Risk premia and seasonality in commodity futures

Constantino Hevia[1] | Ivan Petrella[2,3] | Martin Sola[1]

[1]Department of Economics, Universidad Torcuato Di Tella, Buenos Aires, Argentina

[2]Warwick Business School, Warwick, UK

[3]CEPR, London, UK

Correspondence
Martin Sola, Department of Economics, Universidad Torcuato Di Tella, Av. Pres. Figueroa Alcorta 7350, Ciudad de Buenos Aires, Argentina.
Email: msola@utdt.edu

Summary

We develop and estimate a multifactor affine model of commodity futures that allows for stochastic seasonality. We document the existence of stochastic seasonal fluctuations in commodity futures and that properly accounting for the cost-of-carry curve requires at least three factors. We estimate the model using data on heating oil futures and analyze the contribution of the factors to risk premia. Correctly specifying seasonality as stochastic is important to avoid erroneously assigning those fluctuations to other risk factors. We also estimate a nonlinear version of the model that imposes the zero lower bound on interest rates and find similar results.

1 | INTRODUCTION

Futures prices in many energy and agricultural commodities display seasonal fluctuations. Often, those fluctuations are not perfectly predictable. From the point of view of market participants, stochastic seasonal fluctuations imply a source of risk that manifests itself in futures prices and risk premia. The purpose of this paper is to develop and estimate an affine model of futures prices that allows for stochastic variations in seasonality. We use the model to analyze the implications of stochastic seasonal fluctuations for the pricing of commodity futures and risk premia.

The earlier literature, such as Gibson and Schwartz (1990) and Litzenberger and Rabinowitz (1995), assumes that the price of commodity futures depends on two factors: a spot price factor and a cost-of-carry (or convenience yield) factor. Schwartz (1997) extends the model to include a stochastic interest rate, which became the standard three-factor model of commodity futures. Casassus and Collin-Dufresne (2005) provides a different interpretation of the three-factor model that we use to relate the cost-of-carry and the spot commodity price. Miltersen and Schwartz (1998) use the affine framework to price derivatives on commodity futures, and Hamilton and Wu (2014) study risk premia in oil futures markets. Finally, existing models of seasonal commodity futures only allow for deterministic seasonal fluctuations in prices (e.g., Sorensen, 2002) or in the cost-of-carry (e.g., Borovkova & Geman, 2006).

This paper extends the current literature in three dimensions. First, our model features stochastic seasonal fluctuations in both the spot price and the cost-of-carry. By attaching market prices of risk to seasonal factors, we are able to measure the risks associated with stochastic seasonal shocks. Second, following much of the literature on bond prices, we assume that bond yields depend on three factors. And third, consistent with data on a set of agricultural and energy commodity futures, we assume that the cost-of-carry curve is driven by three factors.

We begin with a preliminary analysis of futures prices for a group of energy and agricultural commodities that provide support to our modeling assumptions. We first document that seasonal fluctuations in this set of commodity futures are stochastic. Second, we perform a principal components analysis on the cost-of-carry curve for each commodity. We conclude that we need at least three factors to appropriately account for the dynamics of the cost-of-carry curve. The first three principal components account for 90% or less of the variability of the cost-of-carry curve, with the third principal component contributing between 4% and 14% of the variability depending on the commodity and method used to compute the curve. Furthermore, the first principal component only accounts for between 16% and 80% of the variability.

These results suggest that a model in which the cost-of-carry depends on a single factor misses important features of the data, and motivates our assumption that the cost-of-carry depends on three factors.

We next describe an affine model of commodity futures with stochastic seasonality in discrete time. The model prices zero-coupon bonds and commodity futures of any maturity, and stochastic variations in seasonality are driven by a pair of random walks (Hannan, 1964). To estimate the model we need to impose identifying restrictions. An approach often used in the bond pricing literature is to impose an identifying assumption such that the yield curve adopts an augmented Nelson and Siegel (1987) specification (Christensen, Diebold, & Rudebusch, 2011). We extend this literature to the pricing of commodity futures and find conditions under which the yield curve and the cost-of-carry curve derived from the model adopt a Nelson–Siegel representation.

We estimate the model of bond and commodity futures prices using data on US zero coupon bonds and futures prices on heating oil for the period January 1984 to April 2017. We concentrate our analysis on heating oil but leave for an online supplement (Supporting Information) estimation results using data on soybeans futures. The model matches the cross-section of futures prices over time, including their seasonal pattern. We find strong evidence of stochastic seasonality: the peaks and troughs of the seasonal cycle vary over the years, and the amplitude of the seasonal fluctuations decreases over time, particularly at the end of the sample. Consistent with the theory of storage, the moderation of the estimated seasonal component coincides with a similar mitigation of the seasonal component in stocks of heating oil inventories.

Expected returns of holding a futures contract fluctuate widely over time, and most of those fluctuations come from variations in the spot, cost-of-carry, and yield curve factors. Although nonnegligible, the contribution of seasonal shocks to risk premia is small. Seasonal shocks account for about 0.5 percentage points of expected returns but become less relevant at the end of our sample period. Therefore, correctly specifying seasonality in futures prices as stochastic is important mostly to avoid erroneously assigning those fluctuations to other risk factors. When we estimate the model imposing deterministic seasonality, the omitted time variation in the seasonal pattern manifests itself as large fluctuations in the cost-of-carry factors, which, in turn, translate into large and spurious fluctuations in estimated risk premia.

It is often argued that interest rate shocks have a minor impact on the time variation in commodity futures risk premia. Schwartz (1997) assumes a constant interest rate because interest rate fluctuations are orders of magnitudes lower than those in futures returns. Using the three-factor model, Casassus and Collin-Dufresne (2005) argue that the market price of interest rate shocks is barely significant. In contrast, we find that yield curve factors have a significant impact on risk premia, mostly at medium and lower frequencies. In our sample period, interest rates declined from about 12% to roughly zero. The contribution of interest rate factors to expected holding returns went from about −10 percentage points to 0 over the same time frame. When the slope of the yield curve is positive, long-term contracts are relatively more expensive than shorter contracts, while the reverse holds when the yield curve is inverted. Thus changes in the slope of the yield curve over time affect futures prices and risk premia. Overall, we find that several measures of risk premia began 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 smaller than that of the other factors.

From December 2008 to December 2015, the US Federal Reserve set the policy interest rate to virtually zero. We explore to what extent our results change if we explicitly impose the zero lower bound on interest rates. To that end, we adapt the yield curve model proposed by Wu and Xia (2016) and jointly estimate a model of commodity futures and bond prices that imposes the zero lower bound constraint. Although this constraint is important for bond pricing, we find that it has a minor impact on futures prices and risk premia relative to the results of the baseline model.

The paper is organized as follows. Section 2 documents two empirical regularities for commodity futures and Section 3 describes an affine model with stochastic seasonality. Section 4 proposes a representation of the model that adopts Nelson–Siegel functional forms and Section 5 describes the estimation method. Section 6 contains the main results and Section 7 concludes. Proofs are relegated to the Appendix (the Supporting Information contains additional material).

2 | SOME EMPIRICAL PROPERTIES OF COMMODITY FUTURES

In this section we document two empirical regularities in a group of energy and agricultural commodity futures. The first is that seasonal fluctuations are stochastic. The second is that we need at least three factors to properly account for the variability of the cost-of-carry curve. Those results provide support to our subsequent modeling assumptions.

2.1 | Stochastic seasonality in commodity futures

Futures prices in many energy and agricultural commodities display seasonal fluctuations.[1] Although it is common to model seasonality as deterministic cycles, in this section we argue that, for many commodities, those seasonal fluctuations are in fact stochastic. This distinction is important because stochastic seasonality implies an additional risk factor that affects risk premia.

Let us decompose a stochastic process $z_t = z_t^n + z_t^s$ into its seasonal (z_t^s) and nonseasonal (z_t^n) components. Deterministic seasonality can be modeled in terms of trigonometric functions:

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

where $2\pi j/12$ are seasonal frequencies and ξ_j and ξ_j^* are parameters. The seasonal effect z_t^s is the sum of six deterministic cycles with periods of $12/j$ months, for $j = 1, 2, \dots, 6$. The frequency $2\pi/12$ corresponds to a period of 12 months and is known as the fundamental frequency. The other frequencies represent waves with periods of less than a year.

Following Hannan (1964), we model stochastic seasonality by letting the parameters ξ_j and ξ_j^* evolve as a random walks. Thus the stochastic seasonal component is given by

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

where $\xi_{jt} = \xi_{jt-1} + \nu_{jt}$ and $\xi_{jt}^* = \xi_{jt-1}^* + \nu_{jt}^*$ for $j = 1, 2, \dots, 6$. The shocks ν_{jt} and ν_{jt}^* are orthogonal and normally distributed with variances σ_j^2 and σ_j^{*2}. When the seasonal cycle is dominated by a single seasonal peak and trough within the year (only the fundamental frequency matters), the seasonality process collapses to

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

where ξ_t and ξ_t^* are two independent random walks.

Table 1 shows the results of three tests of the null hypothesis that seasonality in futures prices is deterministic for a number of energy commodities (gasoil, gasoline, heating oil, and natural gas) and agricultural commodities (corn, soybean, and wheat). The alternative hypothesis is that seasonal fluctuations are stochastic.[2] We use Canova and Hansen's (1995) nonparametric test for parameter stability and its spectral extension developed by Busetti and Harvey (2003). The third is a parametric test also proposed by Busetti and Harvey. For most commodities and contract maturities, the tests strongly reject the null hypothesis of deterministic seasonality.[3]

These results suggest that the common practice of imposing a deterministic seasonal model in commodity futures (at least for those considered here) is flawed. Moreover, deseasonalizing the data prior to any empirical analysis is also problematic for two reasons. First, extracting the seasonal component from each futures contract in isolation does not guarantee that the seasonal factors are consistent across maturities.[4] And second, by deseasonalizing the data prior to any empirical analysis one is unable to measure the contribution of seasonal shocks to risk premia. We deal with both concerns by imposing a common seasonal factor that affects futures prices of all maturities within the context of an affine model of futures prices.

2.2 | Number of factors in the cost-of-carry curve

To estimate an affine model of commodity futures we need to determine the number of risk factors. Researchers agree that three factors capture most of the variation in bond yields and we follow this practice below. But less is known about the factor structure in the cost-of-carry curve. The basic model of commodity futures assumes that a single factor drives variations in the cost-of-carry. In this section we show that, in fact, we need at least three factors to properly account for the dynamics of the cost-of-carry.

---

[1]A Wald test of significance of seasonal dummies in autoregressive models of futures prices strongly reject the null of lack of seasonality for all the commodities considered below.
[2]The data are end-of-month log settlement futures prices on contracts with maturities up to 18 months ahead. Details of the sources and construction of the data are given in the Supporting Information.
[3]The evidence is mixed when considering natural gas, which is often used as an example of a seasonal commodity.
[4]The Supporting Information shows that seasonal patterns extracted from univariate models are not synchronized across maturities.

