Time-Varying Term Structure of Oil Risk Premia

Gonzalo Cortazar,[a] Philip Liedtke,[b] Hector Ortega,[c] and Eduardo S. Schwartz[d]

ABSTRACT

We develop a framework to estimate time-varying commodity risk premia from multi-factor models using futures prices and analysts' forecasts of future prices. The model is calibrated for oil using a 3-factor stochastic commodity-pricing model with an affine risk premia specification. The WTI oil futures price data is from the New York Mercantile Exchange (NYMEX) and analysts' forecasts are from Bloomberg and the U.S Energy Information Administration. Weekly estimations for short, medium, and long-term risk premia between 2010 and 2017 are obtained. Results from the model calibration show that the term structure of oil risk premia moves stochastically through time, that short-term risk premia tend to be higher than long-term ones and that risk premia volatility is much higher for short maturities. An empirical analysis is performed to explore the macroeconomic and oil market variables that may explain the stochastic behavior of oil risk premia, showing that inventories, hedging pressure, term premium, default premium and the level of interest rates all play a significant role in explaining the risk premia.

Keywords: Commodities, Futures, Expected Prices, Pricing Models

https://doi.org/10.5547/01956574.43.5.gcor

1. INTRODUCTION

Risk premia represent the annualized price difference between the futures contracts' prices and the markets' expected prices. Therefore, the use of futures prices as the most likely price for a commodity in the future is only valid if the risk premium is equal to zero. Even though the existence of risk premia in futures contracts is not a new finding, there is no consensus on their magnitude, behavior and appropriate estimation procedure (Baumeister and Kilian, 2016; Bianchi and Piana, 2018; De Roon, Nijman, and Veld, 2000; Melolinna, 2011; Palazzo and Nobili, 2010). Moreover, the recent *financialization* of commodity markets has increased their relevance for investors and strengthened arguments on their time-varying behavior (Hamilton and Wu, 2014; Baker and Routledge, 2017; Ready, 2018; Fattouh, Kilian, & Mahadeva, 2013).

Understanding the stochastic behavior of commodity risk premia is important for several reasons. First, it provides valuable information on investment returns for agents who treat commodities as an asset class. Second, it helps to relate risk-adjusted expected prices, which are readily available in futures markets, with those of true expected prices, which are preferred by many practitioners for net present value calculations and risk management purposes. Third, it may shed light on some public policy implications by uncovering the macroeconomic determinants of risk premia.

a Pontificia Universidad Católica de Chile. E-mail: gcortaza@ing.puc.cl

b Pontificia Universidad Católica de Chile. E-mail: pliedtke@uc.cl

c Pontificia Universidad Católica de Chile. E-mail: hiortega@uc.cl

d Corresponding author. Beedie School of Business, SFU, UCLA Anderson School of Management and NBER, 500 Granville Street, Vancouver, BC, Canada V6C 1W6. E-mail: eduardos@sfu.ca.

There have been various attempts in the literature to estimate commodity risk premia. Many practitioners and researchers use futures prices as proxies for market expectations (as discussed in Baumeister and Kilian, 2016; Bianchi and Piana, 2018), implicitly assuming that risk premia are zero. Keynes (1930) and Hicks (1939) had proposed that if producers and other market participants wanted to hedge their risk by selling future contracts, buyers should get a compensation in the form of a risk premia for taking on that risk. Furthermore, there is already evidence on its time-varying nature (De Roon, Nijman, and Veld, 2000; Sadorsky, 2002; Pagano and Pisani, 2009; Achraya, Lochstoer, and Ramadorai, 2013; Etula, 2013; Hamilton and Wu, 2014; Singleton, 2014).

In the last few years, different methods have been developed to extract risk premia, or equivalently to calculate expected spot prices, from the available data. Even though most of the literature addresses how to get the market's expected interest rates (e.g., Diebold and Li, 2006; Altavilla, Giacomini, and Costantini, 2014; Chun, 2011), some effort has also been oriented to commodities.

In what follows we present one way of characterizing existing methods for estimating risk premia in commodity markets by classifying them into three approaches: *econometric*, *economic*, and *market-based*.

