with the commodity price (9) is an affine and periodic function of the state variables*

$$c_t = \psi_0^{m_t} + \psi_1^{m_t\prime} X_t, \tag{10}$$

where, for $\tilde{m} = 1, 2, ..., 12$, the scalar $\psi_0^{\tilde{m}}$ and vector $\psi_1^{\tilde{m}}$ satisfy

$$\psi_1^{\tilde{m}+1} = \gamma_1^{\tilde{m}+1} - \left((\Phi - \Gamma\lambda_1)\prime\right)^{-1}\left(\gamma_1^{\tilde{m}} + \rho_1\right),$$
$$\psi_0^{\tilde{m}+1} = \left(\gamma_1^{\tilde{m}+1} - \psi_1^{\tilde{m}+1}\right)\prime(\mu - \Gamma\lambda_0) + \frac{1}{2}\left(\gamma_1^{\tilde{m}+1} - \psi_1^{\tilde{m}+1}\right)\prime \Gamma\Gamma\prime\left(\gamma_1^{\tilde{m}+1} - \psi_1^{\tilde{m}+1}\right) - \rho_0.$$

2.3   Pricing commodity futures

A τ-period futures contract entered into at time t is an agreement to buy the commodity at time $t + \tau$ at the settlement price $F_t^{(\tau)}$ (the futures price).[4] The futures contract involves no initial cash flow and a payoff of $S_{t+\tau} - F_t^{(\tau)}$ at time $t + \tau$. Therefore, the principle of no-arbitrage implies

$$E_t\left[M_{t,t+\tau}(S_{t+\tau} - F_t^{(\tau)})\right] = 0.$$

Using the equivalent pricing condition for a futures contract entered into at time $t + 1$ with settlement date at time $t + \tau$ allows us to obtain the following recursive expression for the

---

[4]We follow the conventional approach of pricing futures contracts as if they were forwards. Futures prices may differ from forwards depending on the correlation between bond yields and commodity prices under the risk neutral measure. These differences have been found to be minor.

futures price,[5],[6]

$$F_t^{(\tau)} P_t^{(\tau)} = E_t \left[M_{t,t+1} F_{t+1}^{(\tau-1)} P_{t+1}^{(\tau-1)} \right]. \tag{11}$$

This expression together with equations (1)–(4) imply that the log futures price, $f_t^{(\tau)} = \log F_t^{(\tau)}$, is an affine and periodic function of the state variables,

$$f_t^{(\tau)} = C_\tau^{m_t} + D_\tau^{m_t \prime} X_t, \tag{12}$$

where $C_\tau^{\tilde{m}}$ and $D_\tau^{\tilde{m}}$ are given by

$$C_\tau^{\tilde{m}} = G_\tau^{\tilde{m}} - A_\tau \tag{13}$$

$$D_\tau^{\tilde{m}} = H_\tau^{\tilde{m}} - B_\tau, \tag{14}$$

A_τ and B_τ solve equations (5) and (6), and $G_\tau^{\tilde{m}}$ and $H_\tau^{\tilde{m}}$ solve the recursions

$$G_\tau^{\tilde{m}} = G_{\tau-1}^{\tilde{m}+1} - \rho_0 + (\mu - \Gamma\lambda_0)' H_{\tau-1}^{\tilde{m}+1} + \frac{1}{2} \left(H_{\tau-1}^{\tilde{m}+1} \right)' \Gamma\Gamma' H_{\tau-1}^{\tilde{m}+1}, \tag{15}$$

$$H_\tau^{\tilde{m}} = (\Phi - \Gamma\lambda_1)' H_{\tau-1}^{\tilde{m}+1} - \rho_1, \tag{16}$$

with initial conditions $G_0^{\tilde{m}} = \gamma_0$ and $H_0^{\tilde{m}} = \gamma_1^{\tilde{m}}$ for $\tilde{m} = 1, 2, ..., 12$.

3 Trading strategies and risk premia

Investors in the commodity futures markets are exposed to different risks. Szymanowska et al. (2014) relate simple trading strategies with different concepts of risk premia. In this section we express the different notions of risk premia in terms of the components of the affine model of futures prices. We show how to recover the risk premia associated with some popular trading strategies: holding a futures contract for a number of periods, spreading strategies, and strategies designed to exploit seasonal patterns. In addition, since all the strategies that we consider cost zero when they are entered into, any ex-ante expected return entirely reflects expected risk premia. Appendix B contains the derivation of the formulas.

Expected risk premia are usually estimated by running regressions of ex-post returns on a set of variables, or by computing average returns of portfolios sorted in terms of character-

---

[5] Details of what follows are provided in Appendix A.

[6] Alquist et. al (2013) study a multifactor affine model of oil futures using a setup different from ours. While we price commodity futures by discounting their dollar cash flows—as traditionally done in the finance literature—Alquist et. al (2013) assume that there are oil denominated bonds and introduce two pricing kernels, one expressed in dollars to price dollar bonds, and the other in units of oil to price oil bonds. Then they relate the oil and dollar pricing kernels as in Backus et. al (2001). In addition, their model does not consider seasonal fluctuations.

istics of the assets under consideration. A drawback of these methods, however, is that the estimated risk premia could be quite sensitive to seemingly minor details of the empirical implementation. For example, Cochrane and Piazzesi (2008) estimate risk premia in US forward interest rates first using a VAR in levels and next treating the forwards as a set of cointegrated variables. The two methods produce strikingly different results even though both are reasonable representations of the process followed by the forward interest rates. In contrast, with our framework we can estimate risk premia as a function of the parameters of the affine model and avoid the drawbacks of the conventional reduced form methods.

3.1 Holding strategies

The *1-period log holding return* (open a position on a τ-period futures at time t and close it at time $t+1$) is $f_{t+1}^{(\tau-1)} - f_t^{(\tau)}$. Using the affine structure, the time-t conditional expectation of this strategy is

$$E_t[f_{t+1}^{(\tau-1)} - f_t^{(\tau)}] = J_{\tau-1}^{m_{t+1}} + D_{\tau-1}^{m_{t+1}\prime}\Gamma\Lambda_t, \qquad (17)$$

where

$$J_{\tau-1}^{m_{t+1}} = \frac{1}{2}[B_{\tau-1}'\Gamma\Gamma'B_{\tau-1} - H_{\tau-1}^{m_{t+1}\prime}\Gamma\Gamma'H_{\tau-1}^{m_{t+1}}]$$

is a periodic Jensen inequality term. The second term, $D_{\tau-1}^{m_{t+1}\prime}\Gamma\Lambda_t$, captures the stochastic variation in expected risk premia over time.

The *spot premium* is the expected return of holding a 1-period futures contract until maturity. It is a particular case of the expected return (17),

$$E_t[s_{t+1} - f_t^{(1)}] = J_0^{m_{t+1}} + \gamma_1^{m_{t+1}\prime}\Gamma\Lambda_t, \qquad (18)$$

where we use that a 0-period futures is equivalent to the spot price, $s_{t+1} = f_{t+1}^{(0)}$. Note that the Jensen inequality term, $J_0^{m_{t+1}} = -\frac{1}{2}\gamma_1^{m_{t+1}\prime}\Gamma\Gamma'\gamma_1^{m_{t+1}}$, and the loadings on the prices of risk depend only on $\gamma_1^{m_{t+1}}$, the loading vector in the evolution of the spot price (8).

As in the bond pricing literature, the *term premium* is defined as the 1-period expected holding return of a τ-period futures contract in excess of the spot premium. In terms of the affine model, the term premium is

$$E_t[(f_{t+1}^{(\tau-1)} - f_t^{(\tau)}) - (s_{t+1} - f_t^{(1)})] = J_{\tau-1}^{m_{t+1}} - J_0^{m_{t+1}} + (D_{\tau-1}^{m_{t+1}\prime} - \gamma_1^{m_{t+1}\prime})\Gamma\Lambda_t. \qquad (19)$$

Another strategy is to open a position on a τ-period futures at time t and sell it as a $\tau - h$-period futures at time $t + h$. The ex-post *h-period log holding return* of this strategy

can be expressed as a sum of 1-period holding returns,

$$f_{t+h}^{(\tau-h)} - f_t^{(\tau)} = [f_{t+h}^{(\tau-h)} - f_{t+h-1}^{(\tau-h+1)}] + [f_{t+h-1}^{(\tau-h+1)} - f_{t+h-2}^{(\tau-h+2)}] + ... + [f_{t+1}^{(\tau-1)} - f_t^{(\tau)}].$$

The expected *h-period log holding return* follows from using the expected 1-period returns and the law of iterated expectations,

$$E_t[f_{t+h}^{(\tau-h)} - f_t^{(\tau)}] = \sum_{i=1}^{h} J_{\tau-i}^{m_{t+i}} + \sum_{i=1}^{h} D_{\tau-i}^{m_{t+i}'}\Gamma E_t[\Lambda_{t+i-1}]. \qquad (20)$$

A *short roll* strategy, $SR_{t,t+h}$, consists of rolling over 1-period contracts during h consecutive periods. The ex-post return of this strategy is

$$SR_{t,t+h} = (s_{t+1} - f_t^{(1)}) + (s_{t+2} - f_{t+1}^{(1)}) + ... + (s_{t+h} - f_{t+h-1}^{(1)}).$$

The expected return of the short roll strategy in terms of the affine model is

$$E_t[SR_{t,t+h}] = \sum_{i=1}^{h} J_0^{m_{t+i}} + \sum_{i=1}^{h} \gamma_1^{m_{t+i}'}\Gamma E_t[\Lambda_{t+i-1}].$$

Finally, an *excess holding* strategy, $XH_{t,t+h}$, consists of buying an h-period futures contract and shorting a short roll strategy. The expected return of this strategy is

$$E_t[XH_{t,t+h}] = \sum_{i=1}^{h} \left(J_{\tau-i}^{m_{t+i}} - J_0^{m_{t+i}}\right) + \sum_{i=1}^{h} \left(D_{\tau-i}^{m_{t+i}'} - \gamma_1^{m_{t+i}'}\right) \Gamma E_t[\Lambda_{t+i-1}].$$

3.2 Spreading strategies

Spread strategies are trading rules widely used by practitioners. They consist of buying and selling two futures contracts with different settlement dates, with the hope of earning a return by predicting changes in the slope of the futures curve.

Consider two futures contracts that mature $\bar{\tau}$ and $\underline{\tau} < \bar{\tau}$ periods ahead, and define the slope of the (log) futures curve using those contracts as

$$\mathsf{slope}_t^{\bar{\tau},\underline{\tau}} = f_t^{(\bar{\tau})} - f_t^{(\underline{\tau})}.$$

The *spread strategy 1* (SS1) is designed to produce a gain whenever the slope moves as predicted and loses otherwise. At time t, if $E_t(\mathsf{slope}_{t+1}^{\bar{\tau}-1,\underline{\tau}-1}) > \mathsf{slope}_t^{\bar{\tau},\underline{\tau}}$ buy the futures contract that matures in $\bar{\tau}$ periods and sell the contract with maturity $\underline{\tau}$; if $E_t(\mathsf{slope}_{t+1}^{\bar{\tau}-1,\underline{\tau}-1}) <$

$\mathsf{slope}_t^{\bar{\tau},\underline{\tau}}$ adopt the reverse strategy. At time $t+1$, close the positions and open new positions using contracts with the same maturities $\bar{\tau}$ and $\underline{\tau}$. The expected return of SS1 is the difference between two expected 1-period holding returns using futures that matures in $\bar{\tau}$ and $\underline{\tau}$ periods,

$$
\begin{aligned}
E_t(SS1_{t+1}) &= E_t\left[(f_{t+1}^{(\bar{\tau}-1)} - f_t^{(\bar{\tau})}) - (f_{t+1}^{(\underline{\tau}-1)} - f_t^{(\underline{\tau})})\right] \\
&= J_{\bar{\tau}-1}^{m_{t+1}} - J_{\underline{\tau}-1}^{m_{t+1}} + (D_{\bar{\tau}-1}^{m_{t+1}\prime} - D_{\underline{\tau}-1}^{m_{t+1}\prime})\Gamma\Lambda_t.
\end{aligned}
$$

In an alternative spread strategy, which we call *spread strategy 2* (SS2), the investor chooses the maturities $\bar{\tau}_t$ and $\underline{\tau}_t$ to maximize the expected difference between the slope of the futures curve at time $t+1$ relative to the slope at time t,

$$
(\bar{\tau}_t, \underline{\tau}_t) \in \arg\max\left\{E_t(\mathsf{slope}_{t+1}^{\bar{\tau}_t+1,\underline{\tau}_t-1}) - \mathsf{slope}_t^{\bar{\tau}_t,\underline{\tau}_t}\right\}.
$$

The expected return of this strategy is computed as that of SS1 but changing the maturities of the contracts in every period to maximize the difference in the expected slopes.

The previous spreading strategies maintain the spread positions for a single period. An investor may consider keeping open the position during h periods. As with SS1 and SS2, the investor can choose the maturities $\bar{\tau}$ and $\underline{\tau}$ arbitrarily or maximize the expected difference in the slope of the futures curve at time $t+h$ relative to that at time t. The expected return of this strategy is the difference in the expected h-period holding returns using contracts with maturities $\bar{\tau}$ and $\underline{\tau}$,

$$
E_t\left[f_{t+h}^{(\bar{\tau}-h)} - f_t^{(\bar{\tau})} + f_{t+h}^{(\underline{\tau}-h)} - f_t^{(\underline{\tau})}\right] = \sum_{i=1}^{h}(J_{\bar{\tau}-1}^{m_{t+i}} - J_{\underline{\tau}-1}^{m_{t+1}}) + \sum_{i=1}^{h}(D_{\bar{\tau}-i}^{m_{t+i}\prime} - D_{\underline{\tau}-i}^{m_{t+i}\prime})\Gamma E_t[\Lambda_{t+i-1}].
$$

3.3 Risk premia and the predictive content of futures prices

A traditional view that goes back to Cootner (1960) decomposes the futures price into the expected spot price and a risk premium or discount component

$$
f_t^{(\tau)} = E_t[s_{t+\tau}] + \pi_t^{(\tau)}. \tag{21}
$$

Fama and French (1987), among many others, test for the existence of time-varying risk premia as defined in equation (21). These tests can be interpreted as uncovering the predictive content of futures prices: if the risk premium is zero, futures prices are good predictors of spot prices. The empirical implementations of the tests, however, are inconclusive. Fama and French argue that their tests lack power to prove or disprove the existence of time varying risk premia. Alquist and Kilian (2010) find that oil futures are not good predictors

of subsequent oil prices. The results in Chinn and Coibion (2014), however, suggest that futures are good predictors of spot prices in some periods but not in others.

The affine model can shed light into this question. Shorting a τ-period futures contract and holding the short position until maturity delivers an expected return of $f_t^{(\tau)} - E_t\left[s_{t+\tau}\right]$, precisely the risk premium in equation (21). The risk premium $\pi_t^{(\tau)}$ is the negative of the expected h-period holding return (20) evaluated at $\tau = h$. Equivalently, we can write the risk premium $\pi_t^{(\tau)}$ using equations (1) and (8) as

$$f_t^{(\tau)} - E_t\left[s_{t+\tau}\right] = C_\tau^{m_t} - \gamma_0 - \gamma_1^{m_{t+\tau}\prime}(\sum_{j=0}^{\tau-1}\Phi^j)\mu + \left[D_\tau^{m_t\prime} - \gamma_1^{m_{t+\tau}\prime}\Phi^\tau\right]X_t.$$

4 A parsimonious model of commodity futures

Estimating unrestricted affine models of the sort described in Section 2 is problematic due to their large number of parameters. The likelihood function is flat and the forecasting performance of the model is poor. These are known problems in the literature on government bonds and, if anything, are exacerbated in models of commodity futures: an affine model of commodity futures prices includes an affine model of bond prices.[7] Most of the identification problems can be traced to the large number of parameters in the prices of risk matrix λ_1. For example, below we estimate a nine-factor model which, if left unrestricted, includes 81 free parameters just in λ_1. Some simplification is needed.

One possibility is to consider a small number of factors and to impose restrictions on their evolution under the physical measure. This is the approach taken by most of the literature on commodity futures. We follow a different methodology that is equivalent to imposing restrictions on the evolution of the risk factors under the risk neutral measure. We impose flexible, yet parsimonious, functional forms on the yield curve on government bonds and on the basis of the commodity futures, defined as the log-difference between the futures price and the spot commodity price, $f_t^{(\tau)} - s_t$. Following Miltersen and Schwartz (1998) and Trolle and Schwartz (2011) we *define* the cost-of-carry curve at time t (net of interest rates) as the value $u_t^{(\tau)}$ such that the basis of the commodity futures can be written as

$$f_t^{(\tau)} - s_t = \tau(y_t^{(\tau)} + u_t^{(\tau)}), \tag{22}$$

where $y_t^{(\tau)}$ is the yield curve on zero coupon bonds. This expression is the well known non-arbitrage relation between futures and spot prices when the per-period cost-of-carry c_t

---

[7]Hamilton and Wu (2012) discuss the identification problems of affine models of government bonds and Duffee (2002) highlights their poor forecasting performance.

is deterministic, in which case $u_t^{(\tau)} = (1/\tau) \sum_{j=1}^{\tau} c_{t+j}$. When c_t is stochastic, the relation between $u_t^{(\tau)}$ and c_t is more complex. Next, we impose Nelson and Siegel (1987) functional forms on $y_t^{(\tau)}$ and the non-seasonal component of $u_t^{(\tau)}$. Finally, we show that, under appropriate conditions, the model represented by (22) and the Nelson and Siegel expressions for $y_t^{(\tau)}$ and $u_t^{(\tau)}$ *is* an affine model of commodity futures.

4.1 Seasonality

Seasonal fluctuations can be modeled as deterministic or stochastic patterns that repeat once every year. Let us decompose an arbitrary stochastic process z_t into its seasonal and non-seasonal components z_t^n and z_t^s,

$$z_t = \mathsf{z}_t^n + \mathsf{z}_t^s.$$

If the seasonal pattern is deterministic, z_t^s is a periodic sequence of period 12, so that $\mathsf{z}_t^s = \mathsf{z}_{t+12k}^s$ for any integer k.

Researchers often use dummy variables to model seasonality, imposing that the sum of the seasonal components is zero. Alternatively, we can model seasonality in terms of trigonometric functions,

$$\mathsf{z}_t^s = \sum_{j=1}^{6} \left[\xi_j \cos(\tfrac{2\pi j}{12} m_t) + \xi_j^* \sin(\tfrac{2\pi j}{12} m_t) \right],$$

where $2\pi j/12$ are seasonal frequencies and ξ_j and ξ_j^* are parameters. The two representations of seasonality are equivalent: the right side of the last equation is the Fourier series representation of the periodic sequence z_t^s. The trigonometric approach, however, has two advantages. First, it emphasizes the cyclical nature of the seasonal factor. The seasonal effect z_t^s is the sum of six deterministic cycles with periods of $12/j$ months, for $j = 1, 2, ..., 6$. The frequency $2\pi/12$ corresponds to a period of 12 months and is known as the fundamental frequency. The remaining frequencies, called harmonics, represent waves with periods of less than a year. Second, the trigonometric approach allows for a more parsimonious representation of seasonality. For example, one may want to emphasize only the fundamental frequency or perhaps ignore seasonal fluctuations associated with some of the harmonics.

Seasonal fluctuations in many commodity prices are not perfectly predictable. Following Hannan (1964), we model stochastic seasonality by letting the parameters ξ_j and ξ_j^* evolve

as random walk processes. The seasonal component is assumed to be

$$z_t^s = \sum_{j=1}^{6} \left[\xi_{jt} \cos(\tfrac{2\pi j}{12} m_t) + \xi_{jt}^* \sin(\tfrac{2\pi j}{12} m_t) \right], \tag{23}$$

where, for $j = 1, 2, ..., 6$,

$$\begin{aligned} \xi_{jt} &= \xi_{jt-1} + \nu_{jt}, \\ \xi_{jt}^* &= \xi_{jt-1}^* + \nu_{jt}^*. \end{aligned}$$

The shocks ν_{jt} and ν_{jt}^* are normally distributed with variances σ_j^2 and σ_j^{*2}, and mutually orthogonal. This representation models stochastic seasonality in terms of periodic loadings on random walk processes. If only the fundamental frequency matters (when $\xi_{jt} = \xi_{jt}^* = 0$ for $j = 2, ..., 6$) the seasonality process collapses to

$$z_t^s = \xi_t \cos(\tfrac{2\pi}{12} m_t) + \xi_t^* \sin(\tfrac{2\pi}{12} m_t), \tag{24}$$

where ξ_t and ξ_t^* are two independent random walks. As we show below, since seasonal fluctuations in heating oil prices seem to follow the simpler process (24), for the rest of the paper we focus on seasonal fluctuations associated only with the fundamental frequency.