TABLE 1 Test of stochastic versus deterministic seasonality in futures prices

	Contract maturity τ months					
	1	**2**	**3**	**6**	**9**	**12**
Energy						
Gasoil						
CH	3.7065***	1.0628**	0.9536**	0.6203*	0.3104	0.6865*
BH (nonparam.)	4.4755***	1.7076***	1.9084***	0.9575**	0.1455	0.5695
BH (param.)	11.738***	1.3052***	3.7776***	7.5316***	1.8639***	1.2931***
Gasoline						
CH	0.7967**	0.8284**	0.8368**	0.4411	0.8771**	2.2912***
BH (nonparam.)	1.155***	1.3338***	1.4022***	0.1824	0.2539	2.244***
BH (param.)	3.8417***	0.4951	1.7158***	26.415***	4.5124***	0.4165
Heating oil						
CH	0.8418**	0.768**	0.6652*	0.1954	0.8524**	1.4317***
BH (nonparam.)	1.0308**	1.2491***	1.2211***	0.2055	0.9599**	1.5977***
BH (param.)	5.303***	11.162***	15.177***	9.5591***	1.0832***	15.704***
Natural gas						
CH	1.542***	0.3236	0.2023	0.4106	0.8064**	0.2779
BH (nonparam.)	3.5174***	0.3736	0.283	0.7733**	0.4905	0.398
BH (param.)	4.2392***	14.496***	18.123***	1.2262***	13.781***	6.514***
Agricultural						
Corn						
CH	8.1784***	0.8555**	0.7226*	0.9563**	2.4234***	2.5611***
BH (nonparam.)	18.572***	1.2667***	0.3207	1.0738***	6.5478***	7.4085***
BH (param.)	1.3201***	7.0038***	1.6591***	6.8562***	8.2679***	8.3063***
Wheat						
CH	5.9498***	1.6014***	0.4361	1.0749***	1.0055**	0.2214
BH (nonparam.)	7.8489***	3.061***	0.3709	0.508	1.2175***	0.2855
BH (param.)	1.5193***	5.1433***	1.6911***	0.1694	1.3814***	4.6131***
Soybean						
CH	10.119***	0.4443	0.5835	1.1623***	2.8187***	4.9593***
BH (nonparam.)	0.9757**	0.7387*	0.4419	0.4659	7.8844***	11.549***
BH (param.)	4.3753***	6.7288***	2.0364***	7.0118***	9.9595***	12.379***

Note. The table shows tests of deterministic versus stochastic seasonality in heating oil futures. τ stands for maturity. CH denotes Canova and Hansen's (1995) test, while BH denotes Busetti and Harvey's (2003) parametric and nonparametric tests. Under the null hypothesis of deterministic seasonality, the statistics are distributed as a generalized Von-Mises random variable with 2 degrees of freedom. Asterisks denote significance at ***1%, **5%, *10% levels. For gasoil, data starts in August 1990; for gasoline, in January 1985; for natural gas, in April 1990; for all other commodities, the sample starts on the last day of 1983. For all commodities the ending date is April 2017.

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 $F_t^{(\tau)}$. Let S_t denote the spot commodity price; $f_t^\tau = \log(F_t^{(\tau)})$, the log of the futures price; and $s_t = \log(S_t)$, the log of the spot price. The log-basis of a commodity futures is defined as $f_t^{(\tau)} - s_t$. Following Miltersen and Schwartz (1998) and Trolle and Schwartz (2009) we define the cost-of-carry curve (net of bond yields) at t as the value $u_t^{(\tau)}$ (viewed as a function of maturity τ) such that the basis of the commodity futures can be written as

$$f_t^{(\tau)} - s_t = \tau(y_t^{(\tau)} + u_t^{(\tau)}), \tag{3}$$

where $y_t^{(\tau)}$ is the yield on a zero-coupon bond that matures in τ periods.

To infer the number of factors necessary to capture the variability of the cost-of-carry curve, we compute principal components of log futures prices net of the contribution of the spot, seasonal, and yield curve factors. Based on Equation 3, we subtract from futures prices the contribution of bond yields and define $w_t^{(\tau)} = f_t^{(\tau)} - \tau y_t^{(\tau)}$. We use two procedures to isolate the contribution of the spot and seasonal factors. In the first, we run a series of cross-sectional regressions (one for each date) of $w_t^{(\tau)}$ on a constant and sine and cosine waves according to

$$w_t^{(\tau)} = \mathsf{a}_{0t} + \cos\left(\frac{2\pi}{12}(t + \tau)\right)\mathsf{a}_{1t} + \sin\left(\frac{2\pi}{12}(t + \tau)\right)\mathsf{a}_{2t} + \varepsilon_t^{(\tau)}. \tag{4}$$

TABLE 2 Number of factors in the yield curve and the cost-of-carry curve

	PC1	PC2	PC3
Yield curve			
$y_t^{(\tau)}$	99.67	0.31	0.02
Energy			
Gasoil			
$\varepsilon_t^{(\tau),CS}$	45.52	19.84	11.28
$\varepsilon_t^{(\tau),KF}$	72.46	5.65	4.00
Gasoline			
$\varepsilon_t^{(\tau),CS}$	35.24	28.08	13.61
$\varepsilon_t^{(\tau),KF}$	38.59	19.62	11.98
Heating oil			
$\varepsilon_t^{(\tau),CS}$	50.59	20.10	12.04
$\varepsilon_t^{(\tau),KF}$	63.61	16.42	5.58
Natural gas			
$\varepsilon_t^{(\tau),CS}$	80.55	7.96	5.15
$\varepsilon_t^{(\tau),KF}$	55.92	16.57	14.03
Agricultural			
Corn			
$\varepsilon_t^{(\tau),CS}$	25.31	16.60	12.41
$\varepsilon_t^{(\tau),KF}$	28.22	18.28	12.26
Wheat			
$\varepsilon_t^{(\tau),CS}$	22.59	17.22	14.37
$\varepsilon_t^{(\tau),KF}$	35.48	19.59	10.79
Soybean			
$\varepsilon_t^{(\tau),CS}$	16.07	14.10	13.25
$\varepsilon_t^{(\tau),KF}$	26.83	17.14	11.76

Note. The table reports the contribution of the first three principal components to the variability of the yield curve and the cost-of-carry curve. $\varepsilon_t^{(\tau),CS}$ refers to the cost-of-carry curve obtained from the cross-sectional regressions and $\varepsilon_t^{(\tau),KF}$ is the cost-of-carry curve estimated using the Kalman filter.

The estimated constant ($\hat{a}_{0t}$) is a common shifter that affects futures prices of all maturities and captures the role of the spot factor. The coefficients associated with the sine and cosine waves ($\hat{a}_{1t}$ and $\hat{a}_{2t}$) capture the seasonal variations in the data, and the residuals $\hat{\varepsilon}_t^{(\tau)}$ are an estimate of (τ times) the cost-of-carry curve. The second approach is to extract the spot and seasonal factors using standard state-space techniques. We still use Equation 4 but now assume that the spot factor (a_{0t}) follows a first-order autoregressive process and the seasonal factors (a_{1t} and a_{2t}) follow independent random walks. We estimate the factors using the Kalman filter and construct the cost-of-carry curve using the residuals of Equation 4. Finally, we perform principal components analysis on the two ways of estimating the cost-of-carry curve.

Table 2 reports the contribution of the first three principal components to the variability of the two measures of the cost-of-carry curve.[5] The result is clear: We need at least three factors to properly account for the variability of the cost-of-carry curve. The first three principal components account for about 90% of the variance of the cost-of-carry curve, with the third principal component explaining between 5% and 14% of that variance. In light of these results, we augment the traditional affine model of commodity futures by assuming that fluctuations in the cost-of-carry are driven by three factors.

---

[5]We extract principal components from the unbalanced panel of futures prices as in Stock and Watson (2002).

3 | AFFINE MODEL OF COMMODITY FUTURES

The risk factors are represented by a vector of state variables $X_t \in R^n$ that evolves as

$$X_{t+1} = \mu + \Theta X_t + \Gamma \eta_{t+1}, \tag{5}$$

where $\eta_{t+1} \sim N(0, I)$ and Γ is lower triangular. Time periods are measured in months and the state vector includes factors capturing the stochastic variation in seasonality, which we specify below.

Nominal cash flows received at time $t + 1$ are priced using the stochastic discount factor

$$M_{t,t+1} = e^{-(r_t + \frac{1}{2}\Lambda_t'\Lambda_t + \Lambda_t'\eta_{t+1})}$$
$$\Lambda_t = \lambda_0 + \lambda_1 X_t, \tag{6}$$

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

$$r_t = \rho_0 + \rho_1' X_t, \tag{7}$$

where ρ_0 is a scalar and $\rho_1 \in R^n$. Since there is no evidence of seasonality in interest rates, we set to zero the loading of ρ_1 on the seasonal factors.

3.1 | Pricing government bonds

Let $P_t^{(\tau)}$ be the price of a τ-period zero-coupon bond. Absence of arbitrage implies that bond prices satisfy the pricing condition $P_t^{(\tau)} = E_t[M_{t,t+1}P_{t+1}^{(\tau-1)}]$. Using standard results (Ang & Piazzesi, 2003) one can show that log bond prices are affine functions of the risk factors

$$\log P_t^{(\tau)} = A_\tau + B_\tau' X_t, \tag{8}$$

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

$$A_\tau = A_{\tau-1} - \rho_0 + (\mu - \Gamma\lambda_0)'B_{\tau-1} + \frac{1}{2}B_{\tau-1}'\Gamma\Gamma'B_{\tau-1}, \tag{9}$$

$$B_\tau = (\Theta - \Gamma\lambda_1)'B_{\tau-1} - \delta_1, \tag{10}$$

with initial conditions $A_0 = 0$ and $B_0 = 0$. The yield on a τ-period zero-coupon bond at date t is thus

$$y_t^{(\tau)} = -\log P_t^{(\tau)}/\tau = a_\tau + b_\tau' X_t. \tag{11}$$

3.2 | Spot price and implied cost-of-carry

Consider a storable commodity with spot price S_t and with a net cost-of-carry of c_t, expressed as a continuously compounded rate of the spot price. The net cost-of-carry represents the storage and insurance costs of physically holding the commodity net of any convenience yield on inventory. It is the analog of the negative of the dividend yield of a stock and can be derived from equilibrium models such as Routledge, Seppi, and Spatt (2000).

To capture seasonality in the spot price and the cost-of-carry, we assume that their loadings on X_t are periodic functions of time. Let $\{m_t\}$ be a periodic sequence mapping the time t into the set of months $\{1, 2, \dots, 12\}$. We initialize the sequence by setting $m_t = t$ for $t = 1, 2, \dots, 12$, and let $m_{t+12k} = m_t$ for every t and k. We often use $\tilde{m}$ when referring to a generic month and impose the convention that $\tilde{m} + 1 = 1$ when $\tilde{m} = 12$.

Assume that the log spot commodity price is given by

$$s_t = \gamma_0 + \gamma_1^{m_t'} X_t, \tag{12}$$

where $\gamma_1^{m_t}$ depend on the season m_t. Since the payoff from holding the commodity between t and $t + 1$ is $e^{-c_{t+1}}S_{t+1}$, the principle of no arbitrage implies that the spot price S_t satisfies

$$S_t = E_t\left[M_{t,t+1}e^{-c_{t+1}}S_{t+1}\right]. \tag{13}$$

Following Casassus and Collin-Dufresne (2005), the next proposition (proved in the Supporting Information) states that there exists an affine and seasonal cost-of-carry process c_t such that the pricing condition (Equation 13) is satisfied given the evolution of the spot price (Equation 12).