In what we call the *econometric* approach we include Gorton, Hayashi, and Rouwenhorst (2013), Hong and Yogo (2012), Pagano and Pisani (2009) and Baumeister and Kilian (2016) among others. This approach regresses realized spot commodity prices, or a function of them, on different lagged market variables to infer the expected market's spot price. Then the resulting risk premia are obtained by comparing this expected spot price with the futures price for the same maturity. Given that realized future spot prices and current futures prices with same maturity are compared, the required data-sample gets larger as longer-term risk premia are estimated making it difficult to estimate current risk premia for maturities greater than one or two years.

In what we call the *economic* approach we include Hamilton and Wu (2014), Bianchi and Piana (2018), and Cortazar, Kovacevic, and Schwartz (2015) among others. These models use no-arbitrage or rational expectation models to infer expected spot prices from past and current market variables, typically futures and spot prices. Even though most of these types of models are successful in fitting futures prices, they do not provide appropriate risk premia estimates. To solve for this issue, asset-pricing models have been extensively applied, obtaining mixed results (Dhume, 2010; Erb and Harvey, 2006; Hong and Yogo, 2012).

In what we call the *market-based* approach we include a recent paper by Cortazar, Millard, Ortega, and Schwartz (2019) in which they propose extracting information on expected spot prices directly from market surveys and using them, in addition to spot and futures prices, to calibrate a term structure model. Thus, risk premia are obtained directly from the model as the difference between the expected spot price and the futures price consensus curves. Including survey forecasts in economic models, even though it had not been previously applied to commodities, had been used in various contexts (see Chun (2011), and Altavilla, Giacomini, and Ragusa (2017)). This new approach allows to get risk premia directly from market observations (i.e., analysts' forecasts) as opposed to the traditional methods that usually infer them based on futures prices.

This paper develops a framework to estimate time-varying risk premia using the *market-based* approach by extending Cortazar et al. (2019) to allow for a stochastic specification of risk premia following Duffee (2002). Once the term structures for oil risk premia are estimated, we explore their market determinants performing several regressions on different macroeconomic variables and oil market variables that have been previously proposed in the literature (e.g., Bhar and Lee (2011)).

We find that the risk premia are time varying and are partially explained by market variables, namely inventories, hedging pressure, term premium, default premium and the level of interest rates. In this sense our methodology can estimate significant time-varying ex-ante risk premia directly from futures prices and analysts' forecasts, and search for their relationship with other market variables after having agreed on their levels and structure. It differs from the research that use the *economic* approach (e.g., Cortazar, Kovacevic, and Schwartz (2015)) on the fact that their risk premia are constant over time. On the other hand, it is fundamentally different to those studies that, following an *econometri*c approach, regress ex-post risk premia (e.g., Baumeister and Kilian (2016)) with market variables due to the different nature of ex-ante and ex-post risk premia, where the first are directly related to market expectations, while the second may differ from the latter by the existence of biases.

The remainder of this paper is organized as follows. Section 2 presents the theoretical model used to estimate time-varying term structures of risk premia. Section 3 describes the data. Section 4 provides risk premia estimates. Section 5 discusses the market determinants of risk premia and Section 6 concludes.

2. THE MODEL TO ESTIMATE RISK PREMIA

2.1 Model Definition

We present an N-factor term structure model which is a non-stationary version of the canonical $A_0(N)$ Dai and Singleton (2000) model with stochastic risk premia as in Duffee (2002). We propose calibrating this model using both futures prices and analysts' forecasts to obtain a time-varying term structure of risk premia[1].

Let S_t be the spot price of the commodity at time t, then assume that:

$$lnS_t = Y_t = h'x_t \tag{1}$$

$$dx_t = \left(-Ax_t + \begin{bmatrix} b_1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \right) dt + dw_t \tag{2}$$

where h is an n×1 vector of constants, x_t is an n×1 vector of state variables, b_1 is a scalar, A is an n×n upper triangular matrix with its first diagonal element being zero and the other diagonal elements all different and strictly positive. Let dw_t be an n×1 vector of uncorrelated Brownian motions following

$$dw_t dw_t' = Idt \tag{3}$$

where I is an n×n identity matrix. Dai and Singleton (2000) show that their model has the maximum number of econometrically identifiable parameters and at the same time nests most of the models used in literature.