4.2 A Nelson and Siegel representation of the affine model

Our objective is to construct an affine model of commodity futures that is parsimonious yet flexible enough to match the different shapes of the futures curve and yield curve over time. We first write the log-basis (22) emphasizing the contribution of the seasonal factors,

$$f_t^{(\tau)} = \beta_{0t} + \tau(y_t^{(\tau)} + \tilde{u}_t^{(\tau)}) + e^{-\omega\tau} \left[\xi_t \cos(\tfrac{2\pi}{12} m_{t+\tau}) + \xi_t^* \sin(\tfrac{2\pi}{12} m_{t+\tau}) \right], \tag{25}$$

where we interpret β_{0t} as the deseasonalized spot commodity factor and $\tilde{u}_t^{(\tau)}$ as cost-of-carry curve net of any stochastic seasonal component. The last term on the right side reflects the contribution of the seasonal factors ξ_t and ξ_t^* to futures prices of different maturities. When $\tau = 0$ the futures price is the spot commodity price and equation (25) becomes

$$s_t = \beta_{0t} + \xi_t \cos(\tfrac{2\pi}{12} m_t) + \xi_t^* \sin(\tfrac{2\pi}{12} m_t), \tag{26}$$

which justifies calling β_{0t} the deseasonalized spot factor. To extract the seasonality of a futures contract with τ months to maturity, we compute the expected seasonal component

at time $t+\tau$ conditional on information at time t, and then multiply the resulting expression by a discounting factor $e^{-\omega\tau}$.[8]

We parametrize the yield curve $y_t^{(\tau)}$ using a dynamic Nelson and Siegel (DNS) model. The DNS parametrization is a three factor model which fits well the cross section and time series of zero-coupon bond yields (Diebold and Li, 2006). Yet, in its basic representation the DNS model does not rule out arbitrage opportunities. We follow Christensen et al. (2011) and augment the Nelson and Siegel equation with a maturity-specific constant a_τ that renders the model arbitrage free. The yield curve is thus parametrized as

$$y_t^{(\tau)} = a_\tau + \delta_{1t} + \left(\frac{1-e^{-\theta_1\tau}}{\theta_1\tau}\right)\delta_{2t} + \left(\frac{1-e^{-\theta_1\tau}}{\theta_1\tau} - e^{-\theta_1\tau}\right)\delta_{3t}, \qquad (27)$$

where δ_{1t}, δ_{2t}, and δ_{3t} are latent variables interpreted as level, slope and curvature factors, and the θ_1 determines the shape of the loadings on the factors δ_{2t} and δ_{3t}. The traditional Nelson and Siegel model sets $a_\tau = 0$ for all τ.

We also impose a DNS structure on the cost-of-carry curve,

$$\tilde{u}_t^{(\tau)} = g_\tau^{m_t} + \beta_{1t} + \left(\frac{1-e^{-\theta_2\tau}}{\theta_2\tau}\right)\beta_{2t} + \left(\frac{1-e^{-\theta_2\tau}}{\theta_2\tau} - e^{-\theta_2\tau}\right)\beta_{3t}. \qquad (28)$$

where β_{1t}, β_{2t}, and β_{3t} are level, slope and curvature factors. Even though $\tilde{u}_t^{(\tau)}$ is independent of any seasonal stochastic factor, the term $g_\tau^{m_t}$ depends deterministically on the season m_t. Without this term, the Nelson and Siegel model cannot be rendered arbitrage-free.[9]

We now show that the augmented Nelson and Siegel model can be interpreted as a restricted version of the general affine arbitrage-free model described in Section 2. The state vector of the arbitrage-free DNS model is

$$X_t = [\delta_{1t}, \delta_{2t}, \delta_{3t}, \beta_{0t}, \beta_{1t}, \beta_{2t}, \beta_{3t}, \xi_t, \xi_t^*]',$$

composed of the three yield curve factors, the deseasonalized spot price, the three cost-of-carry factors, and the two seasonal factors. Our task is to find parameters of the affine model

---

[8]This follows because the seasonal factors are random walks,

$$E_t\left[\xi_{t+\tau}\cos(\tfrac{2\pi}{12}m_{t+\tau}) + \xi_{t+\tau}^*\sin(\tfrac{2\pi}{12}m_{t+\tau})\right] = \xi_t\cos(\tfrac{2\pi}{12}m_{t+\tau}) + \xi_t^*\sin(\tfrac{2\pi}{12}m_{t+\tau}).$$

[9]Our Nelson and Siegel parametrization differs substantially from that used by Karstanje et al. (2015). While they impose the usual three factor Nelson and Siegel structure to the log of the futures curve, our model distinguishes the separate contributions of the spot price, the yield curve, and the cost-of-carry curve. Furthermore, we show conditions under which the Nelson and Siegel parametrization is arbitrage free and allow for stochastic seasonality.

(i.e. market prices of risk λ_0 and λ_1, parameters of the short rate equation ρ_0 and ρ_1, and of the log-spot price γ_0 and $\gamma_1^{m_t}$) such that the yields and futures prices adopt the Nelson and Siegel forms (25), (27), and (28). The conditions are summarized in the next proposition, proved in Appendix A.

Proposition 2. *Consider any vector* $\mu^Q \in \mathbb{R}^9$ *and a matrix* Φ^Q *defined as*

$$\Phi^Q = \begin{bmatrix}
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & e^{-\theta_1} & \theta_1 e^{-\theta_1} & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & e^{-\theta_1} & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & \frac{1-e^{-\theta_1}}{\theta_1} & \left(\frac{1-e^{-\theta_1}}{\theta_1} - e^{-\theta_1}\right) & 1 & 1 & \frac{1-e^{-\theta_2}}{\theta_2} & \left(\frac{1-e^{-\theta_2}}{\theta_2} - e^{-\theta_2}\right) & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & e^{-\theta_2} & \theta_2 e^{-\theta_2} & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & e^{-\theta_2} & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & e^{-\omega} & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & e^{-\omega}
\end{bmatrix} \quad (29)$$

where $\theta_1, \theta_2, \omega > 0$. *Define the following prices of risk parameters*

$$\lambda_0 = \Gamma^{-1}\left(\mu - \mu^Q\right) \ \text{and} \ \lambda_1 = \Gamma^{-1}\left(\Phi - \Phi^Q\right),$$

parameters of the short rate equation

$$\rho_0 = 0 \ \text{and} \ \rho_1 = \left[1, \tfrac{1-e^{-\theta_1}}{\theta_1}, \tfrac{1-e^{-\theta_1}}{\theta_1} - e^{-\theta_1}, 0, 0, 0, 0, 0, 0\right]',$$

and parameters of the log-spot price process

$$\gamma_0 = 0 \ \text{and} \ \gamma_1^{\tilde{m}} = \left[0, 0, 0, 1, 0, 0, 0, \cos(\tfrac{2\pi}{12}\tilde{m}), \sin(\tfrac{2\pi}{12}\tilde{m})\right]$$

for $\tilde{m} = 1, 2, ..., 12$. *Then, the yields and futures prices derived from the affine model adopt the Nelson and Siegel parametrization (25), (27), and (28).*

The arbitrage-free Nelson and Siegel representation imposes strong restrictions on the risk neutral evolution of the state variables, $X_{t+1} = \mu^Q + \Phi^Q X_t + \Gamma \eta_{t+1}^Q$, where $\mu^Q = \mu - \Gamma\lambda_0$, $\Phi^Q = \Phi - \Gamma\lambda_1$, and $\eta_{t+1}^Q \sim N(0, I)$. With the proposed matrix Φ^Q and parameters ρ_0, ρ_1, γ_0, and $\gamma_1^{\tilde{m}}$, the recursions (5), (6), (15), and (16) guarantee that the factor loadings of the bond yields and futures prices adopt the Nelson and Siegel parametrization.

The arbitrage-free Nelson and Siegel model greatly reduces the number of parameters to estimate. Without restrictions, λ_1 is a 9×9 matrix of parameters. With the Nelson

and Siegel structure, those 81 parameters collapse to 3: θ_1, θ_2, and ω. Given Φ^Q and the parameters of the physical measure, λ_1 is pinned down by $\lambda_1 = \Gamma^{-1}\left(\Phi - \Phi^Q\right)$. Given a risk-neutral intercept μ^Q, λ_0 is determined as $\lambda_0 = \Gamma^{-1}\left(\mu - \mu^Q\right)$. Rather than imposing dubious identifying restrictions on λ_1 and λ_0, identification of the Nelson and Siegel model requires imposing a few restrictions on μ^Q.[10] We view this massive reduction in the number of free parameters as the main advantage of the Nelson and Siegel approach.

4.3 Risk premia under the Nelson and Siegel representation

We analyze the implications of the Nelson and Siegel representation for the risk premia and the compensation for exposure to the different shocks in the model. For simplicity, we assume that the state matrix Φ is diagonal, an assumption that we drop in the empirical section below. There is, however, valuable intuition obtained from this example.

Risk premia vary over time because market prices of risk fluctuate and because the factor loadings are periodic functions of time. Most of the variation in risk premia, however, comes from variations in the market prices of risk. Using the definition of the market prices of risk Λ_t, we write the expected h-period holding return (20) of a τ-period futures contract as

$$E_t[f_{t+h}^{(\tau-h)} - f_t^{(\tau)}] = \sum_{i=1}^{h}\left(J_{\tau-i}^{m_{t+i}} + D_{\tau-i}^{m_{t+i}\prime}\Gamma[\lambda_0 + \lambda_1(\sum_{j=0}^{i-2}\Phi^j)\mu]\right) + \left(\sum_{i=1}^{h}D_{\tau-i}^{m_{t+i}\prime}\Gamma\lambda_1\Phi^{i-1}\right)X_t.$$

Under risk neutrality ($\lambda_0 = 0$ and $\lambda_1 = 0$) the expected h-period holding return is the usual Jensen inequality term, which varies over time due to the periodic nature of $J_{\tau-1}^{m_{t+1}}$. When $\lambda_1 \neq 0$, fluctuations in the state variables X_t drive fluctuations in expected returns.

With the Nelson and Siegel representation (25), (27), and (28) (using ϕ_{ii} to denote

---

[10] We discuss these restricciones in section 5.1.

element (i, i) of Φ), the loadings on X_t of the expected h-period holding return simplify to

Factor	Loading
δ_{1t}	$(\tau - h)\,\phi_{11}^h - \tau$
δ_{2t}	$\left(\frac{1-e^{-\theta_1(\tau-h)}}{\theta_1}\right)\phi_{22}^h - \left(\frac{1-e^{-\theta_1\tau}}{\theta_1}\right)$
δ_{3t}	$\left(\frac{1-e^{-\theta_1(\tau-h)}}{\theta_1} - (\tau - h)\,e^{-\theta_1(\tau-h)}\right)\phi_{33}^h - \left(\frac{1-e^{-\theta_1\tau}}{\theta_1} - \tau e^{-\theta_1\tau}\right)$
β_{0t}	$\left(\phi_{44}^h - 1\right)$
β_{1t}	$(\tau - h)\,\phi_{55}^h - \tau$
β_{2t}	$\left(\frac{1-e^{-\theta_2(\tau-h)}}{\theta_2}\right)\phi_{66}^h - \left(\frac{1-e^{-\theta_2\tau}}{\theta_2}\right)$
β_{3t}	$\left(\frac{1-e^{-\theta_2(\tau-h)}}{\theta_2} - (\tau - h)\,e^{-\theta_2(\tau-h)}\right)\phi_{77}^h - \left(\frac{1-e^{-\theta_2\tau}}{\theta_2} - \tau e^{-\theta_2\tau}\right)$
ξ_t	$e^{-\omega\tau}\left(e^{\omega h} - 1\right)\cos(\frac{2\pi}{12}m_{t+\tau})$
ξ_t^*	$e^{-\omega\tau}\left(e^{\omega h} - 1\right)\sin(\frac{2\pi}{12}m_{t+\tau}).$

The first three rows are the loading on the yield curve factors δ_{1t}, δ_{2t}, and δ_{3t} (level, slope, and curvature); the fourth row is the loading on the non-seasonal spot factor β_{0t}; the next three rows are the loading on the cost-of-carry factors β_{1t}, β_{2t}, and β_{3t} (level, slope, and curvature); and the last two rows are the loadings on the seasonal factors ξ_t and ξ_t^*.

Figure 1 displays the loadings on the yield curve and cost-of-carry factors 1-month holding return (top panels) and the 1-year holding return (bottom panels) as a function of the maturity τ of the contract. [11]

Commodity prices often manifest near-random walk behavior. If the spot commodity factor has a unit root ($\phi_{44} = 1$), β_{0t} does not contribute at all to fluctuations in expected holding returns. The spot price today is the expectation of the spot price tomorrow. Therefore, the only way for the current spot price to affect expected holding returns is through its interaction with the other factors—an interaction that we allow in the empirical section below. If the spot price is less persistent than a random walk, a positive shock to the spot factor decreases expected returns and the effect is larger for longer maturity contracts. When the holding period is 1 month, $h = 1$, the impact on the holding returns of shocks to β_{0t} is small: $\phi_{44} - 1 = -0.02$. When the holding period increases to $h = 12$, the impact of shocks to β_{0t} drops to $\phi_{44}^{12} - 1 \approx -0.22$.

We now focus on the 1-month holding return ($h = 1$) and the yield curve factors. In the data, the level factor δ_{1t} is very persistent with an autoregressive parameter near one. When $\phi_{11} \approx 1$, the loading of the 1-month expected holding return on δ_{1t} is -1: changes in

---

[11]We set the values of ϕ_{ii} by running independent autoregressions on the factors that we obtain in the empirical section below. The parameter values are $\phi_{11} = 0.99$, $\phi_{22} = 0.94$, $\phi_{33} = 0.96$, $\phi_{44} = 0.98$, $\phi_{55} = 0.83$, $\phi_{66} = 0.69$, $\phi_{77} = 0.68$, $\theta_1 = 0.07$, $\theta_2 = 0.25$, and $\omega = 0.008$. The parameter θ_1, θ_2, and ω are also those estimated below.

the level of interest rates translate into equal changes but of the opposite sign in expected holding returns. If interest rates fall over time, as they do in our sample period, the price of the futures contract also fall (equation (25)) increasing the risk of holding the contract. When $\phi_{11} < 1$, the loading on δ_{1t} is negative and increases with the maturity of the contract, as shown in the left upper panel of Figure 1.[12] The contribution of the slope factor δ_{2t} depends on the maturity of the contract and on the parameter θ_1. If $\phi_{22} \approx 1$ and $\theta_1 \to 0$, the loading of the 1-month expected return on the slope factor is -1 for all maturities. For the calibrated example, the contribution of the slope factor is -1 for a 1-month maturity contract and increases to -0.55 when $\tau = 24$. Thus, the contribution of the slope factor is positive when the yield curve is upward sloping and negative when it is inverted. The contribution of the curvature factor δ_{3t} depends on the values of ϕ_{33} and θ_1. If $\phi_{33} \approx 1$ and $\theta_1 \to 0$, the loading on δ_{3t} is zero for all maturities. In the calibrated example, the loading on the curvature factor decreases with the maturity of the contract, going from zero when $\tau \to 0$ to about -0.6 when $\tau = 24$. Shocks to δ_{3t} increases risk premia when the curvature of the yield curve is positive, and decreases risk premia when the curvature is negative.

The analysis of the contribution of the commodity factors β_{1t}, β_{2t}, and β_{3t} is similar to that of the yield curve factors. The main differences are that the parameter θ_2 is substantially larger than that of the yield curve, and that the calibrated level, slope, and curvature cost-of-carry factors are less persistent than the equivalent yield curve factors. These differences have a major impact on the loadings. First, increases in the level of the cost-of-carry, β_{1t}, lead to a large drop in 1-month expected returns, ranging from -1 when $\tau = 1$ to almost -5 when $\tau = 24$. Second, the contribution of the cost-of-carry slope factor β_{2t} decreases from -1 to -1.24 as the maturity of the contract increases from $\tau = 1$ to $\tau = 24$. Third, the contribution of the curvature factor β_{3t} is also negative and decreasing in the maturity of the contract. The total contribution of the three cost-of-carry factors depend on their sign and volatility. In the empirical section below we find that β_{1t} is mostly negative and that β_{2t} and β_{3t} change sign often over time. The slope factor β_{2t} is the most volatile of the cost-of-carry factors, making its contribution to the expected holding return the largest of the three.

This example also shows that the contribution of the seasonal factors to the variations in risk premia for contracts of any maturity is minor: the term multiplying the sines and cosines in the 1-month expected holding return is tiny, $e^{-\omega\tau}(1 - e^{-\omega}) \approx \omega e^{-\omega\tau} = 0.008 e^{-0.008\tau}$.

The contribution of the different factors on the 12-month expected holding return is

---

[12] This result is different from the effect of the level factor on expected excess return of holding zero coupon government bonds. If the level factor has a unit root and is independent of the slope and curvature factors, changes in this factor are exactly cancelled out by movements in the short interest rate for all maturities.

qualitatively similar to those on the 1-month expected return (lower panel of Figure 1). The difference is that the relative contribution of the level factors (both from the yield curve and the cost-of-carry) are now substantially larger, specially for longer dated contracts.

This example illustrates that fluctuations in the level, slope, and curvature factors of the yield curve have a relevant contribution to expected holding returns. The loadings on the level and curvature factors increase with the maturity of the contract, while that on the slope factor decreases with maturity. The contribution of the spot commodity factor is small and constant, while that of the level of the cost-of-carry inscreases (in absolute value) substantially with the maturity of the contract. The loadings on the slope and curvature factors of the cost-of-carry also increase with maturity, although proportionally less than that on the level factor. Of course, expected risk premia is also determined by the volatility of the factors. The empirical results that follow indicate that the cost-of-carry factors are much more volatile than those of the yield curve. Yet, the contribution of the interest rate factors is still substantial. We consider this finding an important contribution of our analysis since interest rate risks have received relatively less attention than other factors in the commodity futures literature.

5   Estimation method

We estimate the model using the method of maximum likelihood. Since the state variables are unobserved, we use the Kalman filter to evaluate the prediction error decomposition of the likelihood function. The Kalman filter also allows us to handle missing observations, a common feature in the market of commodity futures.[13] We initialize the Kalman filter with a diffuse prior due to the two random walk components associated with the seasonal factors and to account for a possible unit root in the heating oil spot price.

The state variables X_t follow the first order vector autoregressive process (1). The observation equation consists of the arbitrage-free Nelson and Siegel parametrization of the log futures and bond yields evaluated at a set of maturities $\tau \in \mathcal{T}$.[14] Since there are more observed maturities than factors, we augment the observation equations with uncorrelated measurement errors $\varepsilon_{ft}^{(\tau)}$ and $\varepsilon_{yt}^{(\tau)}$ to avoid the problem of stochastic singularity,

$$f_t^{(\tau)} = C_\tau^{m_t} + D_\tau^{m_t\prime} X_t + \varepsilon_{ft}^{(\tau)} \tag{30}$$

$$y_t^{(\tau)} = a_\tau + b_\tau' X_t + \varepsilon_{yt}^{(\tau)}. \tag{31}$$

---

[13]We estimate the model using an unbalanced panel (see the data description below). When there are missing observations we evaluate the likelihood function as explained in Harvey (1989).

[14]We allow for bond yields and commodity futures to be observed at some periods and for some maturities in $\mathcal{T}$ but not for others.

The intercept and factor loadings satisfy the Nelson and Siegel functional forms (25), (27), and (28). Namely, $C_\tau^{m_t} = G_\tau^{m_t} - A_\tau$, $D_\tau^{m_t} = H_\tau^{m_t} - B_\tau$, $a_\tau = -A_\tau/\tau$, $b_\tau = -B_\tau/\tau$,

$$H_\tau^{m_t} = \left[0, 0, 0, 1, \tau, \frac{1-e^{-\theta_2\tau}}{\theta_2}, \frac{1-e^{-\theta_2\tau}}{\theta_2} - \tau e^{-\theta_2\tau}, e^{-\omega\tau}\cos\left(\frac{2\pi}{12}\left(m_t+\tau\right)\right), e^{-\omega\tau}\sin\left(\frac{2\pi}{12}\left(m_t+\tau\right)\right) \right],$$

$$B_\tau = \left[-\tau, -\frac{1-e^{-\theta_1\tau}}{\theta_1}, -\left(\frac{1-e^{-\theta_1\tau}}{\theta_1} - \tau e^{-\theta_1\tau}\right), 0, 0, 0, 0, 0, 0 \right]',$$

and A_τ and $G_\tau^{m_t}$ satisfy the recursions (5) and (15).