Proposition 1. *(Cost-of-carry process) The net cost-of-carry consistent with the commodity price (Equation 13) is an affine and periodic function of the state variables $c_t = \psi_0^{m_t} + \psi_1^{m_t\prime} X_t$, where the scalar $\psi_0^{\tilde{m}}$ and the vector $\psi_1^{\tilde{m}}$ satisfy*

$$\psi_1^{\tilde{m}+1} = \gamma_1^{\tilde{m}+1} - \left[(\Theta - \Gamma\lambda_1)^{-1}\right]' \left(\gamma_1^{\tilde{m}} + \rho_1\right),$$

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

3.3 | 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 $F_t^{(\tau)}$. Entering into a futures contract involves no initial cash flow and a payoff of $S_{t+\tau} - F_t^{(\tau)}$ at time $t + \tau$. Therefore, if we let $M_{t,t+\tau}$ denote the stochastic discount factor for nominal payoffs received τ periods ahead, the principle of no-arbitrage implies

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

In Appendix A we show that log-futures prices satisfy

$$f_t^{(\tau)} = C_\tau^{m_t} + D_\tau^{m_t\prime} X_t, \tag{14}$$

where $C_\tau^{\tilde{m}} = G_\tau^{\tilde{m}} - A_\tau$ and $D_\tau^{\tilde{m}} = H_\tau^{\tilde{m}} - B_\tau$, A_τ and B_τ solve Equations (9) and (10)), and $G_\tau^{\tilde{m}}$ and $H_\tau^{\tilde{m}}$ solve the recursions

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

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

with initial conditions $G_0^{\tilde{m}} = \gamma_0$ and $H_0^{\tilde{m}} = \gamma_1^{\tilde{m}}$ for $\tilde{m} = 1, 2, \cdots, 12$.[6,7]

3.4 | Risk premia in commodity futures

In this section we express different notions of risk premia in terms of the components of the affine model. Since all the strategies that we consider cost zero when they are entered into, ex ante expected return entirely reflects expected risk premia.

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 $f_{t+1}^{(\tau-1)} - f_t^{(\tau)}$. The time t conditional expectation of this strategy is

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

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 the particular case of the expected return (Equation 17) evaluated at $\tau = 1$:

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

where we use that a 0-period futures is equivalent to the spot price, $s_{t+1} = f_{t+1}^{(0)}$.

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

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

---

[6]Alquist, Bauer, and Diez de los Rios (2013) study an affine model of oil futures using a setup different from ours. While we price commodity futures by discounting dollar cash flows—as usually done in the literature—they 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. Moreover, they do not consider seasonal fluctuations.

[7]Some authors claim that no-arbitrage restrictions only matter for the cross-section of bond yields and that they do not affect the time series dynamics of the factors (Duffee, 2011). Nevertheless, our model has several features such as unobserved factors, stochastic seasonality, and restrictions in the physical and risk-neutral measures that, when considered together, should be addressed within an arbitrage-free model.

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:

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

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

$$E_t[f_{t+h}^{(\tau-h)} - f_t^{(\tau)}] = \sum_{i=1}^{h} J_{\tau-i}^{m_{t+i}} + \sum_{i=1}^{h} D_{\tau-i}^{m_{t+i}\prime} \Gamma E_t[\Lambda_{t+i-1}]. \tag{20}$$

4 | A REPRESENTATION OF THE AFFINE MODEL OF COMMODITY FUTURES

Here we extend results in Christensen et al. (2011) to the commodity futures literature by showing that a single stochastic discount factor can be used to simultaneously price bonds and commodity futures displaying stochastic seasonality in such a way that the yield curve and the cost-of-carry curve adopt augmented Nelson and Siegel functional forms. This representation of the model is simple, yet flexible enough to match the different shapes of the cost-of-carry and yield curves.

We first write the log-basis (Equation 3) emphasizing the contribution of the seasonal factors as

$$f_t^{(\tau)} = \beta_{0t} + \tau(y_t^{(\tau)} + \tilde{u}_t^{(\tau)}) + e^{-\omega\tau}\left[\xi_t \cos\left(\frac{2\pi}{12}m_{t+\tau}\right) + \xi_t^* \sin\left(\frac{2\pi}{12}m_{t+\tau}\right)\right], \tag{21}$$

where we interpret β_{0t} as the deseasonalized spot commodity factor and $\tilde{u}_t^{(\tau)}$ as the cost-of-carry curve net of any *stochastic* seasonal component.[8] The last term on the right-hand side reflects the contribution of the seasonal factors to futures prices of different maturities. When $\tau = 0$ the futures price is the spot commodity price and Equation 21 becomes

$$s_t = \beta_{0t} + \xi_t \cos\left(\frac{2\pi}{12}m_t\right) + \xi_t^* \sin\left(\frac{2\pi}{12}m_t\right). \tag{22}$$

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 $e^{-\omega\tau}$.

There is agreement in the bond pricing literature that three factors are sufficient to summarize the evolution of the yield curve over time. Thus we parametrize the yield curve $y_t^{(\tau)}$ using a dynamic Nelson and Siegel model as in Christensen et al. (2011):

$$y_t^{(\tau)} = a_\tau + \delta_{1t} + \left(\frac{1 - e^{-\zeta_1\tau}}{\zeta_1\tau}\right)\delta_{2t} + \left(\frac{1 - e^{-\zeta_1\tau}}{\zeta_1\tau} - e^{-\zeta_1\tau}\right)\delta_{3t}, \tag{23}$$

where δ_{1t}, δ_{2t}, and δ_{3t} are latent variables interpreted as level, slope, and curvature factors, and the parameter ζ_1 determines the shape of the loadings on the factors δ_{2t} and δ_{3t}.

Section 2 shows that we need at least three factors to properly characterize the nonseasonal component of the cost-of-carry curve $\tilde{u}_t^{(\tau)}$. Therefore, we also look for conditions under which the cost-of-carry curve adopts an augmented dynamic Nelson and Siegel form:

$$\tilde{u}_t^{(\tau)} = g_\tau^{m_t} + \beta_{1t} + \left(\frac{1 - e^{-\zeta_2\tau}}{\zeta_2\tau}\right)\beta_{2t} + \left(\frac{1 - e^{-\zeta_2\tau}}{\zeta_2\tau} - e^{-\zeta_2\tau}\right)\beta_{3t}, \tag{24}$$

where β_{1t}, β_{2t}, and β_{3t} are level, slope, and curvature factors. Even though $\tilde{u}_t^{(\tau)}$ is independent of seasonal shocks, rendering the Nelson and Siegel parametrization arbitrage free still requires the term $g_\tau^{m_t}$ to depend on the season (month) m_t, although in a deterministic fashion.

We now show that Nelson and Siegel parametrization belongs to the family of affine models described in Section 3. Although it is known that one can impose restrictions on an affine model of bond prices to obtain a Nelson and Siegel representation of the yield curve (Christensen et al., 2011), the extension to the commodity futures literature has not been proved so far.

---

[8]The Supporting Information shows that the pair of seasonal factors associated with the fundamental frequency is enough to characterize seasonality in heating oil and soybean futures. Thus from now on we assume that seasonal fluctuations follow the process of Equation 2.

The vector of risk factors is $X_t = [\delta_{1t}, \delta_{2t}, \delta_{3t}, \beta_{0t}, \beta_{1t}, \beta_{2t}, \beta_{3t}, \xi_t, \xi_t^*]'$. Our task is to find parameters $\lambda_0, \lambda_1, \rho_0, \rho_1, \gamma_0,$ and $\gamma_1^{m_t}$ such that the yield curve and the cost-of-carry curve adopt the functional forms of Equations 23 and 24. These conditions are summarized in the following proposition.

Proposition 2. *(Nelson and Siegel representation)* *Consider parameters of the risk-neutral measure:*

$$
\Theta^Q = \begin{bmatrix}
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & e^{-\zeta_1} & \zeta_1 e^{-\zeta_1} & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & e^{-\zeta_1} & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & \frac{1-e^{-\zeta_1}}{\zeta_1} & \frac{1-e^{-\zeta_1}}{\zeta_1} - e^{-\zeta_1} & 1 & 1 & \frac{1-e^{-\zeta_2}}{\zeta_2} & \frac{1-e^{-\zeta_2}}{\zeta_2} - e^{-\zeta_2} & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & e^{-\zeta_2} & \zeta_2 e^{-\zeta_2} & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & e^{-\zeta_2} & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & e^{-\omega} & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & e^{-\omega}
\end{bmatrix}
$$

and $\mu^Q \in R^n$, where $\zeta_1, \zeta_2, \omega > 0$. Let $\lambda_0 = \Gamma^{-1}(\mu - \mu^Q)$, $\lambda_1 = \Gamma^{-1}(\Theta - \Theta^Q)$, $\rho_0 = 0$, $\gamma_0 = 0$,

$$
\rho_1 = \left[1, \frac{1-e^{-\zeta_1}}{\zeta_1}, \frac{1-e^{-\zeta_1}}{\zeta_1} - e^{-\zeta_1}, 0, 0, 0, 0, 0, 0\right]',
$$

$$
and \quad \tilde{\gamma}_1 = \left[0, 0, 0, 1, 0, 0, 0, \cos(\frac{2\pi}{12}\tilde{m}), \sin(\frac{2\pi}{12}\tilde{m})\right]'
$$

for $\tilde{m} = 1, 2, \cdots, 12$. Then, the yields and futures prices adopt the parametrization of Equations 21, 23, and 24.

5 | ESTIMATION METHOD

We estimate the model by maximum likelihood using the Kalman filter to evaluate the prediction error decomposition of the likelihood function and to handle missing observations, a common feature in the market of commodity futures. The state variables evolve as in Equation 5, and the observation equation consists of futures prices and bond yields that satisfy Equations 11 and 14 augmented with measurement errors $\varepsilon_{yt}^{(\tau_y)}$ and $\varepsilon_{ft}^{(\tau_f)}$:

$$
y_t^{(\tau_y)} = a_{\tau_y} + b_{\tau_y}' X_t + \varepsilon_{yt}^{(\tau_y)}, \tag{25}
$$

$$
f_t^{(\tau_f)} = C_{\tau_f}^{m_t} + D_{\tau_f}^{m_t'} X_t + \varepsilon_{ft}^{(\tau_f)} \tag{26}
$$

Bond yields are observed for a set of maturities $\tau_y \in \mathcal{T}_y$, and futures prices may be observed for a different set of maturities $\tau_f \in \mathcal{T}_f$. We use bond yields for maturities up to 5 years to properly estimate the shape of the yield curve and futures contracts with maturities from 1 to 18 months.

The intercept and factor loadings satisfy the functional forms of Equations 21, 23, and 24. Namely, $C_\tau^{m_t} = G_\tau^{m_t} - A_\tau$, $D_\tau^{m_t} = H_\tau^{m_t} - B_\tau$, $a_\tau = -A_\tau/\tau$, $b_\tau = -B_\tau/\tau$,

$$
H_\tau^{m_t} = \left[0, 0, 0, 1, \tau, \frac{1-e^{-\zeta_2\tau}}{\zeta_2}, \frac{1-e^{-\zeta_2\tau}}{\zeta_2} - \tau e^{-\zeta_2\tau}, e^{-\omega\tau}\cos(\frac{2\pi}{12}(m_t+\tau)), e^{-\omega\tau}\sin(\frac{2\pi}{12}(m_t+\tau))\right]',
$$

$$
B_\tau = \left[-\tau, -\frac{1-e^{-\zeta_1\tau}}{\zeta_1}, -\left(\frac{1-e^{-\zeta_1\tau}}{\zeta_1} - \tau e^{-\zeta_1\tau}\right), 0, 0, 0, 0, 0, 0\right]',
$$

and A_τ and $G_\tau^{m_t}$ satisfy Equations 9 and 15. In addition, stochastic seasonality enters into Equation 26 as a periodic loading on the factors ξ_t and ξ_t^*.[9]

---

[9]Since $\mu^{Q'}D_{\tau-1}$ is a scalar, we can only identify a single parameter in $\mu^{Q'}$. We thus set $\mu^Q = [\mu_1^Q, 0, 0, 0, 0, 0, 0, 0, 0]'$. Also, note that ζ_1 and ζ_2 are identified because ζ_2 appears only in Equation 26 whereas ζ_1 appears in both Equations 25 and 26.