To specify time-varying risk premia in our constant-volatility model we resort to Duffee (2002) who shows how to use affine risk premia in all types of Dai and Singleton (2000) canonical

1. This paper builds on Cortazar, Millard, Ortega, Schwartz (2019) which also used futures and analysts' forecasts, but assumed constant risk premia. That paper used the Cortazar and Naranjo (2006) N-factor model. In Online Appendix 1 we show that our proposed model is a rotated version of the Cortazar and Naranjo (2006) model.

models, including the ones with non-stochastic volatility. Let RP_t be the commodity risk premia and assume that:

$$RP_t = \lambda + \Lambda x_t \tag{4}$$

and the risk adjusted version of the model shown in Equations 1 and 2, is

$$Y_t = h' x_t \tag{5}$$

$$dx_t = \left[-\left(A + \Lambda\right) x_t + \begin{bmatrix} b_1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} - \lambda \right] dt + dw^Q \tag{6}$$

where λ is a n×1 vector and Λ is a n×n matrix which does not need to be diagonal nor triangular. No further restrictions are set for the elements[2] in λ and Λ.

Note that in our model the risk-adjusted process differs from the true one not only by a constant risk premium, λ , but also by the Λ matrix. Thus, futures prices and expected prices depend on different processes for the state variables, the former with the $A + \Lambda$ matrix, while the latter only with matrix A. However, if the Λ matrix were set to zero, risk premia would be a constant and not time-varying.

Futures prices are the expected value of the spot price, S_t, under the risk-adjusted probability measure, Q (Cox, Ingersoll, and Ross, 1981). Given that the risk-adjusted spot price follows a log-normal distribution, futures prices are given by:

$$F_t\left(T\right) = E_t^Q\left(S_T\right) = e^{E_t^Q\left(Y_T\right) + \frac{1}{2} Var^Q\left(Y_T\right)} \tag{7}$$

where the risk-adjusted expected price and variance of Y_T can be obtained by replacing Equation 1 into 7:

$$F_t\left(T\right) = E_t^Q\left(S_T\right) = e^{h' E_t^Q\left(x_T\right) + \frac{1}{2} h' Cov^Q\left(x_T\right) h} \tag{8}$$

with[3]

$$E_t^Q\left(x_T\right) = e^{-(A+\Lambda)(T-t)} x_t + \left(\int_0^{T-t} e^{-(A+\Lambda)\tau} d\tau\right)\left(b - \lambda\right) \tag{9}$$

$$Cov_t^Q\left(x_T\right) = \int_0^{T-t} e^{-(A+\Lambda)\tau} \left(e^{-(A+\Lambda)\tau}\right)' d\tau \tag{10}$$

Analogous to Equations 7, 8, 9 and 10, the expected price should satisfy the following equations:

$$E_t\left(S_T\right) = e^{E_t\left(Y_T\right) + \frac{1}{2} Var\left(Y_T\right)} \tag{11}$$

2. An equivalent model definition is also used by Casassus and Collin-Dufresne (2005), Dai and Singleton (2002), Duarte (2004), Kim and Orphanides (2012), Palazzo and Nobili (2010) among others, however none of them use observations on analysts' forecasts as expected prices as we propose, having difficulties estimating significant risk premia.

3. See Online Appendix 2.

$$E_t\left(S_T\right) = e^{h'E_t\left(x_T\right)+\frac{1}{2}h'Cov\left(x_T\right)h} \tag{12}$$

$$E_t\left(x_T\right) = e^{-A(T-t)}x_t + \left(\int_0^{T-t} e^{-A\tau}d\tau\right)b \tag{13}$$

$$Cov_t\left(x_T\right) = \int_0^{T-t} e^{-A\tau}\left(e^{-A\tau}\right)' d\tau \tag{14}$$

It can be shown[4] that Equations 9 and 10 have a closed form solution if matrix $A+\Lambda$ is diagonal. The same occurs for Equations 13 and 14, now considering matrix A. In a more general case, as in our model, futures prices and expected prices have to be obtained numerically[5].