We note here two things. First, the parameters θ_1 and θ_2 in the futures equation (30) are identified because θ_1 is also a parameter of the yields equation (31). Second, stochastic seasonality enters into the measurement equation of the futures prices as periodic loading on the factors $\boldsymbol{\xi}_t$ of the state vector X_t.

We estimate the model using a two-step procedure. In the first step, we estimate the block of bond yields. In a second step, we estimate the block of futures prices conditioning on the estimates obtained in the first step. We follow the two step procedure to avoid overfitting the futures block of the model at the expense of distorting the parameters of the yield curve. Our data set contains 24 maturities of futures contracts but only 7 maturities of bond yields. Furthermore, the range and volatility of the commodity futures returns are an order of magnitude larger than those of the bond yields. Therefore, the joint estimation of the model would bias the yield curve parameters to provide a better fit of the futures block of the data. As a result, the estimate of θ_1 that simultaneously enter into the yield curve block and futures block of the model may differ dramatically from that we would obtain by only estimating a panel of bond yields. By estimating the model in two steps, we avoid distorting the yield curve and make sure that the estimated yield curve factors are consistent with those estimated in the bond pricing literature.

The two-step procedure restricts the parameters of the state equation to satisfy

$$\mu = \begin{bmatrix} \boldsymbol{\mu}_\delta \\ \boldsymbol{\mu}_\beta \\ \mathbf{0} \end{bmatrix}; \quad \Phi = \begin{bmatrix} \Phi_{\delta\delta} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \Phi_{\beta\beta} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I}_2 \end{bmatrix}; \quad \Gamma = \begin{bmatrix} \Gamma_{\delta\delta} & \mathbf{0} & \mathbf{0} \\ \Gamma_{\beta\delta} & \Gamma_{\beta\beta} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \Gamma_{\xi\xi} \end{bmatrix} \tag{32}$$

where $\boldsymbol{\mu}_\delta$ is a 3×1 vector, $\boldsymbol{\mu}_\beta$ is a 4×1 vector, $\Phi_{\delta\delta}$ is a 3×3 matrix, $\Phi_{\beta\beta}$ is a 4×4 matrix, $\mathbf{I}_2$ is a 2×2 identity matrix, $\Gamma_{\delta\delta}$ is a 3×3 lower triangular matrix, $\Gamma_{\beta\delta}$ is a 4×3 matrix, $\Gamma_{\beta\beta}$ is a 4×4 lower triangular matrix, and $\Gamma_{\xi\xi}$ is a 2×2 diagonal matrix with entries σ_ξ and σ_{ξ^*}.

In the first step, we estimate θ_1 and the parameters of the state equation $\boldsymbol{\mu}_\delta$, $\Phi_{\delta\delta}$, and $\Gamma_{\delta\delta}$ using the yield curve block of the model. With the estimated factors $\hat{\boldsymbol{\delta}}_t$ and parameter $\hat{\theta}_1$ we construct a fitted yield curve $\hat{y}_t^{(\tau)}$. In the second step, we subtract the fitted bond

yields from the futures equation (30). The measurement equation in the second step is thus

$$f_t^{(\tau)} - \tau \widehat{y}_t^{(\tau)} = G_\tau^{m_t} + H_\tau^{m_t \prime} X_t + \varepsilon_{ft}^{(\tau)}, \tag{33}$$

which we use to estimate the parameters θ_2, ω, $\boldsymbol{\mu}_\beta$, $\Phi_{\beta\beta}$, $\Gamma_{\beta\beta}$, and $\Gamma_{\xi\xi}$ (and the variance of the measurement errors). Note that with the two-step procedure we do not estimate the submatrix $\Gamma_{\beta\delta}$.

5.1 Identification

Affine term structure models suffer from severe identification problems (Dai and Singleton, 2000; Hamilton and Wu, 2012). Most of these problems come from two sources: first, the process for the market prices of risk ($\Lambda_t = \lambda_0 + \lambda_1 X_t$) has many free parameters; and second, the parameters of the short interest rate ($r_t = \rho_0 + \rho_1' X_t$) are not identified. Solving the identification problem usually entails setting $\rho_0 = 0$ and imposing zeros on the parameters of the state equation μ and Φ (Dai and Singleton, 2000) or on the parameters of the market prices of risk λ_0 and λ_1 (Ang and Piazzesi, 2003).

The Nelson and Siegel model is a simple way to solve the identification problem. Given the parameters of the state equation (μ, Φ, and Γ) there is a one to one mapping between the parameters of the risk-neutral measure, μ^Q and Φ^Q, and the parameters of the market prices of risk λ_0 and λ_1—namely, $\lambda_0 = \Gamma^{-1} \left(\mu - \mu^Q \right)$ and $\lambda_1 = \Gamma^{-1} \left(\Phi - \Phi^Q \right)$. In deriving the Nelson and Siegel representation we left μ^Q unrestricted but constrained Φ^Q to depend on 3 parameters rather than 81. This is equivalent to imposing 78 restrictions on the matrix λ_1, all in the context of a model that matches well the cross-section of futures prices and bond yields. To identify the model, we impose a few additional assumptions on the risk-neutral intercept μ^Q and the parameter ρ_0 of the short rate equation:

- **Assumption 1:** The measurement equation is (31), where $a_\tau = -A_\tau/\tau$ and

$$A_\tau = A_{\tau-1} - \rho_0 + \boldsymbol{\mu}_\delta^{Q\prime} B_{\tau-1} + \frac{1}{2} B_{\tau-1}' \Gamma\Gamma' B_{\tau-1}.$$

  As in Dai and Singleton (2000), we set $\rho_0 = 0$. In addition, since $\boldsymbol{\mu}_\delta^{Q\prime} B_{\tau-1}$ is a scalar, we can identify a single parameter in $\boldsymbol{\mu}_\delta^Q$. We thus set $\boldsymbol{\mu}_\delta^Q = [\mu_{\delta,1}^Q, 0, 0]'$, and estimate $\mu_{\delta,1}^Q$ along with the other parameters of the model.

- **Assumption 2:** The measurement equation is (33), where $g_\tau^{m_t} = G_\tau^{m_t}/\tau$ and

$$G_\tau^{\tilde{m}} = G_{\tau-1}^{\tilde{m}+1} - \delta_0 + \boldsymbol{\mu}_\beta^{Q\prime} H_{\tau-1}^{\tilde{m}+1} + \frac{1}{2} H_{\tau-1}^{\tilde{m}+1 \prime} \Gamma\Gamma' H_{\tau-1}^{\tilde{m}+1}$$

for any month $\tilde{m}$. Here, $\boldsymbol{\mu}_\beta^{Q\prime} H_{\tau-1}^{\tilde{m}+1}$ is a scalar varying with the particular month $\tilde{m}$. Therefore, one possible identification strategy entails setting $\boldsymbol{\mu}_\beta^Q = [0, 0, 0, 0, \mu_{\beta,5}^Q, \mu_{\beta,6}^Q]'$ and estimating $\tilde{\mu}_\beta^Q = \mu_{\beta,5}^Q = \mu_{\beta,6}^Q$ as a free parameter.

6 Model estimation

We first describe the data and show that seasonal fluctuations in heating oil futures prices vary over time. Next, we estimate two version of the affine model of futures prices: one with stochastic and one with deterministic seasonality. Using a number of tests, we argue that the model with deterministic seasonality is misspecified. Once we are confident that the data supports the model with stochastic seasonality, we use the model to study risk premia in commodity futures and to tests two traditional theories of commodity futures prices.

6.1 Data description

We estimate the model using monthly data on heating oil futures prices and U.S. zero coupon bond prices. Heating oil is the second most important petroleum product after natural gas in the United States. It is mostly used to fuel building furnaces and its price displays a pronounced seasonality. The seasonality in heating oil prices varies over time as it depends on the severity of the winter season in the U.S. and, more recently, China. Heating oil futures are traded on the New York Mercantile Exchange and contracts are for delivery in New York Harbor. The last trading day of heating oil futures contracts is the last business day of the month preceding the delivery month. Delivery can be made between the sixth business day and the last day before the last business day of the delivery month (NYMEX, 2009).

We construct monthly series from daily prices for the period January 1984–July 2012.[15] Since not all contracts trade every day, we set the monthly price equal to the available price closest to the last business day of the month. We include contracts with maturities up to 24 months, although available maturities have varied over time. In the early part of our sample, contracts were available with maturities up to 12 months. New contracts appeared in 1991 with maturities up to 18 months and it was not until 2007 that longer maturities contracts begun trading. We drop from our sample the contracts closest to expiration and label a 1 month futures contract those that expire in the month after the next month, and likewise for the longer maturities. We impose this convention for two reasons. First, delivery for contracts that are about to expire can be made as early as six days after the last trading

---

[15]Futures data come from the commercial provider `www.price-data.com`.

day. Second, futures contracts become very illiquid a couple of weeks before expiration.[16]

For bond yields, we use last day of month data on U.S. treasury yields of fixed maturities of 3, 6, 12, 24, 36, 48, and 60 months. The data for the 3 and 6 months yields are obtained from the Federal Reserve Bank of St. Louis and the remaining maturities are unsmoothed Fama and Bliss yields. To evaluate the model, we also use data on inventories and on the type of traders that operate in the futures market. Inventories data are U.S. Ending Stock of Distillate Fuel Oil, obtained from the U.S. Energy Information Administration. The data for the type of traders are from the CFTC's Commitments of Traders Reports.[17]

6.2   Seasonality in futures prices

We consider three tests of the null hypothesis that seasonality in the log of heating oil futures prices is deterministic. The alternative hypothesis is that seasonal fluctuations are driven by random walk processes. We apply Canova and Hansen (1995) nonparametric test for parameter stability and its spectral extension of Busetti and Harvey (2003).[18] The third is a parametric test also proposed by Busetti and Harvey (2003). Under the null hypothesis, the tests are distributed as a generalized Von Mises with degrees of freedom equal to the number of seasonal factors. Table 1 shows the results of the tests for selected maturities of the contracts. In most cases, the tests reject the null of deterministic seasonality.

Once we are confident that seasonality is stochastic, we use the parametric test of Busetti and Harvey (2003) to determine the number of seasonal factors. Table 2 shows the Akaike and Bayesian information criteria for different models of seasonality and maturities. In all the cases, the information criteria select a model of seasonality based on fluctuations only at the fundamental frequency and with a different volatility parameter for each seasonal factor.

These results suggest that the common practice of extracting the seasonal component from futures prices prior to any analysis is flawed. First, by deseasonalizing the data prior to the analysis one is unable to measure the contribution of seasonal shocks to risk premia. Second, the results of seasonality tests do not coincide across the different maturities of the contracts. While seasonal variations have a prominent role in the cross-section of the futures curve, they contribute a small fraction to the time series variation of the futures prices. Moreover, extracting seasonal components in isolation does not guarantee that the

---

[16]Passive traders usually roll forward contracts at the beginning of their expiration month.

[17]Inventories data were downloaded from http://www.eia.gov/. The type of traders data are available every two weeks between January 1986 and September 1992, and weekly thereafter. We construct monthly data by taking end-of month values.

[18]To implement the tests we assume that log futures prices follow an AR(2) process with seasonality. The covariance matrix is estimated using the Newey and West (1987) procedure using a Bartlett window with five lags. For the test of Busetti and Harvey (2002) we focus on the annual frequency.

estimated seasonal factors are consistent across maturities. In fact, individually extracted seasonal components display considerably heterogeneity across maturities.

6.3   The statistical relevance of stochastic seasonality

We now compare the models with stochastic (DNS-SS) and with deterministic (DNS-DS) seasonality. The former is consistent with the previous evidence on futures prices while the latter is not. We estimate both models using the two-step maximum likelihood procedure outlined in Section 5. Since U.S. bond yields do not display seasonality, the models differ only in the second step of the procedure.[19]

Table 3 reports the results of the yield curve block of the model (step 1), which is common to both representations of seasonality. Figure 2 displays the estimated factors $\hat{\delta}_{1t}$, $\hat{\delta}_{2t}$ and $\hat{\delta}_{3t}$ multiplied by 1200 (to express them in annualized percentage terms). The level, slope, and curvature factors are similar to those obtained in the literature over the same sample period (e.g. Diebold and Li, 2006). The estimate of the parameter $\hat{\theta}_1 = 0.083$ implies that the loading on the curvature factor is maximized at a maturity of about 22 months.

Tables 4 and 5 report the estimates of the futures block (step 2) of the models with stochastic and deterministic seasonality. The Akaike information criterion favors the model with stochastic seasonality (-46260 versus -42500). The Scharwz and Hannan-Quinn criteria give similar results. The estimated commodity factors in the model with stochastic seasonality are radically different from in the model with deterministic seasonality (Figure 3). The factors $\hat{\boldsymbol{\beta}}_t$ show clear signs of seasonality in the DNS-DS model but they do not in the DNS-SS model. Furthermore, the estimated feedback matrix $\widehat{\Phi}_\beta$ in the DNS-DS model contains complex roots corresponding to a cycle of roughly 12 months, which should have been eliminated if seasonality was properly extracted.

The model with stochastic seasonality also dominates the model with deterministic seasonality in terms of pricing errors (Table 6). The model with deterministic seasonality systematically produces negative pricing errors, a phenomenon particularly severe at the long end of the futures curve. In both models the pricing errors are more volatile at the short end of the futures curve and, for almost all maturities, larger in the model with deterministic seasonality. Likewise, root mean square and mean absolute pricing errors are much larger in the model with deterministic seasonality, especially at longer maturities. The poor performance of DNS-DS model reflects its inability to fully extract the seasonal fluctuations

---

[19]It is standard in the literature to write theoretical models with constant seasonal fluctuations (e.g. Sorensen, 2002) or to deseasonalize the data using dummy variables (Borovkova and Geman, 2006). In our framework, the model with deterministic seasonality amounts to setting $\sigma_\xi = \sigma_{\xi^*} = 0$ (which implies $\xi_t = \xi_0$ and $\xi_t^* = \xi_0^*$ for all t) and estimating ξ_0 and ξ_0^* as free parameters.

from the commodity factors.

We also observe the misspecification of seasonality in the top panel of Figure 5. The seasonal peak in heating oil prices is attained at the beginning of the winter season and the trough, at the beginning of the summer season. The model with stochastic seasonality shows a substantial change in the seasonal pattern over time, a feature ignored by the model with deterministic seasonality. The peak and trough months sometimes change from year to year, possibly depending on the severity and length of the winter season. There is a secular change in the seasonal pattern in the data with a clear drop in the seasonal fluctuations at the end of the sample. This observation is consistent with anecdotal evidence suggesting a vanishing seasonality in oil prices since about 2005.[20] As a result, the model with deterministic seasonality underestimates the seasonality at the beginning of the sample and over estimates it at the end of it.

Finally, the in-sample dominance of the model with stochastic seasonality also carries over to out-of-sample forecast accuracy. We construct recursive 1, 3 and 6-monts ahead forecasts for contracts of different maturities over an expanding sample starting in August 2008.[21] The h-step ahead forecast of the futures prices is calculated as

$$E_t f_{t+h}^{(\tau)} = \widehat{C}_\tau^{m_t} + \widehat{D}_\tau^{m_t\prime} E_t X_{t+h},$$

where $E_t X_{t+h}$ is computed using equation (1). The estimated parameters and factors are conditional on the information available at the time of the forecast. Table 7 summarizes the results. We conclude that, in general, the model with stochastic seasonality also dominates the deterministic one out of sample. The superior performance of the model with stochastic seasonality is apparent in the 1 and 3 months ahead forecasts of long dated contracts.

6.4 The economic relevance of stochastic seasonality

Misspecifying the seasonality process introduces large distortions into the estimated commodity factors $(\beta_{1t}, \beta_{2t}, \beta_{3t})$ which, in turn, determine the evolution of the cost-of-carry curve over time. Also, the estimated parameter θ_2 in the model with stochastic seasonality is two times larger than that in the model with deterministic seasonality. This difference is economically relevant as variations in θ_2 of such magnitude change the shape of the cost-of-carry curve and the estimated risk premia associated with the different trading strategies. We also

---

[20] The Energy Information Administration notes that "looking at data for the last 13 years, it is apparent that the traditional northern hemisphere winter spike in demand [of oil] has become increasingly less pronounced" (EIA, 2013).

[21] Although we could compute forecasts of futures of any maturity, we are constrained by the sample since longer dated contracts are available only since May 2007.

show in Figure 3 that the model with deterministic seasonality is unable to extract all seasonal patterns from the futures curve. Without properly extracting the seasonal component it is often difficult to determine whether the futures curve is normal (upward sloping) or inverted (downward sloping). Finally, consistent with the theory of storage (discussed later), we show that the moderation of the seasonal fluctuations in futures prices is correlated with a similar moderation in heating oil inventories.

The shape of the cost-of-carry curve depends on the parameter θ_2. If we specify seasonality as stochastic, we estimate $\theta_2 \approx 0.25$ while in the model with deterministic seasonality we find $\theta_2 \approx 0.11$. Differences of such magnitudes have a large impact on the loadings on the slope and curvature factors of the cost-of-carry curve (equation (28) and Figure 4). The loading on the slope factor β_{2t} decreases faster as a function of maturity when $\theta_2 = 0.25$ and the maturity that maximizes the loading on the curvature factor β_{3t} in the DNS-SS model is about half of that in the DNS-DS model (7 versus 14 months). The misspecification of the seasonal component is key to understand this difference. The factors β_{1t} and β_{2t} capture the variations of the cost-of-carry curve at the long and short end of the curve, respectively. And in the model with deterministic seasonality, the time variation in the seasonal components is erroneously attributed to fluctuations in the estimated factors $\hat{\beta}_{1t}$ and $\hat{\beta}_{2t}$. Those spurious fluctuations in the factors imply spurious fluctuations in the estimated cost-of-carry curve.

The parameter θ_2 also determines the evolution of the factors under the risk-neutral measure (equation (29) for Φ^Q). The lower is θ_2, the higher is the persistence of the level and slope factors, β_{1t} and β_{2t}, and the lower is the impact of the past curvature β_{3t} on the current slope β_{2t}. The estimated θ_2 is the model with deterministic seasonality implies more persistent factors and a weaker relation between the lagged curvature and the current slope. Moreover, through its impact on the factor loadings, θ_2 is also a key determinant of the different notions of risk premia. Therefore, the misspecification of seasonality also seriously distorts all the economically relevant measures produced by the model. In particular, the results of the DNS-DS show that the seasonality present in the $\hat{\boldsymbol{\beta}}_t$ factors also affect the dynamics of the estimated prices of risk and risk premia, which are also dominated by those misspecified seasonal fluctuations.

In Figure 6 we plot the actual, fitted, and deseasonalized futures curves on different dates using the DNS-SS model. The figure also reveals that our model is able to fit the different shapes of the yield curve observed in our sample. Seasonality is a dominant feature of the futures curve. On some dates, it is not even clear whether the futures curve is upward sloping or inverted unless one strips out the seasonal movements in prices (the broken line in the Figure displays the deseasonalized futures curve). In addition, the amplitude of the seasonal factor, as measured by the distance between the seasonal pick and trough, changes over time,

consistent with the observation that seasonality varies over time.

It is of interest to determine where does the moderation of the seasonal component over time come from. The theory of storage relates the stock of inventories with the convenience yield or the net cost of carry. We use data on the stock of inventories of heating oil to extract its seasonal component and compare it with the estimated seasonal component of our affine model.[22] The bottom panel of Figure 5 shows a striking similarity between the two series: the peaks and troughs are well aligned and the overall decline in the amplitude of the seasonal components is similar in the two series. This result suggests that the moderation in seasonal components of heating oil prices is capturing the same phenomenon in the stock of heating oil inventories.

7   Risk premia in heating oil futures markets

In this section we use the model to decompose risk premia. We perform four exercises: (i) we consider holding futures contracts of different maturities for 6 month and assess the contribution of the different factors to expected risk premia; (ii) we fix a futures contract that matures in 24 months and analyze expected risk premia across different holding periods; (iii) we decompose expected risk premia into a spot premium and a term premium; and (iv) we use the affine model to discuss the forward price unbiasedness hypothesis.

Estimated risk premia vary a lot over time and most of its variation is due to fluctuations in the commodity factors. Interest rate factors, however, also contribute to the time-variation in risk premia but mostly at medium and lower frequencies. In addition, although stochastic seasonal factors are fundamental to fit the cross section of futures prices, they play a modest (although non-negligible) role in the time variation of risk premia.