TABLE 3 Estimates of the model with stochastic seasonality

Parameters of the VAR(1) process for the yield, the spot and cost-of-carry factors

$\mu_\delta(\times 1000)$ $\Theta_{\delta\delta}$ $\Gamma_{\delta\delta}(\times 1000)$

$$
\begin{bmatrix} 0.1305 \\ (0.048) \\ 0.0064 \\ (0.0593) \\ 0.0255 \\ (0.1072) \end{bmatrix}
\begin{bmatrix} 0.964 & 0.006 & 0.026 \\ (0.012) & (0.019) & (0.011) \\ -0.005 & 0.945 & 0.031 \\ (0.014) & (0.021) & (0.014) \\ -0.004 & 0.016 & 0.936 \\ (0.023) & (0.028) & (0.016) \end{bmatrix}
\begin{bmatrix} 0.313 & 0 & 0 \\ (0.011) & & \\ -0.267 & 0.218 & 0 \\ (0.014) & (0.008) & \\ -0.167 & -0.027 & 0.646 \\ (0.021) & (0.026) & (0.022) \end{bmatrix}
$$

$\mu_\beta(\times 100)$ $\Theta_{\beta\beta}$ $\Gamma_{\beta\beta}(\times 100)$

$$
\begin{bmatrix} -0.0134 \\ (0.9908) \\ -0.0654 \\ (0.0229) \\ -0.01864 \\ (0.4807) \\ 0.5343 \\ (0.3542) \end{bmatrix}
\begin{bmatrix} 0.9908 & 0.1691 & 0.8391 & -0.0085 \\ (0.012) & (1.045) & (0.199) & (0.242) \\ 0.0004 & 0.8876 & 0.0185 & 0.0238 \\ (0.0002) & (0.025) & (0.005) & (0.007) \\ 0.0002 & -0.8541 & 0.8254 & 0.5384 \\ (0.0044) & (0.555) & (0.072) & (0.102) \\ 0.0007 & 1.2311 & -0.2151 & 0.3386 \\ (0.0028) & (0.428) & (0.067) & (0.091) \end{bmatrix}
\begin{bmatrix} 9.634 & 0 & 0 & 0 \\ (0.04) & & & \\ -0.141 & 0.139 & 0 & 0 \\ (0.012) & (0.009) & & \\ -0.971 & -0.004 & 2.915 & 0 \\ (0.197) & (0.309) & (0.323) & \\ -0.143 & -0.488 & -2.006 & 1.138 \\ (0.175) & (0.287) & (0.255) & (0.069) \end{bmatrix}
$$

Volatility of seasonal process

$\sigma_\xi = 0.0029 (0.0002) \sigma_{\xi^*} = 0.0017 (0.0001)$

Other parameters and log-likelihood

$\zeta_1 = 0.053 (0.0006) \zeta_2 = 0.416 (0.021) \omega = 0.0096 (0.0014)$ Log-likelihood = 51270.95

6 │ EMPIRICAL RESULTS

We estimate two versions of the model: one with stochastic seasonaliy (SS) and one with deterministic seasonality (DS).[10] In this section we focus on the case of heating oil futures. (The Supporting Information shows results for soybean futures.)

Table 3 reports the estimates of the model with stochastic seasonality.[11] The Akaike information criterion favors the model with stochastic seasonality (−102,383 vs. −97,850), and the Scharwz and Hannan–Quinn criteria give similar results. The estimated commodity factors in the SS model are radically different from those in the DS model (Figure 1), with the cost-of-carry factors showing clear signs of seasonality in the latter but not in the former.[12] Furthermore, the SS model also dominates the DS model in terms of pricing errors (Table 4).

Figure 2 (top panel) displays the estimated seasonal components. The seasonal peak is attained at the beginning of the winter season and the trough, at the beginning of the summer season. The peaks and troughs of the seasonal pattern in the SS model change over time (possibly depending on the severity of the winter season) and there is a drop in seasonal fluctuations at the end of the sample. This drop may be due to a change in the composition of demand: While the residential use of heating oil has decreased (which is mostly seasonal) and exports increased, its use as transportation fuel has increased over time (a mostly nonseasonal use).[13] Moreover, this finding is also consistent with anecdotal evidence suggesting a vanishing seasonality in oil prices.[14] As a result, the DS model underestimates the seasonality at the beginning of the sample and overestimates it at the end of it.

The model with deterministic seasonality also distorts the estimated factors (Figure 1). The DS model is unable capture all seasonal fluctuations from futures prices and erroneously attributes those cycles to fluctuations in the cost-of-carry factors. Moreover, since β_{1t} and β_{2t} determine the variations of the cost-of-carry curve at the short and long end of the curve, spurious fluctuations in these factors lead to spurious fluctuations in estimated risk premia.

The shape of the cost-of-carry curve depends on the parameter ζ_2, which is more than twice as large in the model with stochastic seasonality (0.416 versus 0.197). The parameter ζ_2 also determines the evolution of the risk factors under the

---

[10]In the model with deterministic seasonality we set $\sigma_\xi = \sigma_{\xi^*} = 0$ and estimate ξ_0 and ξ_0^* as free parameters.

[11]The Supporting Information contains the estimates of the model with deterministic seasonality.

[12]The estimated matrix $\widehat{\Theta}_{\beta\beta}$ in the DS model contains complex roots corresponding to a cycle of about 12 months.

[13]According to EIA (2017), the share of heating oil consumption for residential use has decreased by more than two-thirds from the beginning to the end of the sample, whereas the share of transportation distillate use increased by over 50% over the same period. Moreover, the US distillate exports grew to 1.5 million barrels per day in 2017, whereas they had been stable at around 0.2 million barrels per day from the beginning of our sample up to the end of 2005.

[14]For instance, the US Energy Information Administration (EIA, 2013) notes that "looking at data for the last 13 years, it is apparent that the traditional northern hemisphere winter spike in demand [for oil] has become increasingly less pronounced."
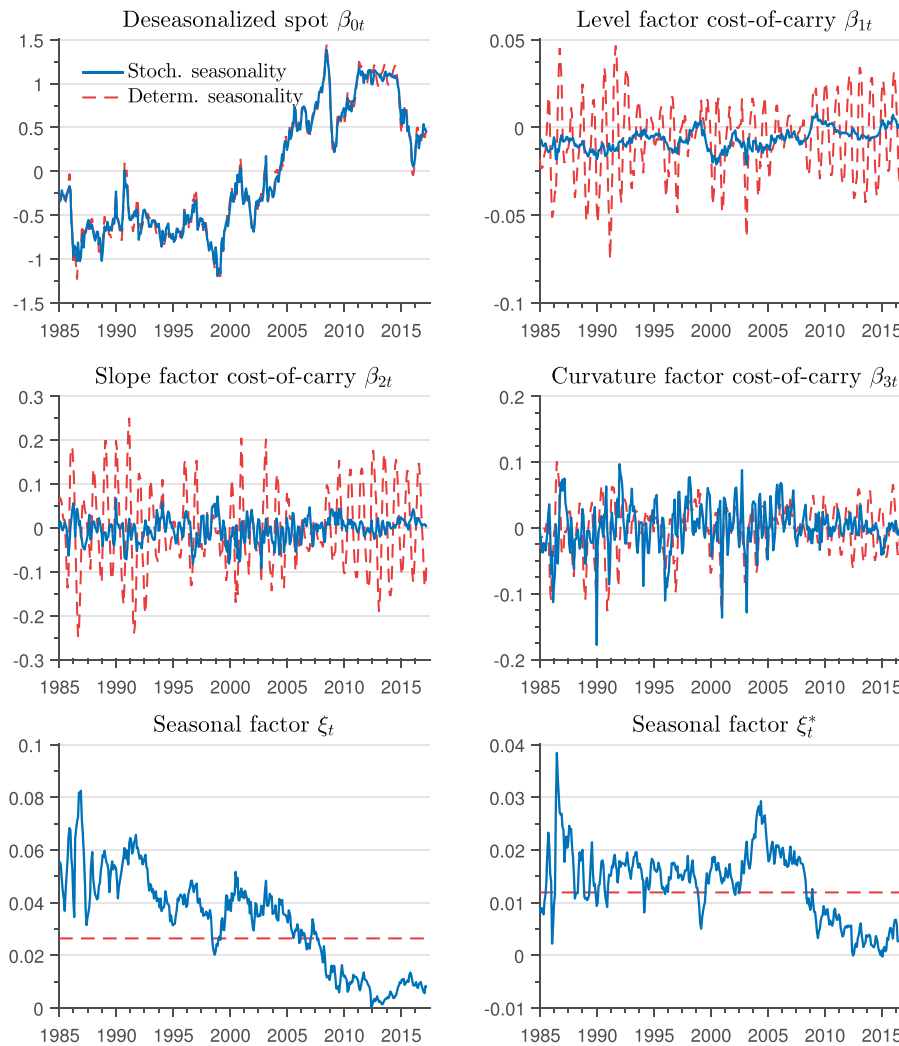
FIGURE 1 Spot, cost-of-carry, and seasonal factors. *Notes*: Estimates of commodity factors of the models with stochastic and deterministic seasonality: deseasonalized spot β_{0t}; factors of the cost-of-carry, β_{1t}, β_{2t}, and $\beta3_{3t}$; and seasonal factors ξ_t and ξ_t^* [Colour figure can be viewed at wileyonlinelibrary.com]

risk-neutral measure and, therefore, the associated market prices of risk. The lower is ζ_2, the higher is the persistence of the level and slope factors β_{1t} and β_{2t} under the risk-neutral measure, and the lower is the impact of the past curvature β_{3t} on the current slope β_{2t}. The estimated ζ_2 in the model with deterministic seasonality implies more persistent factors and a weaker relation between the lagged curvature and the current slope. Therefore, the seasonal cycles leaked into the estimated cost-of-carry factors affect the dynamics of risk premia.

Finally, the model with stochastic seasonality is able to match the different shapes of the futures curve observed in our sample (Figure 3). On some dates, it is not even clear whether the futures curve is upward sloping or inverted unless one strips out the seasonal component. Furthermore, the amplitude of the seasonal factor depends on the particular dates that we choose. This observation is inconsistent with the predictions of a model with deterministic seasonality.

6.1 | Seasonality and the theory of storage

It is of interest to determine the source of the moderation of the seasonal component. The theory of storage relates the stock of inventories to the cost-of-carry.[15] We compare the seasonal component extracted from data on heating oil

---

[15]In the traditional model, the convenience yield measures the option value of holding inventories when there is a positive probability of stock-outs (Deaton & Laroque, 1992; Gorton, Hayashi, & Rouwenhorst, 2013; Routledge et al., 2000). However, the relation between inventories and convenience yields also exists without stock-outs but allowing for other frictions in the production and storage process, such as adjustment costs (Evans & Guthrie, 2017). Since we do not aim to identify the precise interpretation of this relation, any of those models could be consistent with the observed relation.