Annualized risk premia are defined as the log return of the expected spot price over the future price. Let, $\pi_t\left(T-t\right)$ be the instantaneous risk premia at time t for T-t years ahead:

$$\pi_t\left(T-t\right) = \frac{ln\left(\frac{E_t\left(S_T\right)}{F_t\left(T\right)}\right)}{T-t} \tag{15}$$

Then, replacing the expected spot price and the future price from Equations 8 and 12 we obtain

$$\pi_t\left(T-t\right) = \frac{h'\left(E_t\left(x_T\right)-E_t^Q\left(x_T\right)\right)+\frac{1}{2}h'\left(Cov_t\left(x_T\right)-Cov_t^Q\left(x_T\right)\right)h}{T-t} \tag{16}$$

Finally, implied model volatilities for expected spots, σ_E, and for futures prices, σ_F, may be computed using the following expressions[6]:

$$\sigma_E = \sqrt{h'e^{-A(T-t)}\left(e^{-A(T-t)}\right)'h} \tag{17}$$

$$\sigma_F = \sqrt{h'e^{-(A+\Lambda)(T-t)}\left(e^{-(A+\Lambda)(T-t)}\right)'h} \tag{18}$$

2.2 Model Estimation

The parameters of the model and the state variables are estimated using the Kalman Filter (Kalman, 1960), which computes the optimal value of each state variable for any given time taking all past information into account. The procedure can handle a large number of observations (in our case analysts' forecasts and futures prices) and allows for measurement errors.

At any given time-iteration (date), a variable number of observations is available, so we use the incomplete data panel specification of the Kalman filter previously used for Futures (Cortazar and Naranjo, 2006), Bonds (Cortazar, Schwartz, Naranjo, 2007) and Analysts' forecasts (Cortazar et al., 2019):

4. See Online Appendix 2.

5. To solve the equations efficiently we follow Pashke and Prokopczuk (2009) who develop a way of avoiding numerical integration, using a decomposition of matrix $A+\Lambda$ in eigenvalues and eigenvectors. See Online Appendix 3.

6. See Online Appendix 4.

$$z_t = Hx_t + d + v_t \qquad v_t \sim N(0, R) \tag{19}$$

$$x_{t+1} = \bar{A}x_t + \bar{c} + w_t \qquad w_t \sim N(0, Q) \tag{20}$$

where z_t is an m_t×1 vector which contains the log-prices of each futures and analysts' forecast (in that order) observation at week t; H is an m_t×n matrix; d is an m_t×1 vector and v_t is an m_t×1 vector of measurement errors with zero mean and covariance given by R; x_t is the n×1 vector of the state variables from Equation 1; $\bar{A}$ and $\bar{c}$ are an n×n matrix and an n×1 vector, respectively, representing a discretization of the process described in Equation 2 and w_t is an n×n vector of random variables with mean zero and covariance given by the n×n matrix Q. In this specification m_t varies depending on the number of available observations changing the size of z_t, H, d, v_t and R on every iteration.

In contrast to Cortazar et al. (2019) we specify two error terms in Equation (19), with different variances to differentiate between futures prices and forecasts, since the latter include estimations from different analysts' and should be much noisier.

Thus, we define the m_t×m_t matrix R as follows:

$$R_t = \begin{bmatrix} \sigma_f & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma_f & 0 & \cdots & 0 \\ 0 & \cdots & 0 & \sigma_e & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & \sigma_e \end{bmatrix} \tag{21}$$

To estimate the parameters of this model a maximum-likelihood approach is used.

3. DATA

To be able to estimate the risk premia, futures prices and analysts' forecasts for different dates and maturities are required. The data was obtained from different data sources for the period between January 2010 and June 2017. Before 2010 there is no long-term spot price estimates available. The remainder of this section further describes the data used.