Expected risk premia is an affine function of the factors, some of which are non stationary— the level and slope yield curve factors δ_{1t} and δ_{2t}, and the commodity spot factor β_{0t} have unit roots. Yet, since estimated risk premia is stationary, there has to be a linear combination of the factors that is stationary. Our estimated long-run relation is $\delta_{1t} - 0.093\delta_{2t} + 0.0027\beta_{0t} - 0.0041$, and a vector error correction estimation suggests that the interest rate level factor δ_{1t} and commodity spot factor β_{0t} are expected to drop when the long-run relation is positive. This result is important to understand the behavior of risk premia that we discuss next.

Our two step procedure for estimating the model implies that current commodity factors

---

[22]We use a univariate unobserved component model for the logarithm of inventories specifying a local level model with seasonality (Harvey, 1989). The empirical seasonal model is the same that we used to deseasonalize futures prices in Section 6.2.

do not depend on lagged yield curve factors. Yet, this restriction does not apply under the risk-neutral measure (equation (29)). Under such representation, the yield curve factors affect the evolution of the commodity spot factor. This relation under the risk-neutral measure implies that fluctuations in interest rates do have an impact on the market prices of risk of the spot commodity factor. In addition, even though interest rate factors do not affect future commodity factors under the physical measure, the factors are not orthogonal. The residuals are correlated, and this correlation explains the long-run relation between the yield curve and spot factors. It is precisely this long-run relationship between the non-stationary factors what makes the estimated risk premia that we describe below stationary.

7.1 Expected risk premia of 6-months holding returns

We analyze the contribution of the factors to the expected 6-months holding returns for futures contracts that mature in 6 and 24 months (Figure 7).[23]

The block of commodity factors (spot and cost-of-carry) contribute the most to the evolution of risk premia. Yet, the contribution of the yield curve factors is also substantial, particularly at the beginning of the sample when interest rates are high. As interest rates drop over time, so does the contribution of the yield curve factors. Also, although the contribution of the seasonal shocks is relatively modest, it is still relevant: seasonal factors explain variations in risk premia of about 0.5 percentage points at the beginning of the sample but their contribution becomes smaller as the estimated seasonal component gets smaller over time (Figure 5).

Since the off-diagonal terms of the yield-curve block of the transition matrix $\Phi_{\delta\delta}$ are small (and mostly insignificant), we can interpret the contribution of the yield curve factors to risk premia as in the example of Section 4.3. There we argued that the loading on the level factor is negative and its importance increases with the maturity of the contract. Also, the loading on the slope factor is positive when the yield curve is upward sloping, negative when the yield curve is inverted, and the importance of this factor decreases for contracts with longer maturity. The upper right panel of Figure 7 is consistent with those observations. The contribution of the yield curve factors to the risk premium of the 24 months futures contract tracks closely the negative of the level of interest rates, while that of the 6 months contract is smoother and relatively more affected by the slope factor.

The commodity factors are very volatile and, hence, their contribution to risk is larger. Even though the block $\Phi_{\beta\beta}$ of the transition matrix is non-diagonal, we can obtain intuition

---

[23]Although we used a two-step procedure to estimate the model, we do not impose orthogonality between the residuals of the yield curve and futures block of the model. Therefore, we interpret the decompositions that follow as an approximation of the role of the different factors on expected risk premia.

from the example in Section 4.3 where we argued that the importance of the cost-of-carry level factor β_{1t} increases with the maturity of the contract. The bottom left panel of the figure shows that this result still hold in the estimated model: the level factor β_{1t} is highly negatively correlated with the contribution of the commodity factors to the expected return of the 24 months contract (-0.88) but less so with that of the 6 months contract (-0.44).

7.2   Expected holding returns of a 2-years futures contract

Figure 8 displays the expected risk premia of holding a 24 months futures contracts during 6 and 24 months. As the holding period increases, the owner of the contract is exposed to longer term risks. The spot commodity factor and the slope factor of the yield curve become more relevant risk factors. For instance, during periods of inverted yield curves, shorter term futures contracts tend to be more valuable than longer term contracts (equation (22)). Therefore, expected returns of longer term contracts increase.

The upper right panel of Figure 8 shows the greater contribution of the slope yield curve factor as the holding period increases: the long holding return tends to follow more closely the slope of the yield curve than the short holding return. A close inspection of the top panels suggests that periods with inverted yield curves are roughly associated with higher total expected risk premia of holding the futures contract. In addition, the bottom left panel shows that the contribution of the commodity factors tend to follow the medium term movements of the spot commodity factor.

7.3   Decomposing holding risk premia into spot and term premia

The expected 1-month holding return of a τ-period contract can be decomposed into a spot premium (expected return of holding the 1-month contract) and a term premium (the expected 1-month return of a τ-period contract in excess of that of a 1-month contract),

$$
\underbrace{E_t[f_{t+1}^{(\tau-1)} - f_t^{(\tau)}]}_{\text{1-month expected holding return}} = \underbrace{E_t[s_{t+1} - f_t^{(1)}]}_{\text{spot premium}} + \underbrace{E_t[(f_{t+1}^{(\tau-1)} - f_t^{(\tau)}) - (s_{t+1} - f_t^{(1)})]}_{\text{term premium}}.
$$

The top panel of Figure 9 displays the average annualized 1-month expected holding returns and its decomposition into the spot premium and term premium. Holding futures contract of different maturities for one month yield different returns. The average spot premium, of about 4.8 percentage points, is the largest component of all the 1-month expected holding returns. The average term premium attains a maximum value of 2.5 percentage points for contracts with 5 months to maturity. The term premium then decreases monotonically for longed dated contracts reaching 1 percentage points for 24-month contracts.

The spot premium, however, has changed over time. In the bottom panel of Figure 9 we split the sample before and after 2003, the year when heating oil prices started to increase considerably. The spot premium is on average about 8 percentage points before 2003 and it drops to a slightly negative value since then. The drop in the spot premium is driven by the large increase in the spot factor—the loading of the spot premium on the spot factor is negative. This effect is strong enough to compensate for the positive contribution to the spot premium of a higher cost-of-carry level factor and a lower level of interest rates. The shape of the average term premium is drastically different before and after 2003. In the early part of the sample, the average term premium is hump shaped and attains its maximum value of about 2 percentage points for contracts with 4 months to maturity. The average term premium then decreases with maturity and becomes negative for contracts longer than 9 months to maturity. In contrast, the average term premium is positive and upward sloping in the post-2003 sample reaching almost 8 percentage points for contracts with 24 months to maturity.

We now look at the spot and term premiums around the 2008-2009 recession, a period with large and sudden movements in commodity prices (Figure 10). After a sudden increase, heating oil prices reached a peak in July 2008 and the spot premium was negative and large in expectation of a decline in the commodity price (the long-run relation was positive and hence prices were expected to drop). The risk inherent in the abnormally high commodity price was reflected in a positive and increasing term premium for contracts with longer maturities: investors demand a premium to hold long contracts because they contain higher price risk than a shorter contract. At the same time, the downward sloping futures curve reflects the expectation of falling prices. As the financial crisis unfolded, heating oil prices suddenly dropped, reaching a trough in March 2009. This price level was lower than warranted by the long-run relation between the factors and, hence, prices were expected to increase (consistent with an upward sloping futures curve). As heating oil prices were expected to recover, the spot premium turned positive and the term premium negative and downward sloping as investors were willing to accept a negative premium for holding long dated contracts.

7.4 Risk premia and the unbiased forward hypothesis

When a futures contract is held to maturity, the futures price can be decomposed into the expected spot price and an expected risk premium, defined as the negative of the expected holding return of the contract (equation (21)). This equation is the basis for many empirical investigations of the unbiased forward hypothesis, which states that the price of an h-period

futures contract is the best predictor of the spot price h periods ahead,

$$f_t^{(h)} = E_t[s_{t+h}].$$

Since s_t and $f_t^{(h)}$ are usually non stationary, researchers often consider regressions of the form

$$s_{t+h} - s_t = \alpha_0 + \alpha_1(f_t^{(h)} - s_t) + e_{t+h}. \tag{34}$$

The unbiased forward hypothesis holds when $\alpha_0 = 0$ and $\alpha_1 = 1$. Rejection of the null hypothesis implies the existence of a time-varying risk premium $\pi_t^{(h)}$, as in equation (21).

The affine framework produces a risk premium $\pi_t^{(h)}$ in terms of the state variables and the parameters of the model. Therefore, besides running the simple regression (34), we also run regressions of the log-change of spot commodity price adjusted for the estimated risk premium, $s_{t+h} - h_t + \hat{\pi}_t^{(h)}$, on the futures basis, $f_t^{(h)} - s_t$,

$$s_{t+h} - s_t + \pi_t^{(h)} = \alpha_0 + \alpha_1\left(f_t^h - s_t\right) + e_{t+h}. \tag{35}$$

Table 8 reports the results of the regressions for forecast horizons from 1 through 16 months ahead, including a column with the p-value of a test of the null that $\alpha_0 = 0$ and $\alpha_1 = 1$. The top panel displays the standard regression (34). While for short forecast horizons we cannot reject the null hypothesis, as h increases the p-value of the null decreases and, for a forecast horizon of 16 months, we can reject the null hypothesis at the 10 percent confidence level. The lower panel shows the regression adding the estimated risk premium $\pi_t^{(h)}$ to the change in the spot. We now do not reject the null with great confidence, even for the long-horizon forecast of 16 months. The $p-$values are substantially larger than those of the model without the risk-premium. This is particularly relevant in the longest horizon regression ($h = 16$) and we conjecture that for longer horizons this phenomenon would be even more important.[24]

In Figure 8, the risk premium $\pi_t^{(24)}$ can be read as the negative of the expected holding return of the 24-months contract. Likewise, in Figure 7, $\pi_t^{(6)}$ is the negative of the holding return of the 6-month. In both cases, commodity futures trade on average at a discount relative to the expected spot price (this holds for all maturities τ). That is, futures prices are on average negatively biased predictors of spot prices. These risk premia, however, became smaller and less volatile since about 2007. Therefore, futures prices became better predictors of spot prices since then.

---

[24]We are not able to run regressions for horizons greater than 16 months ahead because long futures contracts are relatively new instruments and there are few observations for them. We interpret these results as lending support to the affine model.

8 Fundamentals in Heating Oil Futures markets

In this section we use the model to evaluate two traditional theories of futures prices: the theory of storage, in which variations in commodity futures reflect a convenience yield for holding inventories, storage costs, and interest rates; and the theory of normal backwardation, in which futures risk premia reflect the interaction between hedgers (commodity producers) and speculators. Hedgers use futures markets to insure against uncertain spot prices and speculators demand a compensation for bearing the price risk. We find support for the theory of storage but were unable to find convincing evidence consistent the theory of normal backwardation.

8.1 Cost-of-carry, inventories, and the theory of storage

The theory of storage and the Kaldor-Working hypothesis relate the shape of the futures curve to the level and cost of storing inventories. The convenience yield reflects the marginal benefit accrued to the holder of commodities net of their storage and insurance costs. As inventories drop, the probability of a "stock-out" increases leading to a higher convenience yield on current inventories. Thus, the theory predicts a negative relation between the convenience yield and the stock of inventories. Furthermore, this relation should be convex: when there are large costs associated with a stock-out, the convenience yield rises at an increasing rate as inventories fall towards zero (e.g. Deaton and Laroque, 1992). Equivalently, there are decreasing returns from holding the commodity in storage.

The convenience yield curve is the negative of the net cost-of-carry curve defined in equation (22). After imposing the Nelson and Siegel parametrization of the affine model, we construct the non-seasonal component of the convenience yield curve from equation (28) as a function of the level, slope, and curvature commodity factors β_{1t}, β_{2t}, β_{3t},

$$cy_t^{(\tau)} = - \left[\beta_{1t} + \left(\frac{1 - e^{-\theta_2 \tau}}{\theta_2 \tau} \right) \beta_{2t} + \left(\frac{1 - e^{-\theta_2 \tau}}{\theta_2 \tau} - e^{-\theta_2 \tau} \right) \beta_{3t} \right].$$

Note that we subtract the deterministic component $g_\tau^{m_t}$ from the convenience yield ($cy_t^{(\tau)} = -(\tilde{u}_t^{(\tau)} - g_t^{m_t})$). The level factor β_{1t} can then be interpreted as the negative of the average convenience yield, while the sum of the level and slope factors, $\beta_{1t} + \beta_{2t}$ (obtained when $\tau \to 0$) captures the (negative of) the very short end of the convenience yield curve.

A first evidence in favor of the Kaldor-Working hypothesis is displayed in Figure 11. The figure shows the evolution of the cost-of-carry level factor β_{1t} and the deseasonalized log-level of inventories (both series standardized). The correlation of 0.64 is consistent with the view that the convenience yield drops when the stock of inventories increases. This plot, however,

is unable to capture the non-linear relation between inventories and the convenience yield.

To explore the Kaldor-Working hypothesis, we estimate a non-linear regression of the net convenience yield on the stock of inventories at different maturities. Most empirical tests of the theory have focused on short dated contracts due to the unavailability of enough data on long dated contracts (e.g. Sorensen, 2002, Gorton et al., 2012). In contrast, we are able to test the Kaldor-Working hypothesis over the entire maturity spectrum by reconstructing the implied convenience yield curve (up to 24 months) using the estimated factors.

We consider an empirical model of the form

$$cy_t^{(\tau)} = \alpha + h\left(Inv_t\right) + e_t,$$

where Inv_t denotes the log of the non-seasonal component of inventories, $h\left(\cdot\right)$ is a nonlinear function, and e_t is a residual. We follow Gorton et al. (2012) and estimate the function h using cubic splines regressions.[25]

We report the results of the regressions in Table 9, where we show the slope of the non-linear regression at different levels of inventories. The column labeled "Wald" displays the p-value associated with the test of the null hypothesis that the relation is linear. The last column reports the R^2 of the regressions. Consistent with the theory of storage, the level of inventories emerges as a key driver of the convenience yield, in particular at long horizons. The relationship is clearly negative and significant at all maturities, and the explanatory power of the regression increases with the maturity of the contract. The inventory dynamics explains more than 50 percent of the variation of the convenience yield curve for long maturities. The results in Table 9 also reveal that the nonlinear relation between inventories and the convenience yield curve is stronger at short maturities. In fact, we cannot reject the null hypothesis that the relation between inventories and convenience yield is linear at long maturities. This observation is consistent with the theory, which predicts that stock levels are a key determinant of the probability of stock-out mainly in the near term.

8.2 Risk premia and the theory of storage

The theory of storage also implies a relationship between inventories and risk premia. As inventories fall, the probability of experiencing a physical stock-out increases which, in turn,

---

[25]Cubic splines are piecewise third-order polynomials that are twice continuously differentiable at any point of their domain, and that pass through a set of N prespecified "knots" that partition the domain into $N-1$ disjoint subsets. As in Gorton et al. (2013), we choose a single knot set at the average value of inventories and approximate the nonlinear function $h\left(x\right)$ as $h\left(x_t\right) \approx a_1 x_t + a_2 x_t^2 + a_3 x_t^3 + a_4 \left(x_t - \mu_x\right)^3 I_{\{x_t > \mu_x\}}$, where μ_x denotes the sample average of x_t and $I_{\{\cdot\}}$ is the indicator function that equals 1 if the term in braces is true and 0 otherwise.

is manifested into a higher volatility of future spot prices (Gorton et al., 2012). Therefore, risk-averse investors demand a compensation for the additional risk in future spot prices. That is, the theory of storage implies that the risk premium is a decreasing function of the level of inventories. Here we argue that the data is also consistent with this hypothesis.

Most empirical tests of this hypothesis have focused on short maturities and holding periods, reflecting the lack of long time series to perform the tests. In contrast, we use the affine framework to estimate the risk premia of futures contracts for maturities and holding periods of up to 24 months. As we show next, expanding the maturities and holding periods is important since the relationship between risk premia and inventories becomes stronger as the holding period increases.

Table 10 shows the results of a regression of holding returns risk premia on deseasonalized inventories. The left panel reports results for the risk premium excluding the seasonal component. The right panel displays results for a regression using as dependent variable the risk premium driven only by the commodity factors $\boldsymbol{\beta}_t$ (spot and cost-of-carry factors). The coefficients have the expected negative. In the left panel, the estimates for short holding periods are imprecisely estimated and thus not statistically significant. This result is consistent with those reported in Gorton et al. (2012). For longer dated contracts and holding periods, however, the negative relation is statistically significant. Moreover, the explanatory power of the regression increases with the holding period: inventories explain over 30 percent of the variation in expected risk premia for holding periods of 12 months or more.

Once we subtract the risk premium associated with the interest rate factors (right panel) there is an even tighter negative relationship between inventories and risk premia. Variations in inventories explain over 50 percent of the variability in the risk premia at the longer horizons. The strong relation captured by these regressions is consistent with the evidence showing a connection between the cost-of-carry and the stock of inventories discussed above. In fact, the h-period holding return (20) is a weighted average of the expected market prices of risks of the various factors. Therefore, the strong explanatory power of inventories is related to the comovement between the (deseasonalized) level of inventories and the level factor of the cost-of-carry as documented in Figure 11.

8.3   Risk premia and the theory of normal backwardation

The hypothesis that risk premia in commodity futures is related to the relative position of hedgers in the market dates back to Keynes (1930). Under the theory of normal backwardation, producers of commodities hold long positions in futures contracts to hedge the price risk of their output. In turn, speculators demand a premium to provide insurance to the

producers holding the short position of the underlying contracts. If the demand for insurance increases relative to the number of speculators, hedgers are expected to pay a higher premium to induce speculators to hold the additional risk.

We test the hypothesis of normal backwardation using data from the CFTC's Commitments of Traders Reports. Large traders are classified as "Commercials" (hedgers) and "Non-Commercials" (speculators). Small traders are labeled "Non-Reportables". We construct an empirical measure of "hedging pressure" as the ratio of the net short position collectively taken by "Commercials" to the open interest in the market (Gorton et al., 2012). Support for the theory of normal backwardation requires a positive relation between the risk premium and the measure of hedging pressure.[26]

Table 11 reports the results of the regressions of the risk premium (expected holding returns) on the measure of hedging pressure.[27] The slope of the regressions are mostly negative although in most cases not significantly different from 0. The associated R-squares are generally low, with the largest value (of approximately 4%) in the regression of holding to maturity a 1-month contract. This pattern holds for both, the overall risk premium (left panel) and the risk premium associated only with the commodity factors (right panel). The only coefficients that are significantly different from zero are those associated with short holding periods (of 1 or 3 months) at the very short end (1 month) or the very long end (24 months) of the maturity spectrum. But the estimated coefficients in these two cases have different signs: the estimated coefficient is negative for the risk premium of the 1 month contract and positive for the risk premium of the 24 months contract. The former is inconsistent with the normal backwardation hypothesis and suggests that hedgers increase their short positions as prices increase while speculator increase their long positions in a rising market (see also Gorton et al., 2012). This would make speculators appear to be momentum investors at the very short end of the curve. In contrast, the coefficient of the risk premium of holding 1 or 3 months the 24 months contract is positive and significant, in accordance with the theory. In any case, the evidence in support of the theory of normal backwardation using the proposed measure of hedging pressure is weak at best.

9 Conclusions

The price of commodity futures depends on interest rates, spot commodity prices, convenience yields, and seasonal components driven by systematic fluctuations in supply and

---

[26]We seasonally adjust the measure of hedging pressure as discussed in Section 6.2. The results, however, are qualitatively similar if we use the non-seasonally adjusted measure of hedging pressure.

[27]Using the lagged value of hedging pressure in the regression gives qualitatively similar results..

demand. In this paper we develop a multifactor affine term structure model for seasonal commodity futures and the yield curve. In the model, futures prices are driven by one spot commodity factor, three factors affecting the forward cost-of-carry curve, two factors capturing stochastic fluctuations at seasonal frequencies, and three factors that determine the evolution of the yield curve on government bonds. We argue that those nine factors are useful to match the shape and evolution of futures prices and risk premia over time. Yet, we avoid the usual identification problem of multifactor affine models by proposing a Nelson and Siegel representation of the yield curve and the cost of carry curve, and find conditions under which this representation belongs to the class of arbitrage-free affine models. Following the Nelson and Siegel tradition, we interpret the three yield curve factors and three cost-of-carry factors as level, slope, and curvature factors.

We estimate the model using data on heating oil futures prices and U.S. government bond yields over the period 1984-2012. The model is flexible enough to match the cross-section of bond prices and futures prices over time. We find strong evidence of time variation in the seasonal patterns. In addition, although the estimated level and slope factors of the yield curve, and spot commodity factor seem to contain a unit root, we find evidence of stationary time variation in risk premia—risk premia is an affine function of the factors. That is, there is a linear combination of the non-stationary factors that is stationary.