TABLE 4 Pricing errors

	RMSPE		MAPE	
Maturity	**SS**	**DS**	**SS**	**DS**
1	0.673	1.996	6.847	12.659
3	0.210	0.593	3.905	6.760
6	0.416	0.582	5.696	6.759
9	0.431	0.648	5.750	7.229
12	0.364	0.691	5.171	7.446
16	0.245	0.825	4.391	8.183

Note. The table compares the pricing errors in the models with stochastic (SS) and deterministic (DS) seasonality. RMSPE is the root mean square pricing error and MAPE is the mean absolute pricing error. All entries are multiplied by 100.
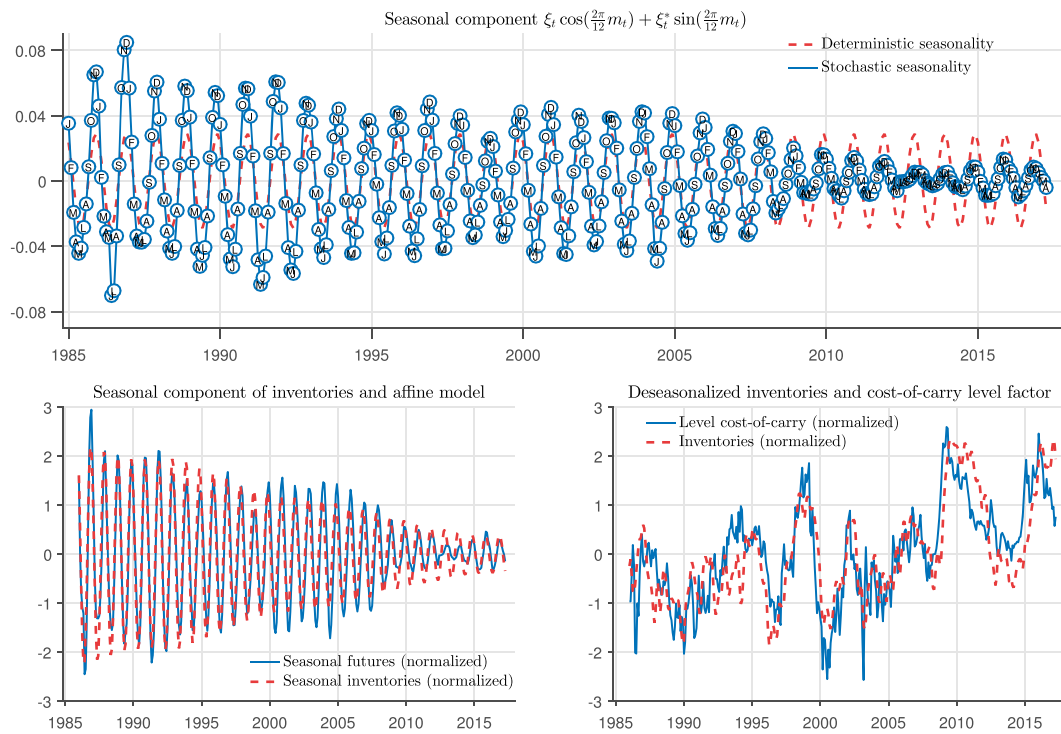
FIGURE 2 Seasonal component and heating oil inventories. _Notes_: The upper panel displays the implied seasonal component $\xi_t \cos(\frac{2\pi}{12} m_t) + \xi_t^* \sin(\frac{2\pi}{12} m_t)$ in the models with stochastic and deterministic seasonality. The figure also displays the months in circles in the model with stochastic seasonality. Most seasonal peaks are in December (D) and troughs in June (J). The bottom left panel displays the seasonal component of log-inventories and the implied seasonal component of the spot commodity price, both standardized. The bottom right panel shows the cost-of-carry level factor β_{1t} and the nonseasonal component of log-inventories, both standardized [Colour figure can be viewed at wileyonlinelibrary.com]

inventories with the estimated seasonal component of the affine model. The bottom left panel of Figure 2 shows a striking similarity between the two series: the peaks and troughs are well aligned and the 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. Furthermore, also consistent with the theory of storage, the estimated level factor of the cost-of-carry β_{1t} is highly correlated with the nonseasonal component of the stock of heating oil inventories (bottom right panel of Figure 2).
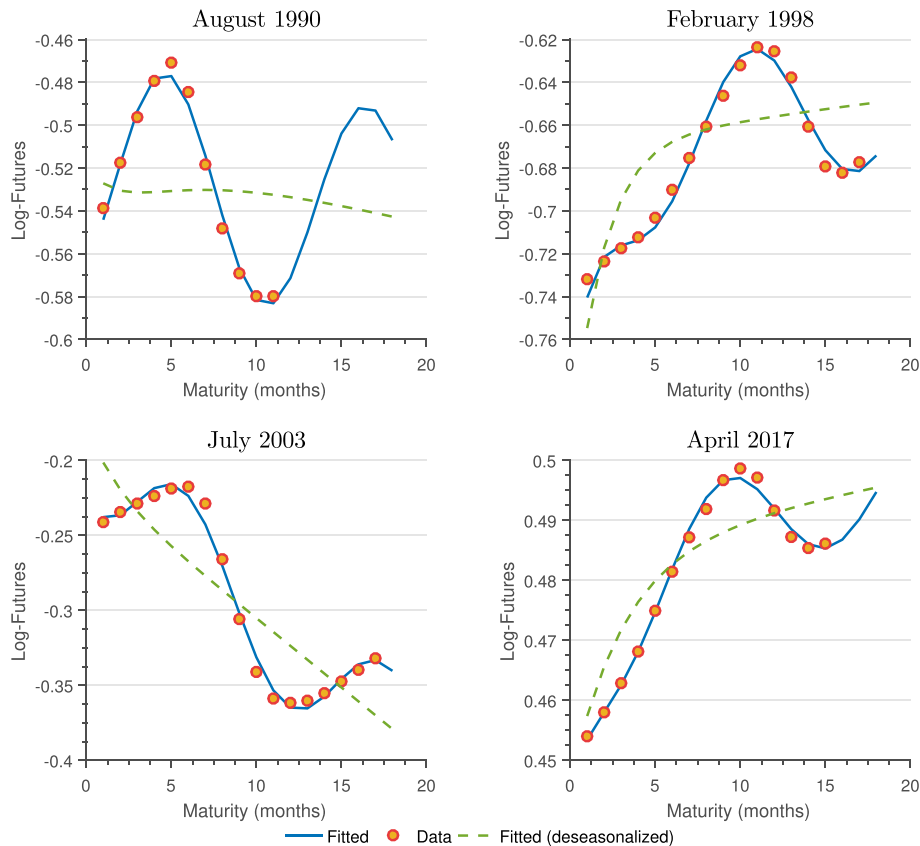
FIGURE 3 Fitted log-futures curves. *Notes*: The figure shows fitted log-futures curves, deseasonalized fitted log-futures curve, and actual log-futures prices on selected dates [Colour figure can be viewed at wileyonlinelibrary.com]

6.2 | Risk premia in heating oil futures markets

Since entering into a futures contract costs zero, any expected return is a risk premium. Here we analyze the risk associated with holding futures contracts for one period and that associated with holding the contract to maturity. We find that risk premia in heating oil futures markets declined over time and that it has been positive on average until 2005 and negative afterwards. The change in the risk premium coincides with an abrupt increase in heating oil spot prices and with an increased participation by financial investors in commodity futures markets.[16] In addition, we find that the contribution of the spot, cost-of-carry, and yield curve factors to the time variation of risk premia is substantial, while seasonal factors play a modest role.

Figure 4 shows the 1-month expected holding return for futures contracts that mature in 1 month and 18 months ahead. Commodity factors (spot and cost-of-carry) explain most of 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 importance of the yield curve factors. The contribution of the seasonal shocks is about 0.6 percentage points at the beginning of the sample and becomes smaller as the seasonal component gets smaller over time (Figure 2).

The estimated loading of expected returns on the yield curve level factor (δ_{1t}) is negative and its importance increases with the maturity of the contract. Furthermore, the loading on the slope factor (δ_{2t}) is positive when the yield curve is upward sloping, negative when the yield curve is inverted, and its importance decreases for longer maturity contracts. The upper right panel of Figure 4 is consistent with those observations. The contribution of the yield curve factors to the risk premium of the 18-month futures contract tracks closely the negative of the level factor (δ_{1t}), whereas that of the 1-month contract is smoother and highly correlated with the slope factor (δ_{2t}).

---

[16]The increased entry of noncommercial traders coincides with an increase in open interest and trading volume over the last decade. The EIA reports that the average daily volume of futures contracts on petroleum products was four times larger in 2017 than the 2000–2006 average (see https://www.eia.gov/finance/markets/products/financial_markets.php). Hamilton and Wu (2014) document a similar shift in crude oil risk premia in 2005 and also attribute it to the entry of noncommercial traders into that market.
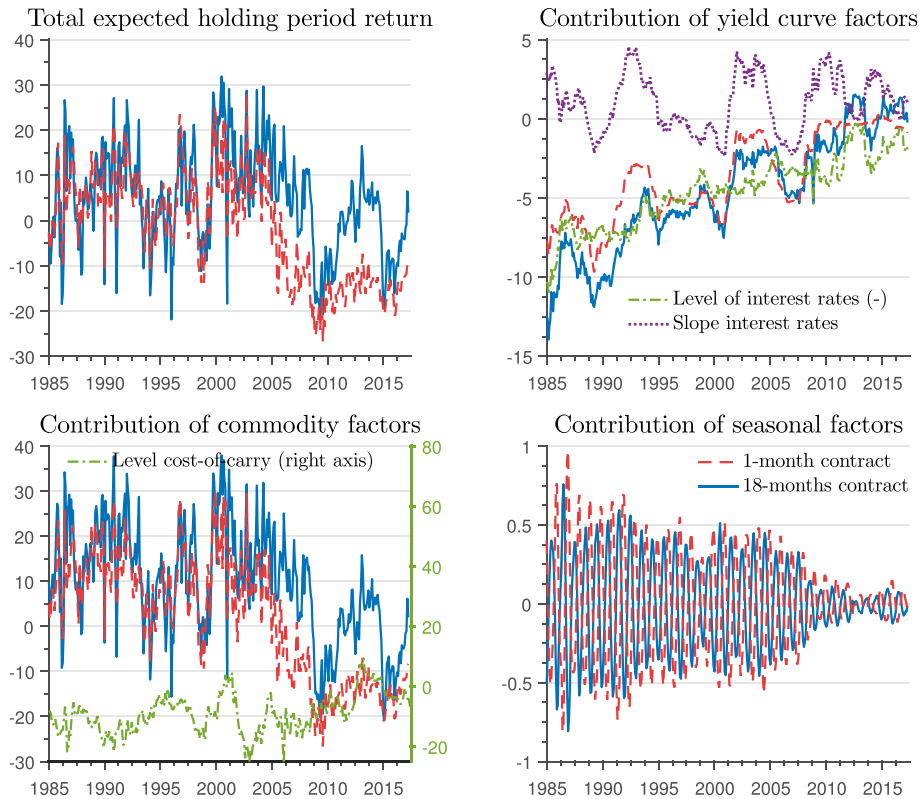
FIGURE 4 Expected 1-month holding futures returns. *Notes*: This figures shows the expected 1-month holding returns of a 1-month and a 18-month futures contract. Returns are expressed in percentage points and on an annualized basis. The upper left panel displays the total expected return. The other panels 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 [Colour figure can be viewed at wileyonlinelibrary.com]

The bottom left panel of the figure shows that commodity factors are very volatile and that their impact on risk premia is large. It shows that the importance of the cost-of-carry level factor (β_{1t}) increases with the maturity of the contract and that this factor is negatively correlated with the component of expected return accounted for by the commodity factors (−0.82 and −0.76 for the 18-month and 1-month contracts, respectively).