3.1 Futures Contracts

WTI crude oil futures prices are obtained from the NYMEX. We used weekly futures prices with expiration every 6 months, including the closest one to maturity. The longest traded contracts expire in approximately 9.2 years. Table 1 presents the futures price data, which has been aggregated into one-year maturity buckets; that is, the first bucket has the contracts maturing in less than a year, the second bucket has the contracts maturing between one and two years, and so on.

3.2 Survey Based Expected Prices

We collected Bloomberg's analysts' WTI price predictions to use them as a noisy proxy of WTI expected future spot prices. Bloomberg's price predictions consist of a list of surveys done to professional analysts on the expected future commodity prices. The expectations are given quar-

Table 1: Futures price observations between January 2010 and June 2017 aggregated into yearly maturity buckets

Maturity Bucket (years)	Mean Price ($/bbl.)	Price SD	Max Price ($/bbl.)	Min Price ($/bbl.)	Mean Maturity (years)	Number of Observations
0–1	77.88	22.28	113.7	26.55	0.45	968
1–2	78.23	19.43	112.83	35.36	1.48	795
2–3	77.55	17.59	109.33	38.66	2.49	821
3–4	77.29	16.41	107.14	41.34	3.51	783
4–5	77.39	15.76	105.8	43.24	4.49	786
5–6	77.48	15.4	105.56	44.42	5.47	809
6–7	78	15.22	105.88	45.77	6.5	767
7–8	78.2	15.2	106.3	46.5	7.49	774
8–9	78.27	15.72	106.95	46.99	8.43	635
9–10	77.15	13.79	95.16	55.08	9.06	44

terly for the next 8 quarters and yearly for the next 4 years. Data is available only when one of the many analysts does a prediction, and may become available any day of the week. Each prediction is grouped on the upcoming Wednesday resulting in weekly groups of observations. If predictions for the same maturity on the same date are available, their mean value is used. On average, there are 220 oil price predictions available every month for different maturities.

In addition to Bloomberg analysts' expectations, EIA's oil price forecasts are also used. Data is available once a year since 2010. EIA's data includes yearly long-term predictions for up to 33 years ahead. Even though both Bloomberg's and EIA's predictions are for the average price of each quarter or year they were assumed to represent the price in the middle of their time period. Data of the current quarter and year were excluded since part of the prediction period might have already been realized. Table 2 describes the forecast data used. The maturity bucket size grows with maturity due to the fewer observations available for longer maturities.

Table 2: Analysts' price forecasts between January 2010 and June 2017 aggregated into maturity buckets

Maturity Bucket (years)	Mean Price ($/bbl.)	Price SD	Max Price ($/bbl.)	Min Price ($/bbl.)	Mean Maturity (years)	Number of Observations
0–1	81.02	22.26	122	35	0.53	1118
1–2	85.62	21.25	135	40	1.43	808
2–3	89.04	23.43	189	44	2.48	289
3–4	88.34	23.03	154	40	3.44	239
4–5	86.21	22.68	150	38.5	4.42	179
5–10	101.22	22.03	152.96	60	6.29	79
10–34	171.56	34.23	265.2	104.68	18.48	134

4. RESULTS

In this section we use the data described in the previous section to estimate all the parameters of the N-factor model using a 3-factor specification.

Table 3 shows the model parameter estimates. It can be noted that half of the parameter estimates are statistically significant at a 1% and 3/4 of them at a 10% significance level. The most relevant parameters of the model correspond to the diagonal elements of matrix A, which are statistically significant. These parameters represent the mean reversion level of the state variables x_t (equation 2) and are the structural pillars of the whole term structure model. The bigger value of the A_{22} variable indicates that oil prices are correctly explained by a term structure model with three

Table 3: Parameter estimates for the 3-factor model. Data between January 2010 and June 2017