Most of the expected return of holding a futures contract reflects variations in the spot and cost-of-carry factors. The risk premium associated with holding a futures contract for several periods is correlated with medium frequency movements in the spot commodity factor. In addition, as the length of the holding period increases, so does the contribution of the level factor of the cost-of-carry curve. In contrast with usual claims in the literature, we find that yield curve factors do have a significant impact on expected holding returns, mostly at medium and lower frequencies. The decline in the level of the interest rates over our sample period is associated with a large decline in expected holding returns. Furthermore, an increase in the slope yield curve factor makes longer term contracts relatively more expensive than shorter contracts, affecting futures prices and risk premia. The contribution of seasonal shocks to risk premia is small, but non-negligible. Most importantly, we show that allowing for time-variation in the seasonal pattern is essential to obtain sensible estimates of the cost-of-carry factors and the associated risk premia.

Finally, we look at the fundamentals driving the cost-of-carry and risk premia. We find strong support for the theory of storage but little support for the theory of normal backwardation. As predicted by the theory of storage, the stock of inventories is a key determinant of the net convenience yield and expected holding returns. The relationship between the stock of inventories and the convenience yield is tighter for longer maturity contracts. In

addition, in line with the hypothesis of decreasing returns from holding the commodity in storage, we find a nonlinear relation between inventories and the net convenience yield which is stronger at shorter maturities. Also consistent with the theory, the decline in the amplitude of seasonal fluctuations observed over our sample coincides with a similar moderation of the seasonal component of stock of heating oil inventories.

References

1. Alquist R., G. H. Bauer, and A. Diez de los Rios (2013), "Macroeconomic Drivers of Crude Oil Futures Risk Premia," Manuscript, Bank of Canada.

2. Ang, A. and M. Piazzesi (2003), "A No-Arbitrage Vector Autoregression of Term Structure Dynamics with Macroeconomic and Latent Variables," Journal of Monetary Economics 50: 745-787.

3. Borovkova S. and H. Geman (2006), "Seasonal and stochastic effects in commodity forward curves" Review of Derivatives Research 9:167–186.

4. Busetti, F. and A.C. Harvey (2003), "Seasonality tests." Journal of Business and Economic Statistics 21, 420-436.

5. Canova F. and B. E. Hansen (1995). "Are seasonal patterns constant over time? A test for stochastic seasonality." Journal of Business and Economic Statistics, 13(3): 237-252.

6. Casassus J. and P. Collin-Dufresne (2005). "Stochastic convenience yield implied from commodity futures and interest rates." Journal of Finance, 60(5), 2283–2331.

7. Chin, M. and Z. Liu (2015). "A joint affine model of commodity futures and US Treasury yields," Bank of England working papers 526, Bank of England.

8. Chinn M.D. and O. Coibion (2014). "The predictive content of commodity futures." Journal of Futures Markets, 34(7): 607-636.

9. Christensen, J. H. E., F. X. Diebold, and G. D. Rudebusch (2011). "The affine arbitrage-free class of Nelson-Siegel term structure models." Journal of Econometrics 164(1): 4-20.

10. Cochrane J. H. and M. Piazzesi (2008). "Decomposing the yield curve." Manuscript, University of Chicago.

11. Cootner P. H. (1960). "Returns to speculators: Telser versus Keynes." Journal of Political Economy, 68(4): 396-404.

12. Dai Q.and K. J. Singleton "Specification Analysis of Affine Term Structure Models", The Journal of Finance Volume 55 Issue 5, pages 1943–1978, October 2000

13. de Roon F. A., T. E. Nijman, and C. Veld (2000). "Hedging pressure effects in futures markets." Journal of Finance, 55(3), 1437–1456.

14. Deaton A. and G. Laroque (1992). "On the behaviour of commodity prices." Review of Economic Studies, 59, 1–23.

15. Diebold F. X. and C. Li (2006). "Forecasting the Term Structure of Government Bond Yields." Journal of Econometrics, 130, 337–364.

16. Duffee G. R. (2002). "Term premia and interest rate forecasts in affine models." Journal of Finance, 57(1): 405-443.

17. EIA (2013). Oil Market Report. https://www.iea.org/media/omrreports/fullissues/2013-12-11.pdf

18. Fama E. and K. French (1987). "Commodity futures prices: Some evidence on forecast power, premiums and the theory of storage." Journal of Business, 60, 55–73.

19. Gibson R. and E. S. Schwartz (1990). "Stochastic convenience yield and the pricing of oil contingent claims." Journal of Finance, 45(3), 959–976.

20. Gorton G., F. Hayashi, and K. G. Rouwenhorst (2012). "The fundamentals of commodity futures returns." Review of Finance, 17(1), 1–71.

21. Hamilton J. D. and J. C. Wu (2012). "Identification and estimation of Gaussian affine term structure models." Journal of Econometrics, 168, 315–331.

22. Hamilton J. D. and J. C. Wu (2013). "Risk premia in crude oil futures markets." Journal of International Money and Finance, 42, 9–37.

23. Hansen L. P. and T. J. Sargent (2014). Recursive models of dynamic linear economies. Princeton University Press, Princeton, New Jersey.

24. Hannan E. J. (1964). "The estimation of a changing seasonal pattern." Journal of the American Statistical Association, 59:1063-1077.

25. Hilliard J. E. and J. Reis (1998). "Valuation of commodity futures and options under stochastic convenience yields, interest rates, and jump diffusions in the spot." Journal of Financial and Quantitative Analysis, 33(1), 61-86.

26. Hong Z., L. Niu, and G. Zeng (2016) "Discrete-time arbitrage-free Nelson-Siegel term structure model and application." Manuscript, Xiamen University.

27. Kaldor N. (1939). "Speculation and economic stability." The Review of Economic Studies, 7, 1–27.

28. Karstanje D., M. van der Wel, and D. van Dijk (2015). "Common factors in commodity futures curves." Manuscript, Erasmus University Rotterdam.

29. Keynes J. (1930). A Treatise on Money, vol. 2. Macmillan, London.

30. Litterman R. and J. Scheinkman (1991). "Common factors affecting bond returns." The Journal of Fixed Income, 1(1): 54-61.

31. Litzenberger R. H. and N. Rabinowitz (1995), "Backwardation in Oil Futures Markets: Theory and Empirical Evidence", The Journal of Finance Volume 50, Issue 5, pages 1517–1545, December 1995

32. Miltersen K. R. and E. S. Schwartz (1998), "Pricing options of Commodities Futures with Stochastic Terms Structures of Convenience Yields and Interest Rates" The Journal of Financial and Quantitative Analysis, 33(1), 33-59.

33. Nelson C. and A. Siegel (1987). "Parsimonious modeling of yield curves." Journal of Business, 60, 473–489.

34. NYMEX (2009), New York Mercantile Exchange. Chapter 150: NY Harbor ULSD Futures. In NYMEX Rulebook. http://www.cmegroup.com/rulebook/NYMEX/

35. Routledge B. R., D. J. Seppi, and C. S. Spatt (2000). "Equilibrium forward curves for commodities." Journal of Finance, 55(3): 1297-1338.

36. Schwartz E. S. (1997). "The stochastic behavior of commodity prices: Implications for valuation and hedging." Journal of Finance, 52(3), 923–973.

37. Sorensen C. (2002). "Modeling seasonality in agricultural commodity futures." Journal of Futures Markets, 22(5), 393–426.

38. Trolle A. B. and E. S. Schwartz (2011)."Unspanned stochastic volatility and the pricing of commodity derivatives." The Review of Financial Studies, 22(11): 4423-4461.

39. Szymanowska M., F. de Roon, T. Nijman, and R. van den Goorbergh (2014). "An anatomy of commodity futures risk premia." Journal of Finance, 69(1), 453–482.

40. Working H. (1949). "The theory of price storage." American Economic Review, 39, 1254-1252.

Appendix A: Proofs

Proof of Proposition 1. The spot commodity price satisfies the pricing condition

$$S_t = E_t \left[M_{t,t+1} e^{-c_{t+1}} S_{t+1} \right]. \tag{A.1}$$

We guess and verify values for $\psi_0^{\tilde{m}}$ and $\psi_1^{\tilde{m}}$ that generate the price process (8). Let $\tilde{m} = m_t$ and use the convention $\tilde{m} + 1 = 1$ for $\tilde{m} = 12$. Then, using equations (1), (2), (8), and (10) into (A.1) gives

$$\exp\left(\gamma_0 + \gamma_1^{\tilde{m}\prime} X_t\right) = \exp\left(-\rho_0 - \rho_1' X_t + \gamma_0 - \psi_0^{\tilde{m}+1} + \left(\gamma_1^{\tilde{m}+1} - \psi_1^{\tilde{m}+1}\right)' (\mu + \Phi X_t) - \tfrac{1}{2}\Lambda_t'\Lambda_t\right)$$
$$\times E_t \left[\exp\left(\left[\left(\gamma_1^{\tilde{m}+1} - \psi_1^{\tilde{m}+1}\right)'\Gamma - \Lambda_t'\right]\eta_{t+1}\right)\right].$$

Solving the expectation,

$$E_t\left[\exp\left(\left[\Gamma'\left(\gamma_1^{\tilde{m}+1} - \psi_1^{\tilde{m}+1}\right) - \Lambda_t\right]'\eta_{t+1}\right)\right]$$
$$= \exp\left(\left(\gamma_1^{\tilde{m}+1} - \psi_1^{\tilde{m}+1}\right)'\tfrac{\Gamma\Gamma'}{2}\left(\gamma_1^{\tilde{m}+1} - \psi_1^{\tilde{m}+1}\right) - \left(\gamma_1^{\tilde{m}+1} - \psi_1^{\tilde{m}+1}\right)'\Gamma\Lambda_t + \tfrac{1}{2}\Lambda_t'\Lambda_t\right).$$

Replacing this expression above, using $\Lambda_t = \lambda_0 + \lambda_1 X_t$, and rearranging gives

$$\gamma_1^{\tilde{m}\prime} X_t = -\rho_0 - \psi_0^{\tilde{m}+1} + \left(\gamma_1^{\tilde{m}+1} - \psi_1^{\tilde{m}+1}\right)'(\mu - \Gamma\lambda_0)$$
$$+ \left(\gamma_1^{\tilde{m}+1} - \psi_1^{\tilde{m}+1}\right)'\tfrac{\Gamma\Gamma'}{2}\left(\gamma_1^{\tilde{m}+1} - \psi_1^{\tilde{m}+1}\right) + \left[\left(\gamma_1^{\tilde{m}+1} - \psi_1^{\tilde{m}+1}\right)'(\Phi - \Gamma\lambda_1) - \rho_1'\right]X_t.$$

Matching coefficients gives the values of $\psi_0^{\tilde{m}+1}$ and $\psi_1^{\tilde{m}+1}$ displayed in Proposition 1.□

Pricing commodity futures

Rewrite the pricing condition of the futures contract that matures in τ periods as

$$F_t^{(\tau)} E_t \left[M_{t,t+\tau}\right] = E_t \left[M_{t,t+\tau} S_{t+\tau}\right]. \tag{A.2}$$

The price of a contract written at time $t + 1$ with settlement at date $t + \tau$ is thus

$$F_{t+1}^{(\tau-1)} E_{t+1} \left[M_{t+1,t+\tau}\right] = E_{t+1} \left[M_{t+1,t+\tau} S_{t+\tau}\right]$$

Multiply both sides of this expression by $M_{t,t+1}$, use $M_{t,t+1} M_{t+1,t+\tau} = M_{t,t+\tau}$, and take expectations conditional on information at time t to obtain

$$E_t \left[M_{t,t+\tau} S_{t+\tau}\right] = E_t \left[M_{t,t+1} F_{t+1}^{(\tau-1)} E_{t+1} \left[M_{t+1,t+\tau}\right]\right].$$

Equation (11) follows by using the previous equation into (A.2) and noting that $E_t \left[M_{t,t+\tau}\right] = P_t^{(\tau)}$ and $E_t \left[M_{t+1,t+\tau}\right] = P_{t+1}^{(\tau-1)}$. Now let $V_t^{(\tau)} = F_t^{(\tau)} P_t^{(\tau)}$ and write (11) as

$$V_t^{(\tau)} = E_t[M_{t,t+1} V_{t+1}^{(\tau-1)}]. \tag{A.3}$$

We conjecture a solution of the form

$$\log V_t^{(\tau)} = G_\tau^{m_t} + H_\tau^{m_t\prime} X_t,$$

Using the proposed solution, introducing equations (1) and (2) into (A.3), and evaluating the conditional expectation gives

$$
\begin{aligned}
G_\tau^{m_t} + H_\tau^{m_t\prime} X_t &= G_{\tau-1}^{m_{t+1}} + H_{\tau-1}^{m_{t+1}\prime}(\mu - \Gamma\lambda_0) + \tfrac{1}{2} H_{\tau-1}^{m_{t+1}\prime} \Gamma\Gamma' H_{\tau-1}^{m_{t+1}} - \rho_0 \\
&\quad + \left(H_{\tau-1}^{m_{t+1}\prime}(\Phi - \Gamma\lambda_1) - \rho_1' \right) X_t.
\end{aligned}
$$

Fix an arbitrary month $\tilde{m}$ and match coefficients to obtain

$$
\begin{aligned}
G_\tau^{\tilde{m}} &= G_{\tau-1}^{\tilde{m}+1} + H_{\tau-1}^{\tilde{m}+1\prime}(\mu - \Gamma\lambda_0) + \tfrac{1}{2} H_{\tau-1}^{\tilde{m}+1\prime} \Gamma\Gamma' H_{\tau-1}^{\tilde{m}+1} - \rho_0 \\
H_\tau^{\tilde{m}} &= (\Phi - \Gamma\lambda_1)' H_{\tau-1}^{\tilde{m}+1} - \rho_1,
\end{aligned}
$$

where $\tilde{m} + 1 = 1$ when $\tilde{m} = 12$. Also, note that $V_t^{(0)} = P_t^{(0)} F_t^{(0)}$, $P_t^{(0)} = 1$, and $F_t^{(0)} = S_t$. Therefore, the conjecture and (8) imply

$$G_0^{m_t} + H_0^{m_t\prime} X_t = \gamma_0 + \gamma_1^{m_t\prime} X_t.$$

Therefore, $G_0^{\tilde{m}} = \gamma_0$ and $H_0^{\tilde{m}} = \gamma_1^{\tilde{m}}$ for $\tilde{m} = 1, 2, ..., 12$. Finally, note that

$$\log F_t^{(\tau)} = \log(V_t^{(\tau)}/P_t^{(\tau)}) = G_\tau^{m_t} - A_\tau + (H_\tau^{m_t} - B_\tau)' X_t.$$

Equation (12) follows by setting $C_\tau^{m_t} = G_\tau^{m_t} - A_\tau$ and $D_t^{m_t} = H_\tau^{m_t} - B_\tau$.

Proof of Proposition 2. Given the proposed μ^Q and Φ^Q, define the following parameters of the affine model

$$\lambda_0 = \Gamma^{-1}\left(\mu - \mu^Q\right), \ \lambda_1 = \Gamma^{-1}\left(\Phi - \Phi^Q\right),$$

$$\rho_0 = 0, \ \rho_1 = \left[1, \tfrac{1-e^{-\theta_1}}{\theta_1}, \tfrac{1-e^{-\theta_1}}{\theta_1} - e^{-\theta_1}, 0, 0, 0, 0, 0, 0 \right]',$$

and, for $\tilde{m} = 1, 2, ..., 12$,

$$\gamma_0 = 0, \ \gamma_1^{\tilde{m}} = \left[0, 0, 0, 1, 0, 0, 0, \cos(\tfrac{2\pi}{12}\tilde{m}), \sin(\tfrac{2\pi}{12}\tilde{m}) \right]$$

Log bond prices and log futures prices are affine functions of the states,

$$
\begin{aligned}
p_t^{(\tau)} &= A_\tau + B_\tau' X_t, \\
f_t^{(\tau)} &= C_\tau^{\tilde{m}} + D_\tau^{\tilde{m}\prime} X_t.
\end{aligned}
$$

Define also

$$
\begin{aligned}
C_\tau^{\tilde{m}} &= G_\tau^{\tilde{m}} - A_\tau, \\
D_\tau^{\tilde{m}} &= H_\tau^{\tilde{m}} - B_\tau.
\end{aligned}
\tag{A.4}
$$

The pricing of bonds and futures imply that the loadings B_τ and $H_\tau^{\tilde{m}}$ satisfy

$$
\begin{aligned}
B_\tau &= \Phi^{Q\prime} B_{\tau-1} - \rho_1 & \text{(A.5)} \\
H_\tau^{\tilde{m}} &= \Phi^{Q\prime} H_{\tau-1}^{\tilde{m}+1} - \rho_1 & \text{(A.6)}
\end{aligned}
$$

with initial conditions $B_0 = \mathbf{0}$ and $H_0^{\tilde{m}} = \gamma_1^{\tilde{m}}$ for $\tilde{m} = 1, 2, ..., 12$.

Partition the risk-neutral feedback matrix, Φ^Q, as follows

$$
\Phi^Q = \begin{bmatrix}
\Phi_{\delta\delta}^Q & \Phi_{\delta\beta}^Q & \Phi_{\delta\xi}^Q \\
\Phi_{\beta\delta}^Q & \Phi_{\beta\beta}^Q & \Phi_{\beta\xi}^Q \\
\Phi_{\xi\delta}^Q & \Phi_{\xi\beta}^Q & \Phi_{\xi\xi}^Q
\end{bmatrix}
$$

where the sizes of the sub-matrices conforms to the size of the vectors $\boldsymbol{\delta}_t$, $\boldsymbol{\beta}_t$ and $\boldsymbol{\xi}_t$.

We begin with the recursion for bonds. The Nelson and Siegel parametrization implies that bond prices depend only on $\boldsymbol{\beta}_t$ (i.e. $B_\tau = [-\tau, -\frac{1-e^{-\theta_1\tau}}{\theta_1}, -(\frac{1-e^{-\theta_1\tau}}{\theta_1} - \tau e^{-\theta_1\tau}), 0_{1\times 6}]'$). Therefore, the non-zero elements of (A.5) requires

$$
\begin{bmatrix}
-\tau \\
-\frac{1-e^{-\theta_1\tau}}{\theta_1} \\
-\left(\frac{1-e^{-\theta_1\tau}}{\theta_1} - \tau e^{-\theta_1\tau}\right)
\end{bmatrix}
= \left(\Phi_{\delta\delta}^Q\right)'
\begin{bmatrix}
-(\tau-1) \\
-\frac{1-e^{-\theta_1(\tau-1)}}{\theta_1} \\
-\left(\frac{1-e^{-\theta_1(\tau-1)}}{\theta_1} - \tau e^{-\theta_1(\tau-1)}\right)
\end{bmatrix}
-
\begin{bmatrix}
1 \\
\frac{1-e^{-\theta_1}}{\theta_1} \\
\left(\frac{1-e^{-\theta_1}}{\theta_1} - e^{-\theta_1}\right)
\end{bmatrix},
$$

which in turns implies

$$
\Phi_{\delta\delta}^Q = \begin{bmatrix}
1 & 0 & 0 \\
0 & e^{-\theta_1} & \theta_1 e^{-\theta_1} \\
0 & 0 & e^{-\theta_1}
\end{bmatrix}.
$$

This condition is the usual one for arbitrage-free DNS models (see e.g. Hong, Niu, and Zeng, 2016).