Figure 5 displays the expected risk premia of holding an 18-month futures contract for 1 and 18 months. As the holding period increases, the owner of the contract is exposed to longer-term risks. As a result, the spot commodity factor (β_{0t}) and the slope factor of the yield curve (δ_{2t}) 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 3). Therefore, expected returns of longer-term contracts increase.

The upper right panel of Figure 5 shows that the contribution of the slope yield curve factor increases with the holding period: Long holding returns tend to follow more closely the slope of the yield curve than short holding returns. Moreover, periods with inverted yield curves are roughly associated with higher total expected risk premia (top panels). In addition, the bottom left panel shows that the contribution of the commodity factors tends to follow the short- and medium-term movements of the spot commodity factor. Interestingly, the longer-term movements are negatively correlated: Periods when the spot factor is relatively high are periods when expected risk premia are negative.

6.3 | The zero lower bound on interest rates

From December 2008 to December 2015, the US Federal Reserve set the policy interest rate to 0.25 percentage points, virtually hitting the zero lower bound (ZLB) on interest rates. If one imposes the ZLB constraint, the model ceases to be affine in the risk factor, which causes analytical and econometric challenges. In this section we explore to what extent our results are distorted by ignoring the ZLB on interest rates.
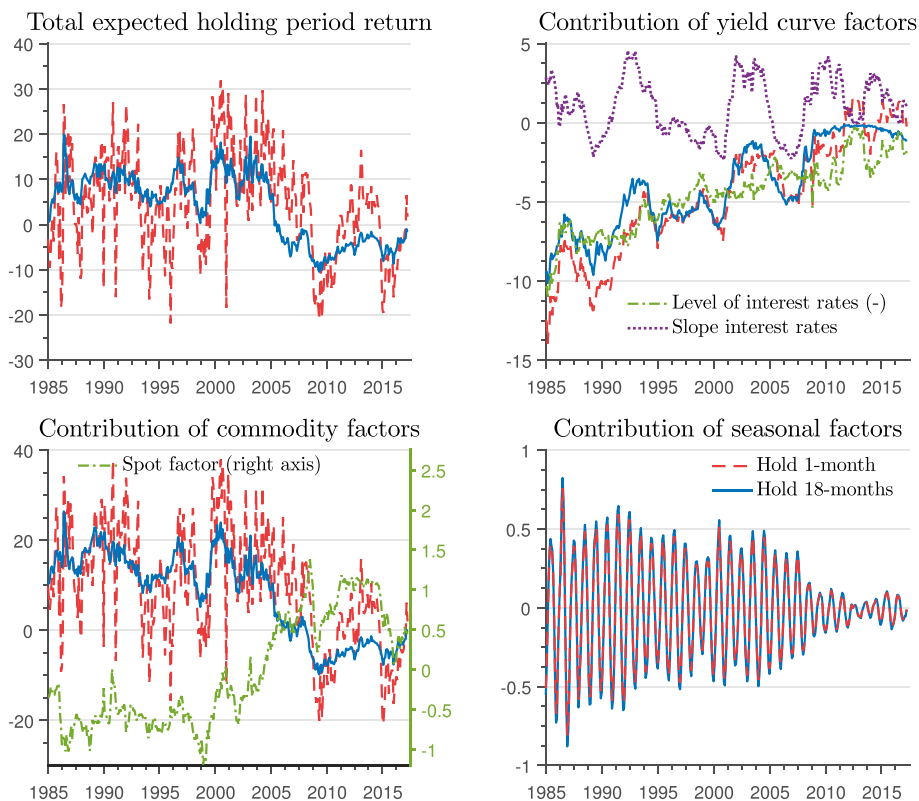
FIGURE 5 Expected holding futures returns of an 18-month futures contract. *Notes*: Expected 1-month and 18-month holding returns of an 18-month futures contract. Returns are expressed in percentage points and on an annualized basis. The upper left panel displays the total expected return. The other 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 [Colour figure can be viewed at wileyonlinelibrary.com]

The model is similar to that described in Section 3 but replacing the interest rate process (Equation 7) with an affine shadow interest rate of the form

$$z_t = \rho_0 + \rho_1' X_t.$$

The short-term interest rate is the maximum of the shadow rate z_t and a lower bound $\underline{r} = 0$:

$$r_t = \max(z_t, \underline{r}).$$

The stochastic discount factor and the market prices of risk are still given by Equations 5 and 6, and the parameters λ_0 and λ_1 still satisfy the restrictions imposed in Proposition 2.

This modification renders the model nonlinear. While we could approximate the solution using numerical techniques, estimating the parameters becomes challenging because each evaluation of the likelihood function requires numerically solving the model and performing a nonlinear filtering procedure. Instead, we adapt the methodology proposed by Wu and Xia (2016) and use an approximate solution whereby the yield on a τ-period zero coupon bond can be written as

$$y_t^{(\tau)} = \underline{r} + \frac{1}{\tau} \sum_{j=0}^{\tau} \sigma_j^Q g\left(\frac{\tilde{a}_j + \tilde{b}_j' X_t - \underline{r}}{\sigma_j^Q} \right). \tag{27}$$

Here, $\tilde{a}_j = \rho_0 + \rho_1'\left(\sum_{i=0}^{j-1} \Theta^{Qi}\right)\mu^Q - 0.5\rho_1'\left(\sum_{i=0}^{j-1} \Theta^{Qi}\right)\Gamma\Gamma'\left(\sum_{i=0}^{j-1} \Theta^{Qi}\right)'\rho_1$; $\tilde{b}_j' = \rho_1'(\Theta^Q)^j$; $\sigma_j^Q = \left[\rho_1'\left(\sum_{i=0}^{j-1} (\Theta^{Qi})\Gamma\Gamma'(\Theta^{Qi})'\right)\rho_1\right]^{1/2}$; and $g(x) = x\Phi(x) + \phi(x)$, where Φ is the standard normal cumulative distribution and ϕ is the associated density.[17] The function $g(x)$ is nonnegative, increasing, approaches zero as x decreases, and approaches $g(x) = x$ as x increases.

---

[17]The Supporting Information contains a detailed description of the model and the estimation procedure. Wu and Xia (2016) actually solve the model in terms of one-period forward interest rates. To make it comparable with the previous section, we chose to compute bond yields by adding one-period forward interest rates.

Let $y_t^{(\tau)} = \mathcal{Y}^{(\tau)}(X_t)$, where $\mathcal{Y}^{(\tau)}(X_t)$ is defined as the right-hand side of Equation 27. With this notation, the observation equations of the state-space system are

$$
\begin{aligned}
y_t^{(\tau_y)} &= \mathcal{Y}^{(\tau_y)}(X_t) + \varepsilon_{yt}^{(\tau_y)}, \\
f_t^{(\tau_f)} &= \beta_{0t} + \tau_f \left(\mathcal{Y}^{(\tau_f)}(X_t) + \tilde{u}_t^{(\tau_f)} \right) + e^{-\omega\tau} \left[\xi_t \cos(\frac{2\pi}{12} m_{t+\tau}) + \xi_t^* \sin(\frac{2\pi}{12} m_{t+\tau}) \right] + \varepsilon_{ft}^{(\tau_f)},
\end{aligned}
\tag{28}
$$

where $\varepsilon_{yt}^{(\tau_y)}$ and $\varepsilon_{ft}^{(\tau_f)}$ are measurement errors, $\tau_y \in \mathcal{T}_y$, $\tau_f \in \mathcal{T}_f$, and $\tilde{u}_t^{(\tau_f)}$ is given by Equation 24. As before, the state variables evolve according to Equation 5. To estimate the parameters of the model, we use the the extended Kalman filter to evaluate the likelihood function. Details of the algorithm and the estimated parameters are described in the Supporting Information.

Figure 6 compares results from the affine model with those of the model that imposes the ZLB on interest rates. The upper left panel displays the short-term interest rate r_t in the baseline model and the shadow interest rate imposing the ZLB. Both lines are almost identical except between mid-2013 and the end of 2015, when the shadow rate becomes negative while the short rate in the baseline model is roughly zero. The upper right and middle left panels compare the level and slope yield curve factors (δ_{1t} and δ_{2t}). Again, they are very similar even during the zero lower bound period. The curvature factor (δ_{3t}) differs between models but it only explains a negligible fraction of the variation in bond yield and risk premia. The level factors of the cost-of-carry (β_{1t}) are indistinguishable even during the zero lower bound period
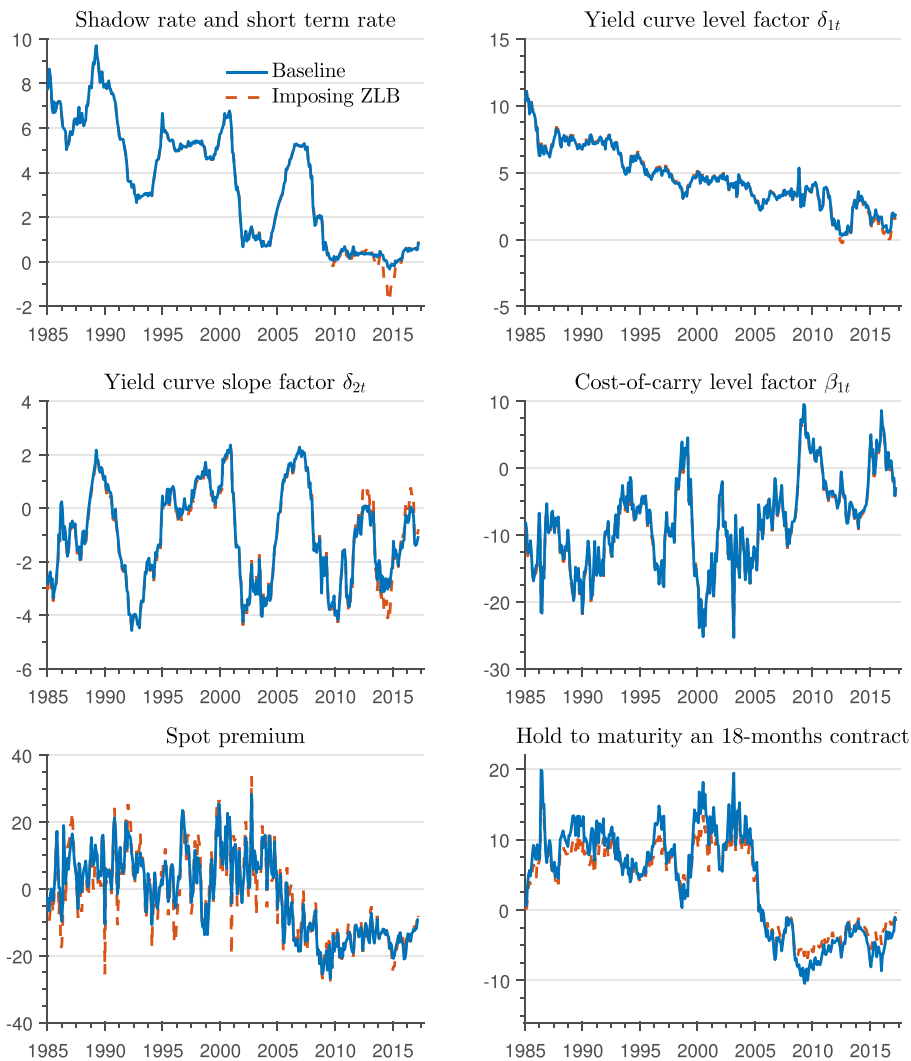

FIGURE 6 Comparing models with and without imposing the zero lower bound on interest rates. *Notes*: The figure compares some results for the baseline model (solid lines) with those estimated imposing the zero lower bound (dashed lines) on interest rates (ZLB) [Colour figure can be viewed at wileyonlinelibrary.com]

(middle right panel). Although not shown in the figure, the same pattern is true for the other cost-of-carry factors (β_{2t} and β_{3t}), the spot factor β_{0t}, and the seasonal factors ξ_t and ξ_t^*. Lastly, the bottom panels display the spot premium and the expected return from holding an 18-month contract to maturity. Both measures of risk premia are very similar between models and their pairwise correlation is 0.97 for the former and over 0.99 for the latter.