	Estimate	Deviation	tStat	pValue
A_{11}	0	—	—	—
A_{12}	0.728*	0.3676	1.9802	0.0564
A_{13}	1.4204	0.9677	1.4678	0.1358
A_{22}	1.4929***	0.185	8.0674	0
A_{23}	2.7146*	1.3379	2.0291	0.0512
A_{33}	0.163***	0.0238	6.8577	0
Λ_{11}	0.2267***	0.0076	29.7516	0
Λ_{12}	–0.7768*	0.3877	–2.0037	0.0539
Λ_{13}	–1.5684*	0.9397	–1.669	0.0992
Λ_{21}	–0.044	0.0423	–1.0404	0.2319
Λ_{22}	–1.3074***	0.2862	–4.5686	0
Λ_{23}	–2.2669	1.4022	–1.6166	0.108
Λ_{31}	–0.0306	0.0248	–1.2314	0.1867
Λ_{32}	0.2826***	0.0673	–4.2015	0.0001
Λ_{33}	0.4187***	0.111	3.7708	0.0004
h_1	0.1521***	0.0184	8.2626	0
h_2	0.2146*	0.117	1.8333	0.0745
h_3	0.7469***	0.0431	17.3302	0
λ_1	–6.081***	0.5953	–102143	0
λ_2	1.2692	1.2894	0.9843	0.2454
λ_3	1.0407*	0.6039	1.7233	0.0905
b_1	0.1767***	0.0543	3.2549	0.0021
σ_f	0.0058***	0	302.9228	0
σ_e	0.1***	0.0005	195.163	0

* Significant at the 10% level; ** significant at the 5% level; *** significant at the 1% level.

state variables, the first of them non-stationary, the second with a fast mean reversion (x_2) and the third one with a slower mean reversion (x_3).

Additionally, it is worth noting that the diagonal elements of matrix Λ and the first element of vector λ are also statistically significant. The diagonal elements of Λ are again related to the risk adjusted mean reversion of the state variables, and therefore represent the existence of time-varying component of risk premia, and the first element of λ represents the risk adjusted non-stationarity of the first state variable, and the existence of a constant component of risk premia.

Figure 1 shows the term structure (from 1 month to 10 years) of annualized risk premia over the whole sample period (01/2010 to 06/2017). Three things are worth noting. First, the term structure of risk premia varies stochastically through time which is the starting point of our analysis. This opens the possibility of risk premia being influenced by other market variables. Second, short-term risk premia tend to be higher than long-term ones. This result suggests that on average investors have higher hedging demands for short term contracts allowing speculators to require higher risk premia. Third, risk premia volatility is much higher for short maturities. This result suggests that there is more disagreement between market participants for short term contract prices.

Figure 2 shows the term structure, the mean and the volatility of risk premia. Figure 2a) compares our model's mean risk premia to those of Cortazar et al. (2019) constant risk premia model (with our same data) and to the data average. It can be noted that our model's mean risk premia levels are similar to those of Cortazar et al. (2019) and both fit the data risk premia well. Additionally, both premia decrease with maturity. The similarity in both models of the average risk premia, however, is misleading since as Figure 3 shows that while the mean might be similar over our sample period, the time variation is dramatically different.

Figure 1: Annualized risk premia term structure from 1 month to 10 years. Data between January 2010 and June 2017