Consider next the recursions of the parameters associated with the commodity factors. The Nelson and Siegel structure implies that the first 3 elements of $H_\tau^{\tilde{m}}$ are zero. Thus, satisfying (A.6) requires

$$
\begin{bmatrix}
1 \\
\tau \\
\frac{1-e^{-\theta_2\tau}}{\theta_2} \\
\left(\frac{1-e^{-\theta_2\tau}}{\theta_2} - \tau e^{-\theta_2\tau}\right) \\
e^{-\omega\tau}\cos\left(\frac{2\pi}{12}(\tilde{m}+\tau)\right) \\
e^{-\omega\tau}\sin\left(\frac{2\pi}{12}(\tilde{m}+\tau)\right)
\end{bmatrix}
=
\begin{bmatrix}
\Phi_{\beta\beta}^Q & \Phi_{\beta\xi}^Q \\
\Phi_{\xi\beta}^Q & \Phi_{\xi\xi}^Q
\end{bmatrix}'
\begin{bmatrix}
1 \\
(\tau-1) \\
\frac{1-e^{-\theta_2(\tau-1)}}{\theta_2} \\
\left(\frac{1-e^{-\theta_2(\tau-1)}}{\theta_2} - (\tau-1)e^{-\theta_2(\tau-1)}\right) \\
e^{-\omega(\tau-1)}\cos\left(\frac{2\pi}{12}(\tilde{m}+1+\tau-1)\right) \\
e^{-\omega(\tau-1)}\sin\left(\frac{2\pi}{12}(\tilde{m}+1+\tau-1)\right)
\end{bmatrix}
\quad \text{(A.7)}
$$

and

$$\begin{bmatrix} 1 \\ \frac{1-e^{-\theta_1}}{\theta_1} \\ \frac{1-e^{-\theta_1}}{\theta_1} - e^{-\theta_1} \end{bmatrix} = \begin{bmatrix} \left(\Phi_{\beta\delta}^Q\right)' & \left(\Phi_{\xi\delta}^Q\right)' \end{bmatrix} \begin{bmatrix} 1 \\ (\tau-1) \\ \frac{1-e^{-\theta_2(\tau-1)}}{\theta_2} \\ \left(\frac{1-e^{-\theta_2(\tau-1)}}{\theta_2} - (\tau-1)e^{-\theta_2(\tau-1)}\right) \\ e^{-\omega(\tau-1)}\cos\left(\frac{2\pi}{12}(\tilde{m}+\tau)\right) \\ e^{-\omega(\tau-1)}\sin\left(\frac{2\pi}{12}(\tilde{m}+\tau)\right) \end{bmatrix}. \quad (A.8)$$

It is easy to verify that (A.7) implies $\Phi_{\xi\beta}^Q = 0_{2\times4}$, $\Phi_{\beta\xi}^Q = 0_{4\times2}$,

$$\Phi_{\xi\xi}^Q = \begin{bmatrix} e^{-\omega} & 0 \\ 0 & e^{-\omega} \end{bmatrix}, \text{ and}$$

$$\Phi_{\beta\beta}^Q = \begin{bmatrix} 1 & 1 & \frac{1-e^{-\theta_2}}{\theta_2} & \frac{1-e^{-\theta_2}}{\theta_2} - e^{-\theta_2} \\ 0 & 1 & 0 & 0 \\ 0 & 0 & e^{-\theta_2} & \theta_2 e^{-\theta_2} \\ 0 & 0 & 0 & e^{-\theta_2} \end{bmatrix},$$

whereas (A.8) requires $\Phi_{\xi\delta}^Q = 0_{2\times3}$ and

$$\Phi_{\beta\delta}^Q = \begin{bmatrix} 1 & \frac{1-e^{-\theta_1}}{\theta_1} & \left(\frac{1-e^{-\theta_1}}{\theta_1} - e^{-\theta_1}\right) \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Lastly, (A.5) and (A.6) together with (A.4) imply $\Phi_{\delta\beta}^Q = 0_{3\times4}$ and $\Phi_{\delta\xi}^Q = 0_{3\times2}$.$\square$

Appendix B: Trading strategies and risk premia

The (realized) 1-period *holding return* of a τ-month contract is

$$f_{t+1}^{(\tau-1)} - f_t^{(\tau)} = C_{\tau-1}^{m_{t+1}} + D_{\tau-1}^{m_{t+1}'}X_{t+1} - C_\tau^{m_t} - D_\tau^{m_t'}X_t.$$

Using (1), (5), (6), (13), (14), (15), and (16) we can write the holding return as

$$f_{t+1}^{(\tau-1)} - f_t^{(\tau)} = \frac{1}{2}[B_{\tau-1}'\Gamma\Gamma'B_{\tau-1} - H_{\tau-1}^{m_{t+1}'}\Gamma\Gamma'H_{\tau-1}^{m_{t+1}}] + D_{\tau-1}^{m_{t+1}'}\Gamma\Lambda_t + D_{\tau-1}^{m_{t+1}'}\Gamma\eta_{t+1}.$$

The time$-t$ expected 1-period holding return is thus

$$E_t[f_{t+1}^{(\tau-1)} - f_t^{(\tau)}] = \frac{1}{2}[B_{\tau-1}'\Gamma\Gamma'B_{\tau-1} - H_{\tau-1}^{m_{t+1}'}\Gamma\Gamma'H_{\tau-1}^{m_{t+1}}] + D_{\tau-1}^{m_{t+1}'}\Gamma\Lambda_t.$$

The *spot premium* is the 1-period holding return of a futures contract with settlement

date in the following period. Since $f_{t+1}^{(0)} = s_{t+1}$, the spot premium is

$$E_t[s_{t+1} - f_t^{(1)}] = \frac{1}{2}[B_0'\Gamma\Gamma'B_0 - H_0^{m_{t+1}'}\Gamma\Gamma'H_0^{m_{t+1}}] + D_0^{m_{t+1}'}\Gamma\Lambda_t$$

But using $B_0 = 0$ and $H_0^{m_{t+1}} = D_0^{m_{t+1}} = \gamma_1^{m_{t+1}}$ gives

$$E_t[s_{t+1} - f_t^{(1)}] = -\frac{1}{2}\gamma_1^{m_{t+1}'}\Gamma\Gamma'\gamma_1^{m_{t+1}} + \gamma_1^{m_{t+1}'}\Gamma\Lambda_t.$$

The *term premium, h-period holding returns, short roll, excess short roll,* and *spreading strategies* follow from simple manipulations of the 1-period holding returns and the spot premium as described in the text.

Table 1
Test of stochastic versus deterministic seasonality in futures prices

τ stands for maturity. CH denotes Canova and Hansen (1995) test while BH denotes Busetti and Harvey (2003) parametric and nonparametric tests. Under the null hypothesis of deterministic seasonality, the three test statistics are distributed as a generalized Von-Mises random variable with 2 degrees of freedom. ****** denotes the significance at 1/5/10% level.

	$\tau = 1$	$\tau = 2$	$\tau = 3$	$\tau = 6$	$\tau = 9$	$\tau = 12$
CH	0.8894**	0.7781**	0.6819*	0.3773	0.8364**	1.0029**
BH (Nonparam)	1.0975***	1.2511***	1.2390***	0.2286	0.2282	0.4505
BH (Param)	3.7307***	6.0447***	11.2695***	11.0700***	2.5658***	7.5472***

Table 2
Selecting the number of seasonal factors

The stochastic seasonal component is driven by pairs of seasonal factors:

$$z_t^s = \sum_{j=1}^{6} [\xi_{jt} \cos\left(\frac{2\pi}{12}j\right) + \xi_{jt}^* \sin\left(\frac{2\pi}{12}j\right)],$$

where $\xi_{jt} = \xi_{jt-1} + \nu_{jt}$, $\xi_{jt}^* = \xi_{jt-1}^* + \nu_{jt}^*$ for $j = 1, 2..., 6$. The shocks ν_{jt} and ν_{jt}^* are independent and normally distributed with mean zero and variances σ_j^2 and σ_j^{*2}. Each column of the table corresponds to a different specification of the seasonal component. j indicates the number of pairs of seasonal factors (i.e. the harmonics) included. τ stands for maturity. For each specification we report the log likelihood value, the AIC, and BIC. Bold numbers denote the minimum values on the information criteria.

	$j = 1, \sigma_1^2 \neq \sigma_1^{*2}$	$j = 1, \sigma_1^2 = \sigma_1^{*2}$	$j = 2, \sigma_j^2 = \sigma_j^{*2} \, \forall j$	$j = 6, \sigma_j^2 = \sigma_j^{*2} \, \forall j$
$\tau = 1$				
LogLik	-986.4754	-986.6895	-994.3225	-1027.9003
AIC	**1988.9507**	1989.3790	2008.6449	2097.8007
BIC	**2019.6759**	2020.1042	2047.0514	2178.4541
$\tau = 2$				
LogLik	-965.5135	-965.8760	-970.8954	-1006.1485
AIC	**1947.0270**	1947.7521	1961.7908	2054.2971
BIC	**1977.7522**	1978.4772	2000.1972	2134.9506
$\tau = 3$				
LogLik	-949.1721	-949.1967	-951.5601	-986.9808
AIC	**1914.3442**	1914.3934	1923.1202	2015.9616
BIC	**1945.0693**	1945.1185	1961.5266	2096.6150
$\tau = 6$				
LogLik	-890.0019	-890.0986	-898.4592	-935.5490
AIC	**1796.0038**	1796.1972	1816.9185	1913.0980
BIC	**1826.7290**	1826.9223	1855.3249	1993.7515
M9				
LogLik	-841.7108	-842.7229	-846.0776	-881.3148
AIC	**1699.4216**	1701.4458	1712.1552	1804.6297
BIC	**1730.1467**	1732.1709	1750.5616	1885.2831
$\tau = 12$				
LogLik	-681.4070	-681.4403	-691.0648	-728.8841
AIC	**1378.8139**	1378.8806	1402.1296	1499.7682
BIC	**1409.5391**	1409.6057	1440.5360	1580.4216

Table 3
Estimates of the yield curve block of the model

The table contains the estimates of the yield curve block of the affine model of futures prices using the Nelson and Siegel parametrization (first step of the 2-step procedure). The parameter θ_1 determines the shape of the factor loadings and μ_1^Q is the free parameter associated with the non-arbitrage restrictions. The parameters of the vector μ_δ, the matrix $\Gamma_{\delta\delta}$, and μ_1^Q are multiplied by 1000 for easier reading.

Parameters of the VAR(1) process for the yield curve factors

$\mu_\delta\,(\times 1000)$ $\qquad\qquad\qquad$ $\Phi_{\delta\delta}$ $\qquad\qquad\qquad\qquad\qquad$ $\Gamma_{\delta\delta}\,(\times 1000)$

$$
\begin{bmatrix} 0.0201 \\ (0.0371) \\ 0.1046 \\ (0.0457) \\ -0.0149 \\ (0.0714) \end{bmatrix}
\quad
\begin{bmatrix} 0.999 & 0.026 & 0.002 \\ (0.008) & (0.015) & (0.013) \\ -0.037 & 0.901 & 0.080 \\ (0.010) & (0.018) & (0.015) \\ 0.005 & 0.003 & 0.961 \\ (0.016) & (0.031) & (0.025) \end{bmatrix}
\quad
\begin{bmatrix} 0.270 & 0 & 0 \\ (0.015) & & \\ -0.240 & 0.201 & 0 \\ (0.018) & (0.009) & \\ 0.111 & 0.021 & 0.507 \\ (0.044) & (0.041) & (0.041) \end{bmatrix}
$$

Other parameters and Log-likelihood

$\theta_1 = 0.067\,(0.003)$ $\qquad\qquad$ $\mu_1^Q\,(\times 1000) = 0.0160\,(0.0026)$ $\qquad\qquad$ Log-likelihood $= 18847.24$

Table 4
Estimates of the futures block of the model: stochastic seasonality

The table contains the estimates of the futures curve block of the model assuming stochastic seasonality and using the Nelson and Siegel parametrization (second step of the 2-step procedure). The parameter θ_2 determines the shape of the factor loadings, ω is the discounting factor on the seasonal shocks, and μ_8^Q and μ_9^Q are the parameters associated with the non-arbitrage restrictions.

Parameters of the VAR(1) process for the cost-of-carry factors $\boldsymbol{\beta}_t$

μ_β $(\times 100)$ $\qquad\qquad$ $\Phi_{\beta\beta}$ $\qquad\qquad$ $\Gamma_{\beta\beta}$ $(\times 100)$

$$
\begin{bmatrix}
1.670 \\
(1.275) \\
-0.148 \\
(0.042) \\
-0.052 \\
(0.037) \\
0.075 \\
(0.036)
\end{bmatrix}
\begin{bmatrix}
0.992 & 2.270 & 0.725 & -0.224 \\
(0.012) & (1.508) & (0.294) & (0.259) \\
0.0002 & 0.702 & 0.057 & 0.045 \\
(0.0004) & (0.049) & (0.012) & (0.009) \\
-0.0002 & -1.122 & 0.983 & 0.362 \\
(0.0032) & (0.451) & (0.064) & (0.070) \\
0.002 & 1.660 & -0.348 & 0.479 \\
(0.003) & (0.424) & (0.083) & (0.075)
\end{bmatrix}
\begin{bmatrix}
9.655 & 0 & 0 & 0 \\
(0.400) & & & \\
-0.120 & 0.249 & 0 & 0 \\
(0.023) & (0.035) & & \\
-0.807 & 0.877 & 1.580 & 0 \\
(0.129) & (0.205) & (0.145) & \\
0.027 & -1.761 & -1.136 & 0.773 \\
(0.002) & (0.002) & (0.140) & (0.058)
\end{bmatrix}
$$

Volatility of seasonal process and parameters of the risk-neutral intercept
$\sigma_\xi = 0.0036 \, (0.0002)$ $\quad$ $\sigma_{\xi^*} = 0.0018 \, (0.0002)$ $\quad$ $\mu_8^Q = \mu_9^Q \, (\times 1000) = 0.125 \, (0.055)$

Other parameters and Log-likelihood
$\theta_2 = 0.2531 \, (0.0145)$ $\quad$ $\omega = 0.0082 \, (0.0028)$ $\quad$ Log-likelihood $= 23189.20$

Table 5
Estimates of the futures block of the model: deterministic seasonality

The table contains the estimates of the futures curve block of the model assuming deterministic seasonality and using the Nelson and Siegel parametrization (second step of the 2-step procedure). The parameter θ_2 determines the shape of the factor loadings, ω is the discounting factor on the seasonal shocks, and μ_5 is the free are the parameters associated with the non-arbitrage restrictions restrictions.

Parameters of the VAR(1) process for the cost-of-carry factors $\boldsymbol{\beta}_t$

$\mu_\beta\,(\times 100)$

$$
\begin{bmatrix}
-0.444 \\
(1.360) \\
0.868 \\
(0.867) \\
-0.031 \\
(0.037) \\
-0.232 \\
(0.212)
\end{bmatrix}
$$

$\Phi_{\beta\beta}$

$$
\begin{bmatrix}
0.997 & 1.695 & 0.323 & 0.228 \\
(0.012) & (0.957) & (0.299) & (0.358) \\
-0.0018 & -0.203 & 0.219 & -0.446 \\
(0.0032) & (0.230) & (0.073) & (0.102) \\
0.0004 & 0.605 & 0.862 & 0.331 \\
(0.0016) & (0.177) & (0.042) & (0.069) \\
0.007 & 2.628 & -0.622 & 1.857 \\
(0.008) & (0.565) & (0.199) & (0.238)
\end{bmatrix}
$$

$\Gamma_{\beta\beta}\,(\times 100)$

$$
\begin{bmatrix}
9.882 & 0 & 0 & 0 \\
(0.435) & & & \\
-0.283 & 1.724 & 0 & 0 \\
(0.156) & (0.231) & & \\
-0.666 & -0.634 & 0.840 & 0 \\
(0.102) & (0.225) & (0.059) & \\
0.883 & -4.375 & -0.891 & 0.537 \\
(0.004) & (0.005) & (0.075) & (0.028)
\end{bmatrix}
$$

Parameters of seasonal fluctuations and and parameters of the risk-neutral intercept

$\xi_0 = 0.0044\,(0.0001)$ $\xi_0^* = 0.0037\,(0.0002)$ $\mu_5^Q\,(\times 1000) = -0.576\,(0.263)$

Other parameters and Log-likelihood

$\theta_2 = 0.1149\,(0.0063)$ $\omega = 0.0734\,(0.0032)$ Log-likelihood $= 21309.21$

Table 6
Pricing errors

The table reports various measures of pricing errors in the models with stochastic (SS) and deterministic (DS) seasonality. The column "Bias" reports the average pricing error while "St. Dev." their standard deviations. RMSPE denotes the root mean square pricing error while MAPE is the mean absolute pricing error. All entries are multiplied by 100.

	Bias		St. Dev.		RMSPE		MAPE	
Maturity	SS	DS	SS	DS	SS	DS	SS	DS
1	-0.0056	-0.1054	1.0443	2.7240	1.0428	2.7221	0.7201	2.0395
3	0.0115	-0.0827	0.3283	0.2720	0.3280	0.2839	0.2398	0.2206
6	-0.0140	-0.2045	0.4812	0.6304	0.4807	0.6618	0.3772	0.5229
9	0.0024	-0.2649	0.4394	0.7820	0.4388	0.8245	0.3504	0.6654
12	-0.0057	-0.3961	0.5248	0.6875	0.5239	0.7924	0.4022	0.6193
17	0.0067	-0.4939	0.4176	0.9821	0.4166	1.0971	0.3370	0.8907

Table 7
Out-of-sample forecasting results

The table contains the results of out-of-sample forecasting using the models stochastic and deterministic seasonality for different horizons. We estimate both models recursively from 1984:1 to 2008:8, the period when the first forecast is made, through 2012:8. We consider forecast errors of the log of futures prices, and we report the mean, standard deviation, and root mean squared errors of the forecast errors.

Maturity (τ)	Stochastic seasonality			Deterministic seasonality		
	Mean	Std. Dev.	RMSE	Mean	Std. Dev.	RMSE
1-month ahead forecasts						
2 months	0.001	0.098	0.097	0.007	0.101	0.100
6 months	0.007	0.093	0.092	0.013	0.092	0.092
18 months	−0.001	0.085	0.084	−0.003	0.088	0.086
24 months	−0.004	0.083	0.082	−0.018	0.134	0.133
3-months ahead forecasts						
2 months	0.007	0.228	0.225	0.015	0.232	0.230
6 months	0.014	0.217	0.215	0.024	0.211	0.210
18 months	−0.005	0.216	0.213	−0.012	0.223	0.220
24 months	0.007	0.217	0.213	−0.041	0.292	0.290
12-months ahead forecasts						
2 months	−0.090	0.331	0.339	−0.075	0.325	0.329
6 months	−0.075	0.325	0.330	−0.065	0.319	0.321
18 months	−0.093	0.347	0.352	−0.088	0.345	0.349
24 months	−0.052	0.377	0.370	−0.033	0.363	0.355

Table 8
The predictive content of futures prices

The table reports the results of forecasting regressions of the log change in the spot $s_{t+h} - s_t$ on the the futures basis $f_t^{(h)} - s_t$ for different forecast horizons h. The top panel consider the conventional prediction regression, while the bottom panel adjusts the change in the spot by the risk premium term $\pi_t^{(h)}$ estimated from the model.

A. Conventional predictive regression

$$s_{t+h} - s_t = \alpha_0 + \alpha_1(f_t^h - s_t) + e_{t+h}$$

h	α_0	s.e.	α_1	s.e.	$p - value$ $[H_0:\alpha_0=0;\ \alpha_1=1]$	R^2
1	0.0051	0.0057	1.1143	0.2463	0.6290	0.0865
3	0.0167	0.0161	1.1102	0.3047	0.5673	0.1426
6	0.0349	0.0291	0.9894	0.3018	0.4836	0.1247
12	0.0874	0.0502	0.8677	0.3454	0.2148	0.0867
16	0.1364	0.0711	1.0263	0.3895	0.0942	0.1480

B. Model consistent predictive regressions

$$s_{t+h} - s_t + \pi_t^h = \alpha_0 + \alpha_1(f_t^h - s_t) + e_{t+h}$$

h	α_0	s.e.	α_1	s.e.	$p - value$ $[H_0:\alpha_0=0;\ \alpha_1=1]$	R^2
1	0.0010	0.0056	1.0806	0.2324	0.9316	0.0813
3	0.0013	0.0162	1.0289	0.2972	0.9930	0.1262
6	0.0016	0.0292	0.9673	0.2967	0.9919	0.1212
12	0.0251	0.0503	1.0820	0.3385	0.8520	0.1300
16	0.0587	0.0690	1.2645	0.3747	0.3527	0.2158

Table 9
Convenience Yields and Inventories

The table reports the a cubic splines regression of the (seasonally adjusted) convenience yield on (seasonally adjusted log) inventories. The regression takes the form

$$cy_t^{(\tau)} = \alpha + a_1 Inv_t + a_2 Inv_t^2 + a_3 Inv_t^3 + a_4 \left(Inv_t - \overline{Inv} \right)^3 I_{\left\{ Inv_t > \overline{Inv} \right\}}$$

where $cy_t^{(\tau)}$ is the convenience yield, Inv_t denotes the log of the deseasonalized stock of inventories, $\overline{Inv}$ is the sample average of the inventories, and $I_{\left\{ Inv_t > \overline{Inv} \right\}}$ is an indicator function that equals 1 if the level of inventories at time t is larger than the sample average and zero otherwise. Columns 2 to 7 report the slope and its associated t-statistics at different values of the inventories (the 10th percentile, the mean, and the 90th percentile). For example, the reported slope at the 10th percentile is

$$slope_{10\%}^\tau = a_1 + 2a_2 Inv_{10\%} + 3a_3 Inv_{10_\%}^2 + 3a_4 \left(Inv_{10\%} - \overline{Inv} \right)^2 I_{\left\{ Inv_{10\%} > \overline{Inv} \right\}},$$

where $Inv_{10\%}$ is the 10th percentile of inventories and, of course, here $I_{\left\{ Inv_{10\%} > \overline{Inv} \right\}} = 0$. The standard errors associated with the t-values are corrected for possible serial correlation in the residuals following the procedure of Newey and West (1987) with a bandwidth of 12 months. Column 8 reports the p-value of a Wald test of the null hypothesis that the quadratic and cubic terms are zero. The last column reports the R^2 of the regression.