From these results we conclude that the baseline model performs very well even during the periods in which the interest rates were at the zero lower bound. The commodity factors, which are the main drivers of risk premia, are virtually identical in both models. This is because the periods in which interest rate factors differ the most (when the ZLB is binding) are precisely those periods in which changes in interest rates have the lowest impact on futures prices (Equation 28).

6.4 | Decomposing risk premia during the 1997–1998 oil price shock

We now look at the evolution of the futures curve, and the spot and term premia around the 1997–1998 oil price shock. This was a period with large movements in commodity prices, which is particularly interesting because the futures curve was in backwardation before the 1997 peak and in contango at the 1998 trough. Falling oil prices were associated with the emergence of a contango in the futures market, which led to the usual time spread strategies (long a futures contract and short a contract with a different maturity).

To assess the risk inherent in time spread strategies, it is useful to decompose the expected 1-month holding return of a τ-period contract into a spot premium (expected return of holding a 1-month contract) and a term premium (expected 1-month return of a τ-period contract in excess of that of a 1-month contract):

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

Heating oil prices reached a peak in January 1997 and the associated spot premium was large and positive (Figure 7). The reason is that periods of backwardation are usually associated with a drop in stocks of heating oil inventories and
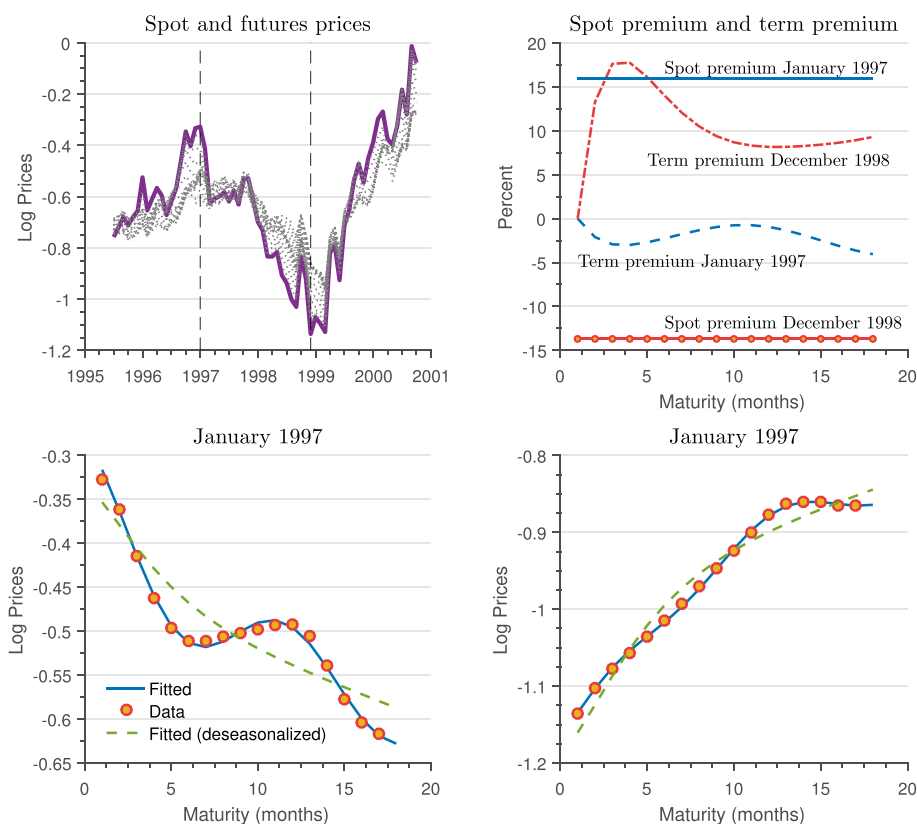

FIGURE 7 Spot and term premium: a case study. *Notes*: The upper left panel displays the spot price (solid line) and all futures prices (dotted lines). The upper right panel displays the spot and term premium during the peak and trough months of heating oil prices. The bottom panels display the fitted and actual futures curves on the same dates [Colour figure can be viewed at wileyonlinelibrary.com]

increases in the spot price. In the 1998 trough, the spot premium was negative and large, consistent with the futures curve being in contango, accumulation of inventories, and an expected fall in the spot price.

The risk inherent in the abnormally high commodity price of 1997 was reflected in a negative and decreasing term premium (top left panel). As prices were too high, investors demanded a lower premium to hold long contracts because they expected an eventual drop in the price and hence were willing to accept a negative premium. In effect, the downward-sloping futures curve of January 1997 reflected the expectation of lower spot price in the future. The futures curve switched from backwardation to contango in April 1997 while at the 1998 trough, as heating oil prices were expected to recover, the term premium turned positive and hump-shaped as investors demanded a positive premium for holding long-dated contracts.

7 | CONCLUDING REMARKS

In this paper we document that seasonal fluctuations in futures contracts for a group of energy and agricultural commodities are stochastic. In addition, we show evidence that properly accounting for the variability in the cost-of-carry curve for the same group of commodities requires at least three factors. These observations motivate the main contribution of the paper: developing and estimating a multifactor affine model of commodity futures and bond prices that allows for stochastic variations in seasonality. We show that there is single pricing kernel that prices bonds and commodity futures in such a way that the yield curve and the cost-of-carry curve adopt Nelson–Siegel functional forms.

We estimate the model using data on heating oil futures and US bond prices. We observe a decline in the amplitude of seasonal fluctuations in the futures prices that coincides with a similar moderation in the amplitude of the seasonal pattern of stocks of heating oil inventories. This observation is consistent with the theory of storage. Next, we measure the contribution of the different factors to risk premia. We find that most high-frequency fluctuations in risk premia are due to variations in the spot price factor and other factors associated with the cost-of-carry curve. Correctly specifying seasonality in futures prices as stochastic is important mostly to avoid erroneously assigning those fluctuations to other risk factors. In addition, in contrast to usual claims in the literature, we find that factors associated with bond yields have a significant contribution to risk premia in commodity futures, mostly at medium and lower frequencies.

Finally, we estimate a nonlinear version of the model that imposes the zero lower bound on interest rates. Although the zero lower bound constraint is important for bond pricing, it has a minor impact on futures prices and risk premia.

REFERENCES

Alquist, R., Bauer, G. H., & Diez de los Rios, A. (2013). Macroeconomic drivers of crude oil futures risk premia. (*Technical report*). Ottawa, ON: Bank of Canada.

Ang, A., & Piazzesi, M. (2003). A no-arbitrage vector autoregression of term structure dynamics with macroeconomic and latent variables. *Journal of Monetary Economics*, *50*(4), 745–787.

Borovkova, S., & Geman, H. (2006). Seasonal and stochastic effects in commodity forward curves. *Review of Derivatives Research*, *9*(2), 167–186.

Busetti, F., & Harvey, A. (2003). Seasonality tests. *Journal of Business and Economic Statistics*, *21*(3), 420–436.

Canova, F., & Hansen, B. E. (1995). Are seasonal patterns constant over time? A test for seasonal stability. *Journal of Business and Economic Statistics*, *13*(3), 237–252.

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

Christensen, J. H., Diebold, F. X., & Rudebusch, G. D. (2011). The affine arbitrage-free class of Nelson–Siegel term structure model. *Journal of Econometrics*, *164*(1), 4–20.

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

Duffee, G. (2011). Forecasting with the term structure: The role of no-arbitrage restrictions. (*Economics Working Paper Archive No.576*). Baltimore, MD: Department of Economics, Johns Hopkins University.

EIA (2013). Oil market report. (Technical report). Retrieved from https://www.iea.org/media/omrreports/fullissues/2013-12-11.pdf

EIA (2017). Seasonal swings in U.S.distillate inventories lower as consumption patterns change. In Today in Energy. Retrieved from https://www.eia.gov/todayinenergy

Evans, L., & Guthrie, G. (2017). Commodity prices and the option value of storage. In Secomandi, N. (Ed.), *Real options in energy and commodity markets* (pp. 3–29). Singapore: World Scientific.

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

Gorton, G. B., Hayashi, F., & Rouwenhorst, K. G. (2013). The fundamentals of commodity futures returns. *Review of Finance*, 17(1), 35–105.

Hamilton, J. D., & Wu, J. C. (2014). Risk premia in crude oil futures prices. *Journal of International Money and Finance*, 42(C), 9–37.

Hannan, E. J. (1964). The estimation of a changing seasonal pattern. *Journal of the American Statistical Association*, 59(308), 1063–1077.

Hong, Z., Niu, L., & Zeng, G. (2016). *Discrete-time arbitrage-free Nelson–Siegel term structure model and application*. Xiamen, China, Wang Yanan Institute for Studies in Economics.

Joslin, S., Singleton, K., & Zhu, H. (2011). A new perspective on Gaussian dynamic term structure models. *Review of Financial Studies*, 24(3), 926–970.

Litzenberger, R. H., & Rabinowitz, N. (1995). Backwardation in oil futures markets: Theory and empirical evidence. *Journal of Finance*, 50(5), 1517–1545.

Miltersen, K. R., & Schwartz, E. S. (1998). Pricing of options on commodity futures with stochastic term structures of convenience yields and interest rates. *Journal of Financial and Quantitative Analysis*, 33(01), 33–59.

Nelson, C. R., & Siegel, A. F. (1987). Parsimonious modeling of yield curve. *Journal of Business*, 60(4), 473–489.

Routledge, B. R., Seppi, D. J., & Spatt, C. S. (2000). Equilibrium forward curves for commodities. *Journal of Finance*, 55(3), 1297–1338.

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

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

Stock, J., & Watson, M. (2002). Forecasting using principal components from a large number of predictors. *Journal of the American Statistical Association*, 97, 1167–1179.

Trolle, A. B., & Schwartz, E. S. (2009). Unspanned stochastic volatility and the pricing of commodity derivatives. *Review of Financial Studies*, 22(11), 4423–4461.