τ	at 10%	t-Stat.	at Mean	t-Stat.	at 90%	t-Stat.	Wald	R^2
1	-0.1813	-3.2685	-0.1208	-3.0485	-0.0156	-0.7035	0.0002	0.2491
3	-0.1458	-3.7754	-0.1078	-3.8104	-0.0314	-1.3458	0.0003	0.3401
6	-0.1100	-4.7194	-0.0914	-3.9239	-0.0417	-2.5649	0.0032	0.4400
9	-0.0871	-4.3069	-0.0788	-5.2016	-0.0445	-2.2504	0.0244	0.4906
12	-0.0718	-4.6050	-0.0692	-4.7601	-0.0443	-2.8136	0.1036	0.5156
17	-0.0560	-4.8382	-0.0581	-4.4883	-0.0422	-3.5239	0.2943	0.5355
24	-0.0439	-3.9173	-0.0488	-5.7516	-0.0392	-3.1138	0.4384	0.5434

Table 10

The table reports the results of a regression of the risk premium (expected holding return) on deseasonalized log inventories. The dependent variable in the left panel is the risk premium excluding the seasonal factors, while that of the right panel is the risk premium driven only by the commodity factors $\boldsymbol{\beta}_t$. The parameter b is the estimated slope coefficient and s.e. is its associated standard error; τ refers to the maturity of the contract and h is the holding period. Standard errors are computed using the Newey-West method for correcting serial correlation with a bandwidth of 12 months.

		Risk premium (exc. Seasonality)				Risk premium (Only Comm. Factors)			
τ	h	b	s.e.	p-value	R^2	b	s.e.	p-value	R^2
1	1	-2.6073	8.3780	0.7558	0.0006	-12.8176	9.2945	0.1688	0.0143
3	1	-14.6025	8.3817	0.0824	0.0404	-25.2134	9.9989	0.0122	0.1004
3	3	-10.9980	8.0595	0.1733	0.0279	-21.4216	9.7276	0.0283	0.0862
6	1	-16.6253	7.6650	0.0308	0.0536	-27.6669	9.1404	0.0027	0.1254
6	3	-20.0479	6.2030	0.0014	0.1205	-30.9873	7.9064	0.0001	0.2146
6	6	-19.3494	5.6429	0.0007	0.1478	-30.0268	7.4444	0.0001	0.2431
12	1	-12.8503	6.9123	0.0639	0.0330	-24.2690	7.4913	0.0013	0.1094
12	3	-15.3533	4.9455	0.0021	0.0865	-26.7880	5.7291	0.0000	0.2186
12	6	-18.0205	3.1250	0.0000	0.2279	-29.3955	4.1815	0.0000	0.3996
12	12	-19.7086	2.7203	0.0000	0.3252	-30.6592	4.3417	0.0000	0.4367
18	1	-15.3899	9.1443	0.0933	0.0271	-26.8063	9.4488	0.0048	0.0798
18	3	-17.1427	6.4325	0.0081	0.0645	-28.6472	6.5796	0.0000	0.1671
18	6	-16.8412	4.4219	0.0002	0.1314	-28.4110	4.3400	0.0000	0.3150
18	12	-15.9515	2.5534	0.0000	0.2702	-27.4126	2.8292	0.0000	0.5051
18	18	-17.3102	2.0026	0.0000	0.3518	-28.3396	3.2744	0.0000	0.4978
24	1	-22.2433	13.3964	0.0978	0.0219	-33.4645	13.6798	0.0150	0.0488
24	3	-25.1948	8.8814	0.0048	0.0591	-36.5482	8.8753	0.0000	0.1191
24	6	-22.9285	6.3155	0.0003	0.1146	-34.4252	5.8633	0.0000	0.2376
24	12	-16.7550	4.1791	0.0001	0.1669	-28.3428	3.5019	0.0000	0.3986
24	18	-15.0129	2.7181	0.0000	0.2455	-26.4353	2.4650	0.0000	0.5134
24	24	-16.2620	1.8656	0.0000	0.3507	-27.2573	2.6331	0.0000	0.5337

Table 11

The table reports the results of a regression of the risk premium (expected holding return) on the deseasonalized "hedging pressure." The dependent variable in the left panel is the risk premium excluding the seasonal factors, while that of the right panel is the risk premium driven only by the commodity factors $\boldsymbol{\beta}_t$. The parameter b is the estimated slope coefficient and s.e. is its associated standard error; τ refers to the maturity of the contract and h is the holding period. Standard errors are computed using the Newey-West method for correcting serial correlation with a bandwidth of 12 months.

		Risk premium (exc. Seasonality)				Risk premium (Only Comm. Factors)			
τ	h	b	s.e.	p-value	R^2	b	s.e.	p-value	R^2
1	1	-0.3150	0.0969	0.0013	0.0474	-0.3012	0.1028	0.0036	0.0413
3	1	-0.0704	0.0871	0.4200	0.0049	-0.0555	0.0931	0.5515	0.0025
3	3	-0.0655	0.0796	0.4113	0.0052	-0.0525	0.0851	0.5378	0.0027
6	1	-0.0187	0.0865	0.8287	0.0004	-0.0024	0.0917	0.9789	0.0000
6	3	0.0076	0.0693	0.9126	0.0001	0.0221	0.0762	0.7716	0.0006
6	6	0.0024	0.0640	0.9699	0.0000	0.0145	0.0713	0.8385	0.0003
12	1	-0.0430	0.0827	0.6039	0.0020	-0.0245	0.0858	0.7753	0.0006
12	3	-0.0089	0.0578	0.8783	0.0002	0.0080	0.0626	0.8989	0.0001
12	6	-0.0140	0.0420	0.7398	0.0007	0.0006	0.0505	0.9903	0.0000
12	12	-0.0359	0.0461	0.4372	0.0056	-0.0251	0.0566	0.6582	0.0015
18	1	0.0829	0.1056	0.4333	0.0041	0.1028	0.1060	0.3331	0.0060
18	3	0.0630	0.0683	0.3572	0.0046	0.0814	0.0699	0.2448	0.0069
18	6	0.0242	0.0444	0.5866	0.0014	0.0406	0.0485	0.4032	0.0033
18	12	-0.0148	0.0339	0.6630	0.0012	-0.0019	0.0424	0.9634	0.0000
18	18	-0.0273	0.0382	0.4762	0.0046	-0.0172	0.0486	0.7236	0.0009
24	1	0.3030	0.1577	0.0556	0.0213	0.3237	0.1568	0.0398	0.0235
24	3	0.2012	0.0942	0.0335	0.0196	0.2207	0.0944	0.0199	0.0222
24	6	0.0920	0.0591	0.1205	0.0095	0.1097	0.0611	0.0735	0.0123
24	12	0.0186	0.0390	0.6330	0.0011	0.0331	0.0435	0.4475	0.0028
24	18	-0.0021	0.0325	0.9492	0.0000	0.0097	0.0400	0.8089	0.0004
24	24	-0.0138	0.0346	0.6908	0.0013	-0.0042	0.0445	0.9246	0.0001

Figure 1: Loadings of 1-month (upper panels) and 12-months (lower panels) holding period returns on the yield curve and commodity factors as a function of the maturity of the contract in the calibrated example. The parameter values are $\phi_{11} = 0.99$, $\phi_{22} = 0.94$, $\phi_{33} = 0.96$, $\phi_{44} = 0.98$, $\phi_{55} = 0.83$, $\phi_{66} = 0.69$, $\phi_{77} = 0.68$, $\theta_1 = 0.07$, $\theta_2 = 0.25$, and $\omega = 0.008$.
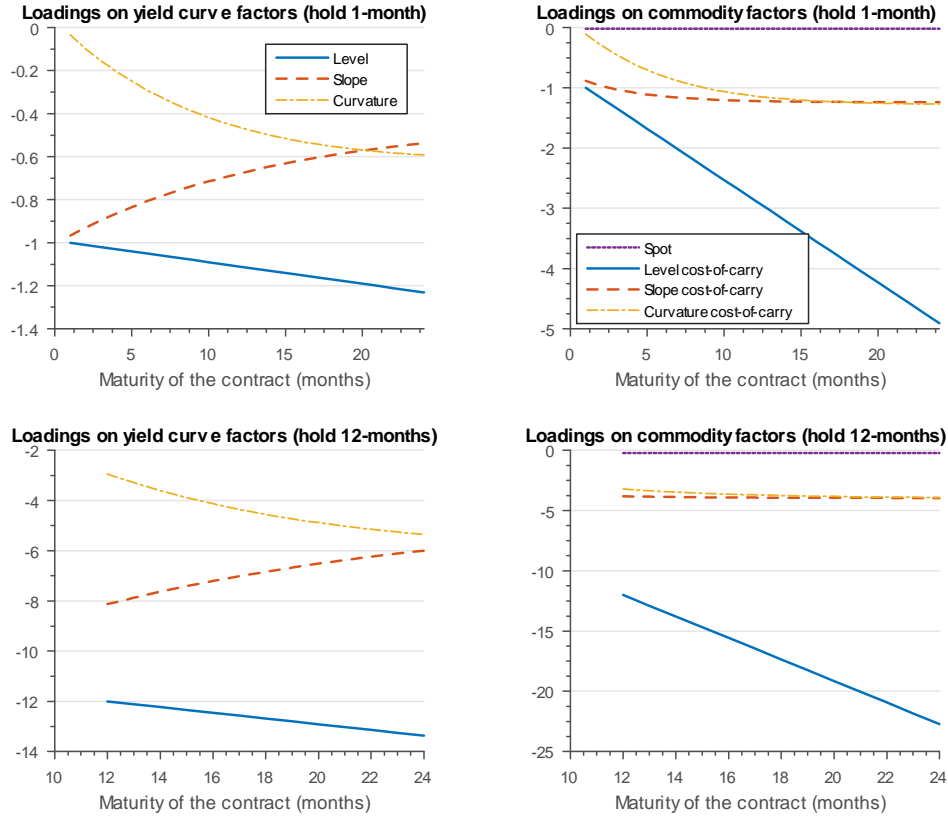
Figure 2: Estimates of yield curve factors (Step 1): level, slope, and curvature (δ_{1t}, δ_{2t}, and δ_{3t}) expressed in percentage points and annualized.
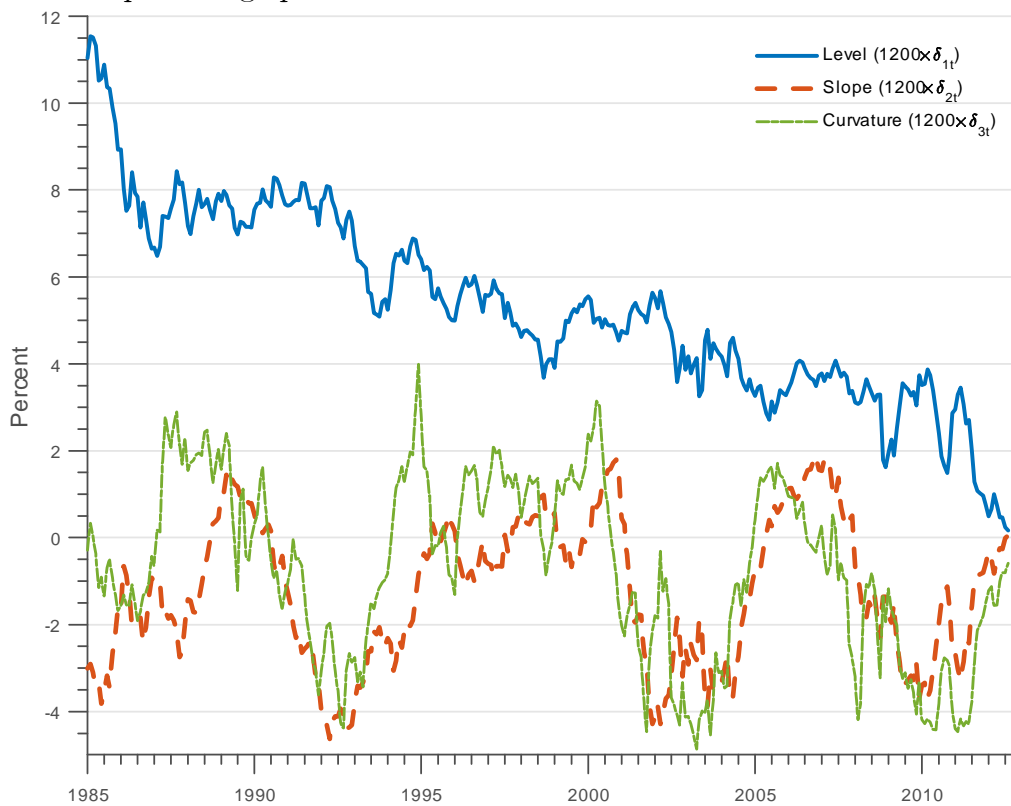
Figure 3: Estimates of commodity factors (Step 2) of the models with stochastic and deterministic seasonality: deseasonalized spot β_{0t}; level, slope, and curvature of the cost-of-carry curve (β_{1t}, β_{2t}, and β_{3t}), and seasonal factors ξ_{1t} and ξ_{2t}. The values of the commodity factors β_{1t}, β_{2t}, and β_{3t} in the model with deterministic seasonality are displayed in the right axis.
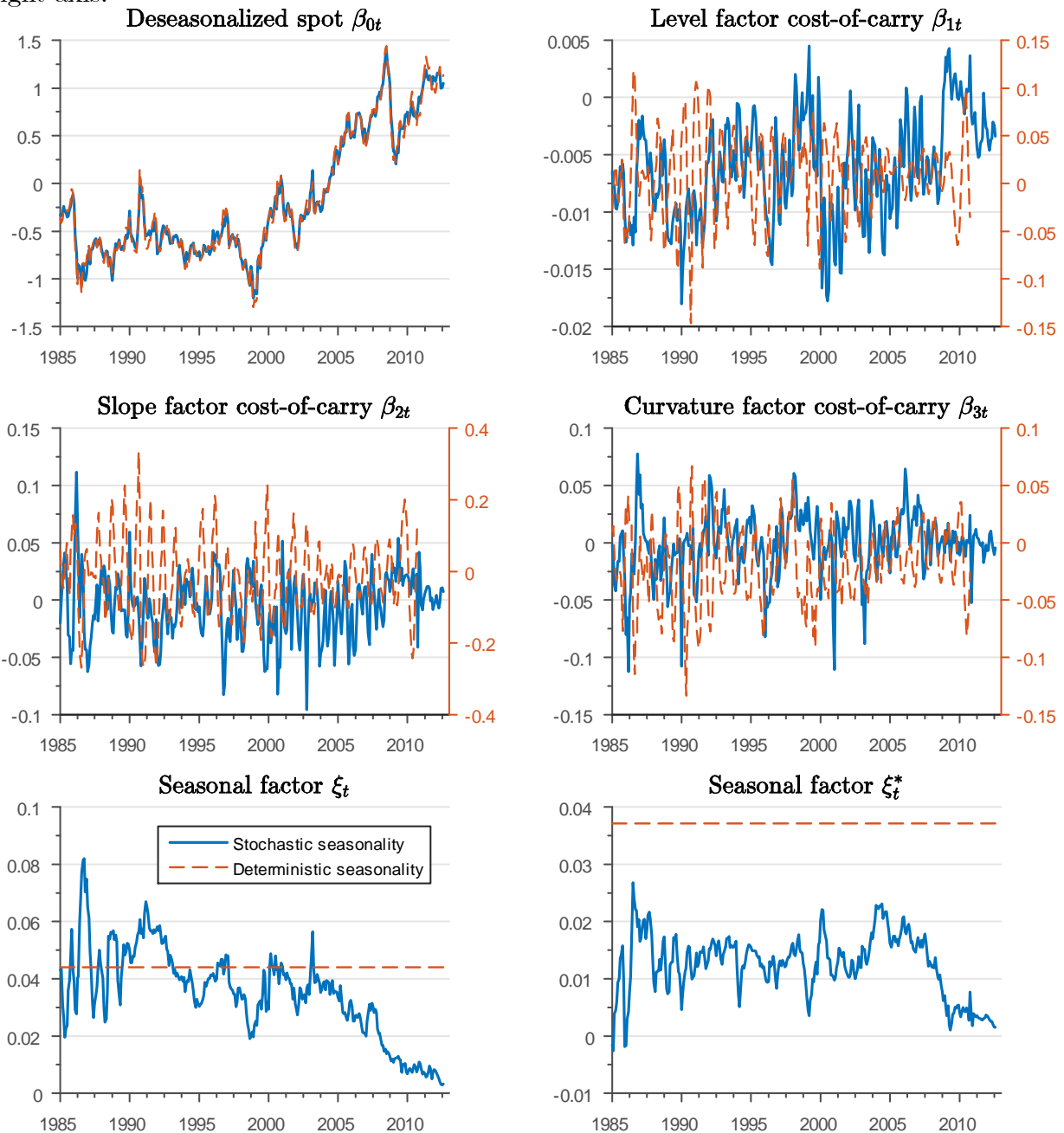
Figure 4: Factor loadings on the slope and curvature commodity factors β_{2t} and β_{3t}.
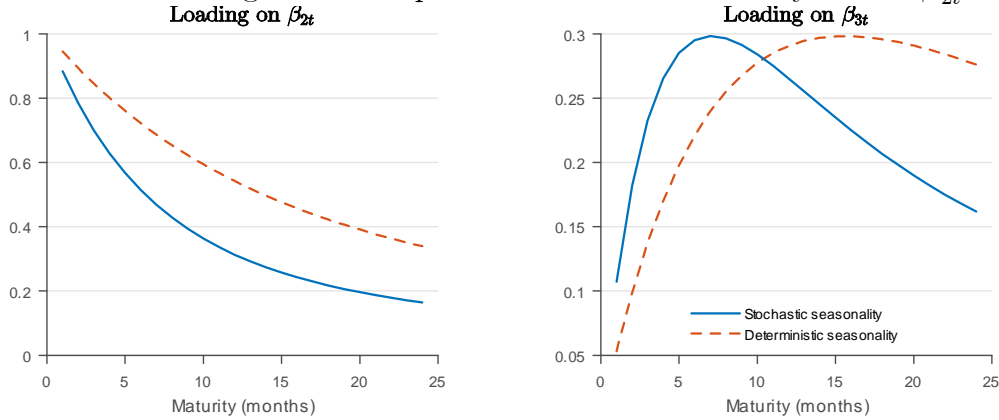
Figure 5: Seasonal components. The upper panel displays the implied seasonal component of the spot commodity price $s_t - \beta_{0t}$ in the models with deterministic and stochastic seasonality. The figure also displays the months in circles in the model with stochastic seasonality. Most seasonal peaks are in December (D) and troughs in June (J). The lower panel displays the seasonal component of log-inventories and the implied seasonal component of the spot commodity price. Both series are standardized.
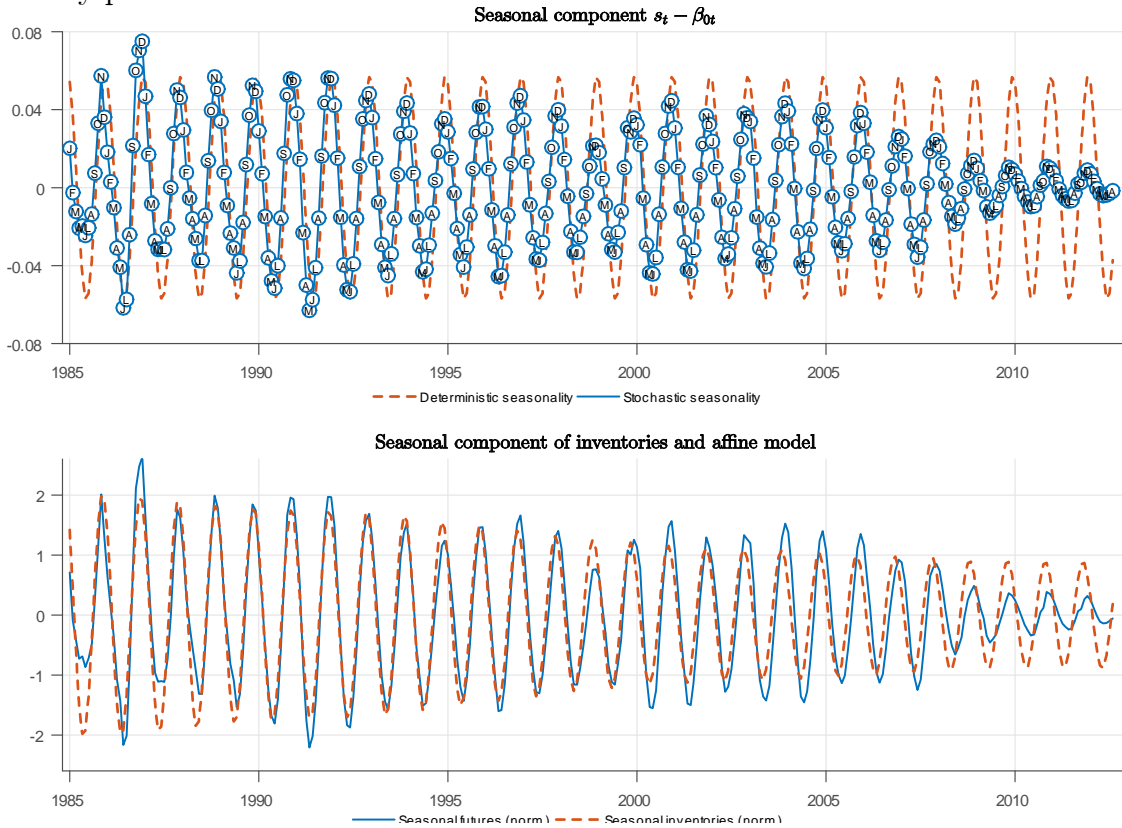
Figure 6: Selected fitted log-futures curves. The figure shows fitted log-futures curves, deseasonalized fitted log-futures curve, and actual log-futures prices for selected dates.
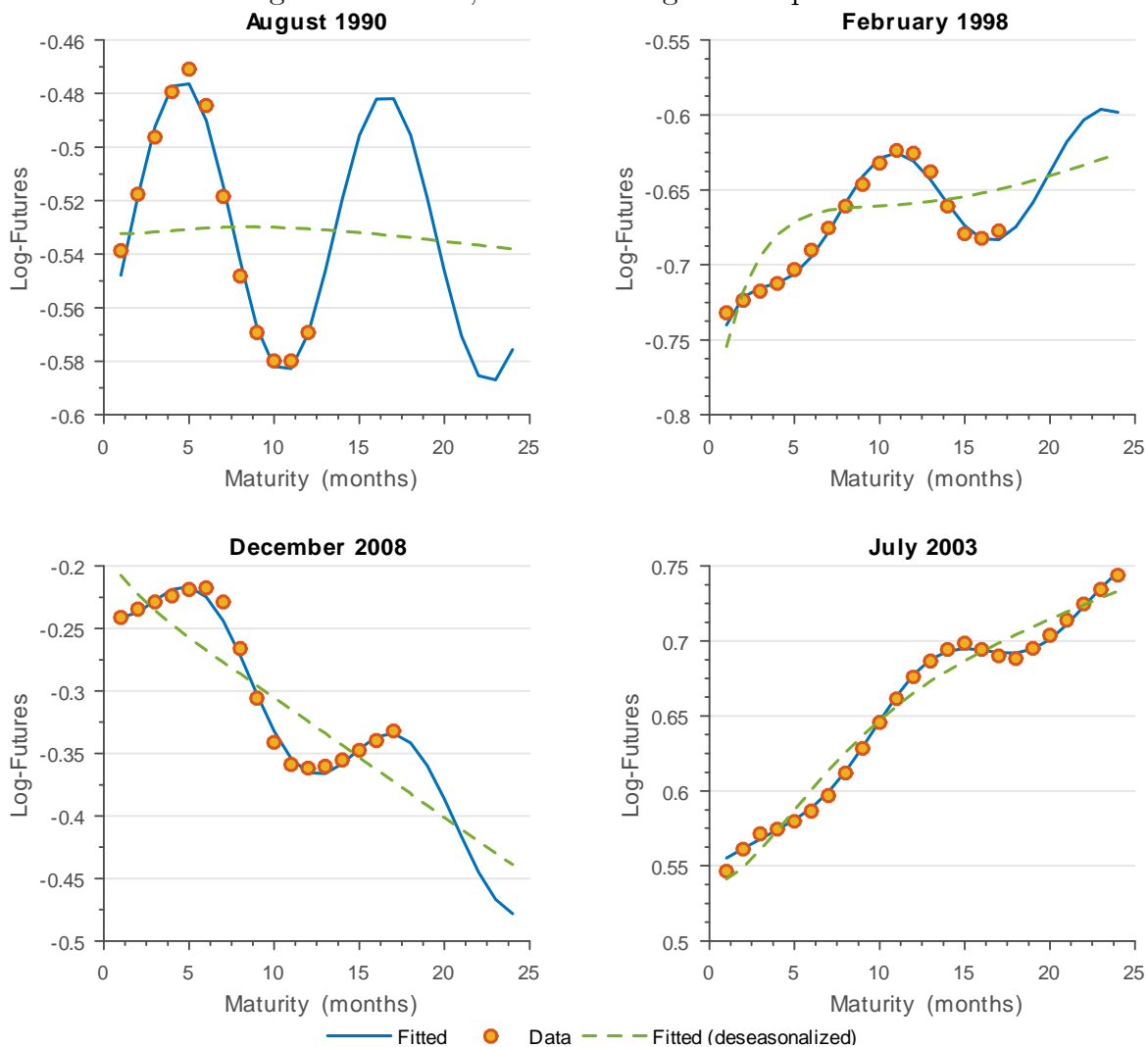
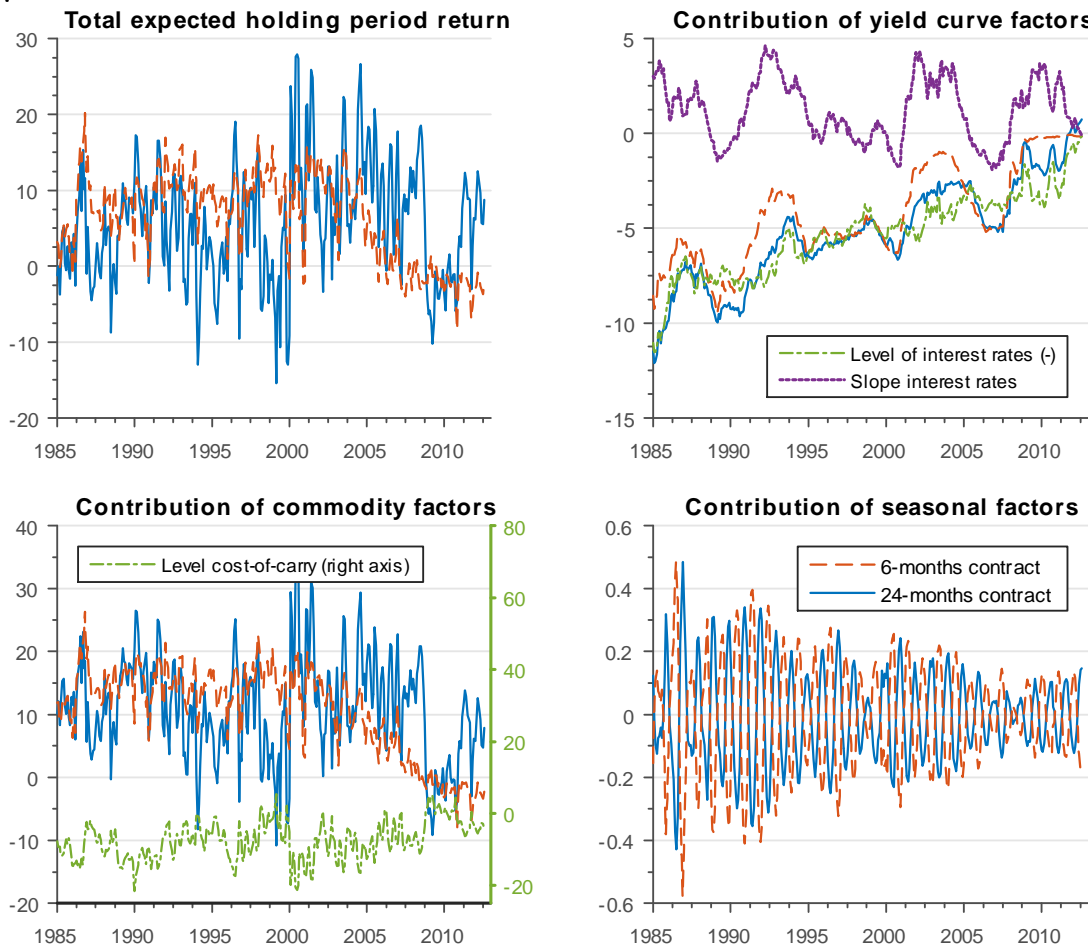Figure 7: This figures shows the expected 6-months holding returns of a 6 months and a 24 months futures contract. Returns are expressed in percentage points and on an annualized basis. The upper left panel displays the total expected return. The other panels display the contribution of the different factors. The upper right panel also shows the negative of the level factor of the interest rates and the bottom left panel adds the level of the cost-of-carry factor.

Figure 8: Expected 6-months and 24-months holding returns of a 24 months futures contract. Returns are expressed in percentage points and on an annualized basis. The upper left panel displays the total expected return. The other panels display the contribution of the different factors. The upper right panel also shows the negative of the level factor and the slope of the interest rates and the bottom left panel adds the commodity spot factor.
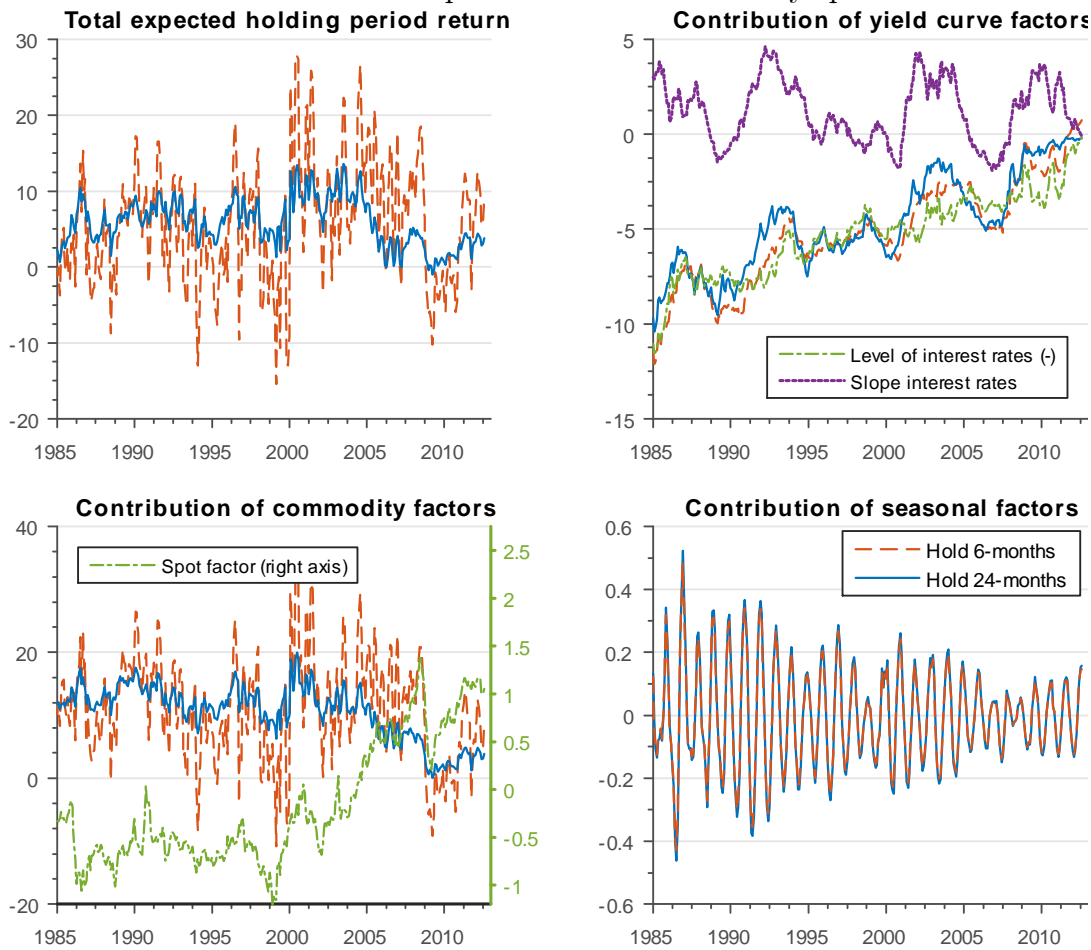
Figure 9: Decomposition of average 1-month expected holding return into spot and term premia. The figure shows the average 1-month expected log-holding return and its decomposition into spot premium and term premium for the entire sample period and dividing the sample before and after 2003. Returns are annualized and expressed in percentage terms.
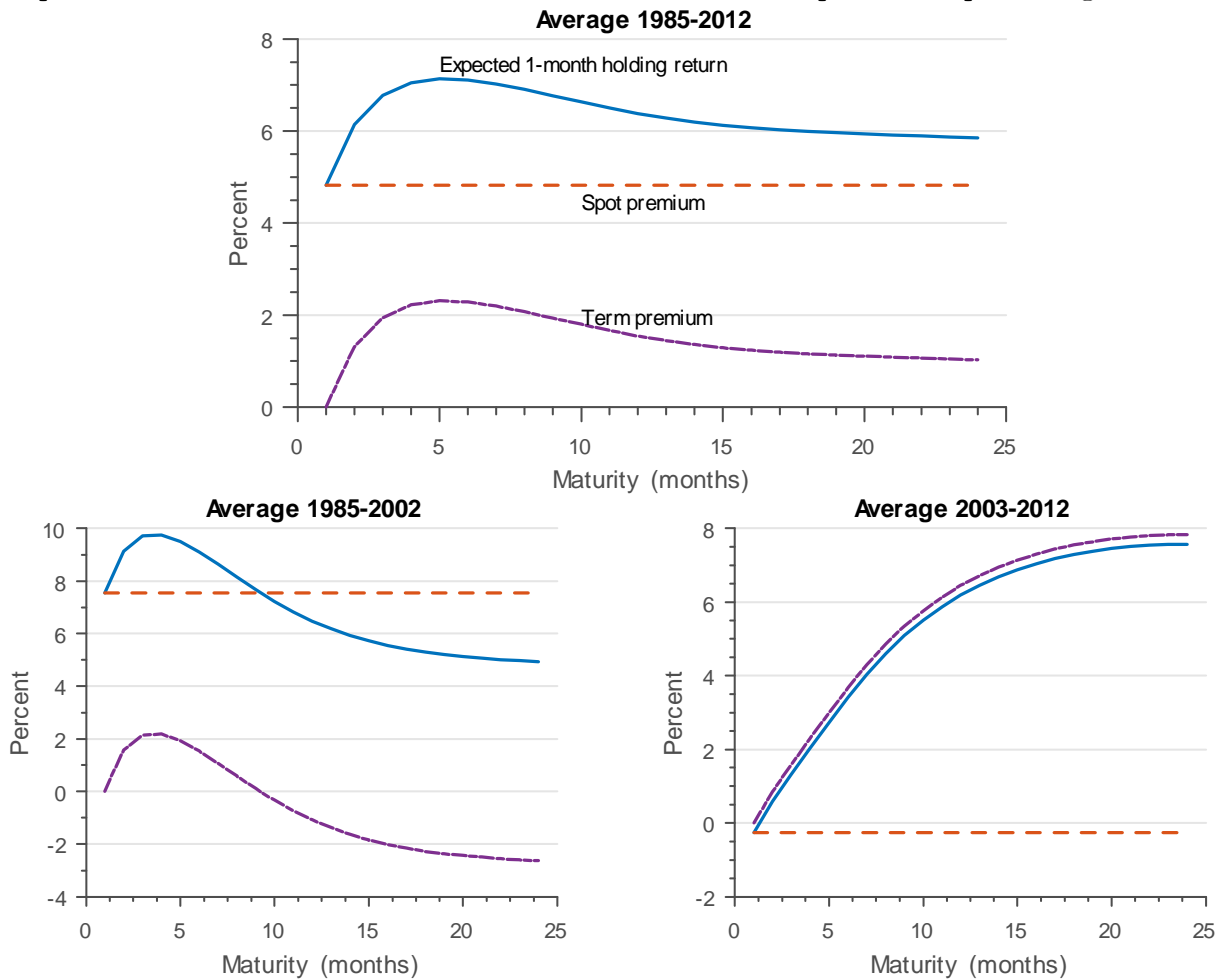
Figure 10: Spot and term premium during the 2008-2009 recession. The upper left panel display the spot price (solid line) and all futures prices (dotted line) together with the cointegration relation between the factors δ_{1t}, δ_{2t}, and β_{0t} (left axis). The upper right figure displays the spot and term premiums during the peak and trough months on heating oil prices. The bottom panels displays the fitted and actual futures curves in those two dates.
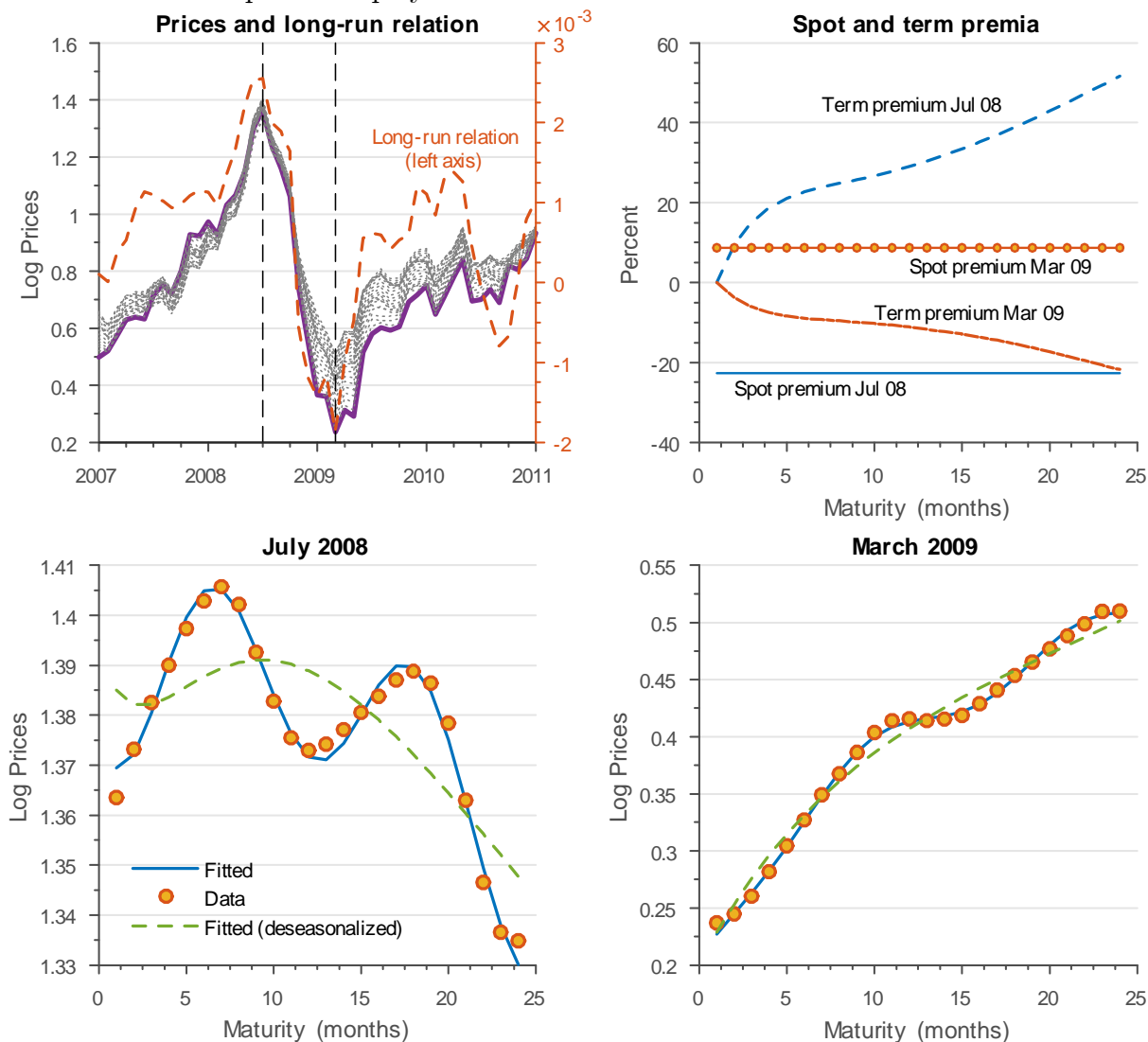
Figure 11: Inventories and cost-of-carry level factor. The figure displays the cost-of-carry level factor β_{1t} and the non-seasonal component of inventories. Series are standardized.