Wu, J. C., & Xia, F. D. (2016). Measuring the macroeconomic impact of monetary policy at the zero lower bound. *Journal of Money, Credit and Banking*, 48(2–3), 253–291.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

APPENDIX A: PROOFS

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

$$F_t^{(\tau)} E_t[M_{t,t+1}] = E_t[M_{t,t+\tau} S_{t+\tau}]. \qquad (A.1)$$

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

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

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

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

Incorporating this equation into Equation A.1, and noting that $E_t[M_{t,t+\tau}] = P_t^{(\tau)}$ and $E_{t+1}[M_{t+1,t+\tau}] = P_{t+1}^{(\tau-1)}$, gives

$$V_t^{(\tau)} = E_t[M_{t,t+1} V_{t+1}^{(\tau-1)}], \qquad (A.2)$$

where $V_t^{(\tau)} = F_t^{(\tau)} P_t^{(\tau)}$. Conjecture a solution of the form $\log V_t^{(\tau)} = G_\tau^{m_t} + H_\tau^{m_t\prime} X_t$. Incorporating Equations 5 and 6 into Equation A.2 and evaluating the conditional expectation gives

$$G_\tau^{m_t} + H_\tau^{m_t\prime} X_t = G_{\tau-1}^{m_{t+1}} + H_{\tau-1}^{m_{t+1}\prime}(\mu - \Gamma\lambda_0) + \frac{1}{2} H_{\tau-1}^{m_{t+1}\prime} \Gamma\Gamma\prime H_{\tau-1}^{m_{t+1}} - \rho_0 + \left(H_{\tau-1}^{m_{t+1}\prime}(\Theta - \Gamma\lambda_1) - \rho_1\prime\right) X_t.$$

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

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

$$H_\tau^{\tilde{m}} = (\Theta - \Gamma\lambda_1)' H_{\tau-1}^{\tilde{m}+1} - \rho_1.$$

Also, note that $V_t^{(0)} = F_t^{(0)} = S_t$. Therefore, the conjecture and Equation 12 imply

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

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

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

Equation 14 follows by setting $C_\tau^{m_t} = G_\tau^{m_t} - A_\tau$ and $D_\tau^{m_t} = H_\tau^{m_t} - B_\tau$.

Proof of Proposition 2: Nelson and Siegel representation

Log-bond prices and log-futures prices are affine functions of the state $p_t^{(\tau)} = A_\tau + B_\tau' X_t$ and $f_t^{(\tau)} = C_\tau^{\tilde{m}} + D_\tau^{\tilde{m}\prime}X_t$. Given the proposed parameters of the affine model, set $B_0 = \mathbf{0}$ and $H_0^{\tilde{m}} = \gamma_1^{\tilde{m}}$, and let the factor loadings B_τ and $H_\tau^{\tilde{m}}$ satisfy the recursions

$$B_\tau = \Theta^{Q\prime}B_{\tau-1} - \rho_1 \tag{A.3}$$

$$H_\tau^{\tilde{m}} = \Theta^{Q\prime}H_{\tau-1}^{\tilde{m}+1} - \rho_1, \tag{A.4}$$

and define $C_\tau^{\tilde{m}} = G_\tau^{\tilde{m}} - A_\tau$ and $D_\tau^{\tilde{m}} = H_\tau^{\tilde{m}} - B_\tau$. Now partition the matrix Θ^Q as

$$\Theta^Q = \begin{bmatrix} \Theta_{\delta\delta}^Q & \Theta_{\delta\beta}^Q & \Theta_{\delta\xi}^Q \\ \Theta_{\beta\delta}^Q & \Theta_{\beta\beta}^Q & \Theta_{\beta\xi}^Q \\ \Theta_{\xi\delta}^Q & \Theta_{\xi\beta}^Q & \Theta_{\xi\xi}^Q \end{bmatrix}$$

where the size of the submatrices conforms to the size of the vectors δ_t, β_t, and ξ_t.

We begin with the recursion for bonds. The Nelson and Siegel parametrization implies that bond prices depend only on δ_t (i.e., $B_\tau = -[\tau, \frac{1-e^{\zeta_1\tau}}{\zeta_1}, \frac{1-e^{\zeta_1\tau}}{\zeta_1} - \tau e^{-\zeta_1\tau}, \mathbf{0}_{1\times 6}]'$). Therefore, the nonzero elements of Equation A.3 satisfy

$$\begin{bmatrix} -\tau \\ -\frac{1-e^{\zeta_1\tau}}{\zeta_1} \\ \frac{1-e^{\zeta_1\tau}}{\zeta_1} - \tau e^{-\zeta_1\tau} \end{bmatrix} = \Theta_{\delta\delta}^{Q\prime} \begin{bmatrix} -(\tau-1) \\ -\frac{1-e^{\zeta_1(\tau-1)}}{\zeta_1} \\ \frac{1-e^{\zeta_1(\tau-1)}}{\zeta_1} - (\tau-1)e^{-\zeta_1(\tau-1)} \end{bmatrix} - \begin{bmatrix} 1 \\ -\frac{1-e^{\zeta_1}}{\zeta_1} \\ \frac{1-e^{\zeta_1}}{\zeta_1} - e^{-\zeta_1} \end{bmatrix}$$

which, in turn, implies

$$\Theta_{\delta\delta}^Q = \begin{bmatrix} 1 & 0 & 0 \\ 0 & e^{-\zeta_1} & \zeta_1 e^{-\zeta_1} \\ 0 & 0 & e^{-\zeta_1} \end{bmatrix}.$$

This condition for the arbitrage-free Nelson–Siegel model in discrete time was obtained by Hong, Niu, and Zeng (2016).[18]

Consider the parameters associated with the commodity factors. The Nelson–Siegel parametrization implies that the first three elements of $H_\tau^{\tilde{m}}$ are zero. Thus Equation A.4 imposes

$$\begin{bmatrix} 1 \\ \tau \\ \frac{1-e^{\zeta_2\tau}}{\zeta_2} \\ \frac{1-e^{\zeta_2\tau}}{\zeta_2} - \tau e^{-\zeta_2\tau} \\ e^{-\omega\tau}\cos(\frac{2\pi}{12}(\tilde{m}+\tau)) \\ e^{-\omega\tau}\sin(\frac{2\pi}{12}(\tilde{m}+\tau)) \end{bmatrix} = \begin{bmatrix} \Theta_{\beta\beta}^Q & \Theta_{\beta\xi}^Q \\ \Theta_{\xi\beta}^Q & \Theta_{\xi\xi}^Q \end{bmatrix}' \begin{bmatrix} 1 \\ \tau-1 \\ \frac{1-e^{\zeta_2(\tau-1)}}{\zeta_2} \\ \frac{1-e^{\zeta_2(\tau-1)}}{\zeta_2} - (\tau-1)e^{-\zeta_2(\tau-1)} \\ e^{-\omega(\tau-1)}\cos(\frac{2\pi}{12}(\tilde{m}+1+\tau-1)) \\ e^{-\omega(\tau-1)}\sin(\frac{2\pi}{12}(\tilde{m}+1+\tau-1)) \end{bmatrix} \tag{A.5}$$

---

[18]Joslin, Singleton, and Zhu (2011) show that the arbitrage-free Nelson–Siegel model imposes a single restriction on their maximally flexible (identified) model. In particular, the unconstrained case replaces the 1 in the $\Theta_{\delta\delta}^Q$ matrix with a free parameter.

and

$$
\begin{bmatrix} 1 \\ \dfrac{1-e^{\zeta_1 \tau}}{\zeta_1} \\ \dfrac{1-e^{\zeta_1 \tau}}{\zeta_1} - \tau e^{-\zeta_1 \tau} \end{bmatrix} = \begin{bmatrix} \Theta^{Q\prime}_{\beta\delta} & \Theta^{Q\prime}_{\xi\delta} \end{bmatrix} \begin{bmatrix} 1 \\ \tau - 1 \\ \dfrac{1-e^{\zeta_2(\tau-1)}}{\zeta_2} \\ \dfrac{1-e^{\zeta_2(\tau-1)}}{\zeta_2} - (\tau - 1)e^{-\zeta_2(\tau-1)} \\ e^{-\omega(\tau-1)}\cos(\tfrac{2\pi}{12}(\tilde{m} + 1 + \tau - 1)) \\ e^{-\omega(\tau-1)}\sin(\tfrac{2\pi}{12}(\tilde{m} + 1 + \tau - 1)) \end{bmatrix}.
\tag{A.6}
$$

It is easy to verify that Equation A.5 implies $\Theta^Q_{\xi\beta} = \mathbf{0}_{2\times 4}$, $\Theta^Q_{\beta\xi} = \mathbf{0}_{4\times 2}$:

$$
\Theta^Q_{\xi\xi} = \begin{bmatrix} e^{-\omega} & 0 \\ 0 & e^{-\omega} \end{bmatrix} \quad \text{and} \quad \Theta^Q_{\beta\beta} = \begin{bmatrix} 1 & 1 & \frac{1-e^{-\zeta_2}}{\zeta_2} & \frac{1-e^{-\zeta_2}}{\zeta_2} - e^{-\zeta_2} \\ 0 & 1 & 0 & 0 \\ 0 & 0 & e^{-\zeta_2} & \zeta_2 e^{-\zeta_2} \\ 0 & 0 & 0 & e^{-\zeta_2} \end{bmatrix}
$$

while Equation A.6 implies $\Theta^Q_{\xi\delta} = \mathbf{0}_{2\times 3}$ and

$$
\Theta^Q_{\beta\delta} = \begin{bmatrix} 1 & \frac{1-e^{-\zeta_1}}{\zeta_1} & \frac{1-e^{-\zeta_1}}{\zeta_1} - e^{-\zeta_1} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.
$$

Lastly, Equations A.3 and A.4 and $D^{\tilde{m}}_\tau = H^{\tilde{m}}_\tau - B_\tau$ imply $\Theta^Q_{\delta\beta} = \mathbf{0}_{3\times 4}$ and $\Theta^Q_{\delta\xi} = \mathbf{0}_{3\times 2}$.

APPENDIX B: RISK PREMIA

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

$$
f^{\tau-1}_{t+1} - f^\tau_t = C^{m_{t+1}}_{\tau-1} + D^{m_{t+1}\prime}_{\tau-1} X_{t+1} - C^{m_t}_\tau - D^{m_t\prime}_\tau X_t.
$$

Using Equations 5, 9, 10, 15, and 16, the time-t expected 1-period holding return can be written as

$$
E_t[f^{\tau-1}_{t+1} - f^\tau_t] = \frac{1}{2}[B'_{\tau-1}\Gamma\Gamma' B_{\tau-1} - H^{m_{t+1}\prime}_{\tau-1}\Gamma\Gamma' H^{m_{t+1}}_{\tau-1}] + D^{m_{t+1}\prime}_{\tau-1}\Gamma\Lambda_t.
$$

The spot premium is the 1-period holding return of a futures contract with settlement date in the next month. Using $f^0_{t+1} = s_{t+1}$, $B_0 = 0$ and $H^{m_{t+1}}_0 = D^{m_{t+1}}_0 = \gamma^{m_{t+1}}_1$, the spot premium is

$$
E_t[s_{t+1} - f^1_t] = -\frac{1}{2}[\gamma^{m_{t+1}\prime}_1 \Gamma\Gamma' \gamma^{m_{t+1}}_1] + \gamma^{m_{t+1}\prime}_1 \Gamma\Lambda_t.
$$

The term premium and the h-period holding returns follow from simple manipulations of the 1-period holding returns and the spot premium, as described in the text.