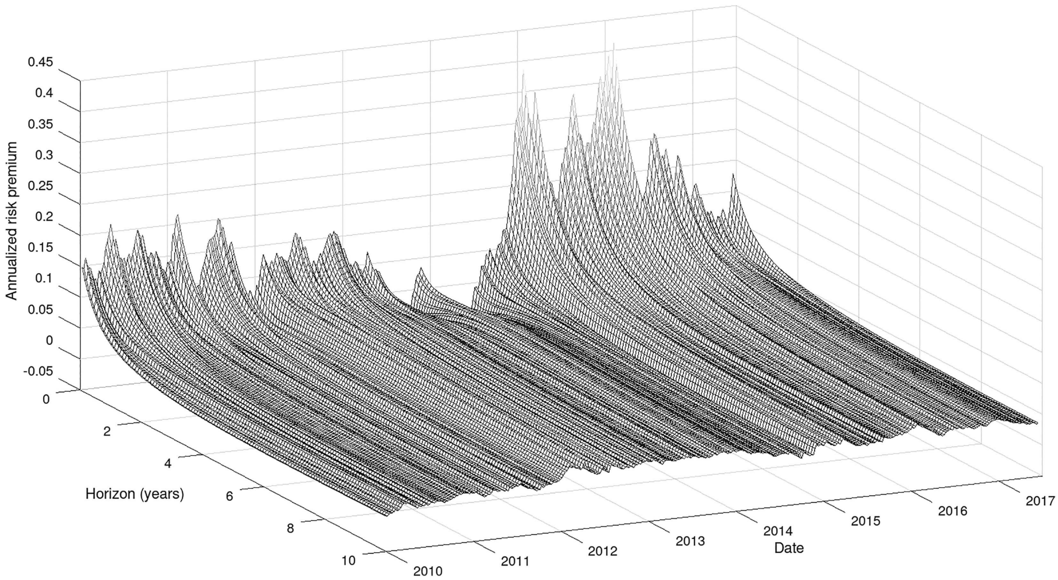
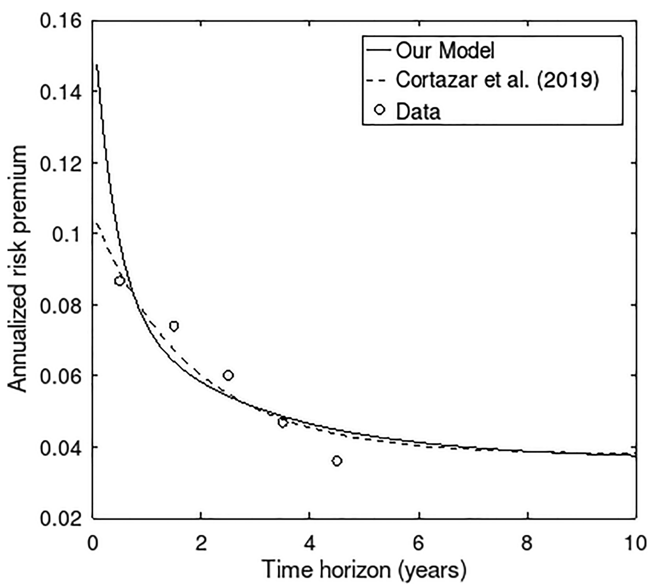


Figure 2a: Mean risk premia for our model and for the constant risk premia model in Cortazar et al. (2019). Data mean risk premia are also included calculated as the mean of the difference between each analysts' forecast and the closest futures contract. Data between January 2010 and June 2017.



Therefore, where the two models really differ is in Figure 2b) that shows no risk premia volatility for Cortazar et al. (2019), but a time-varying risk premia for ours. This shows that even though Cortazar et al. (2019) was able to estimate risk premia that on average adjust the data, it was not able to capture their variations through time.

Figure 2b: Risk premia volatility for our model and for the constant risk premia model in Cortazar et al. (2019). Data between January 2010 and June 2017.

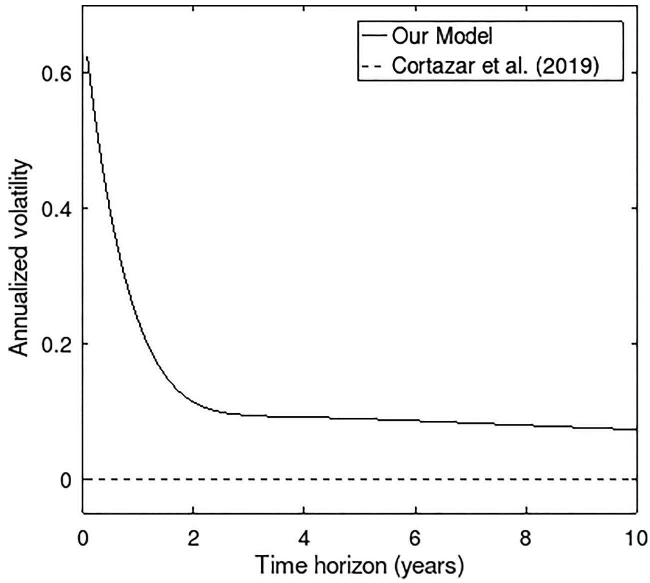
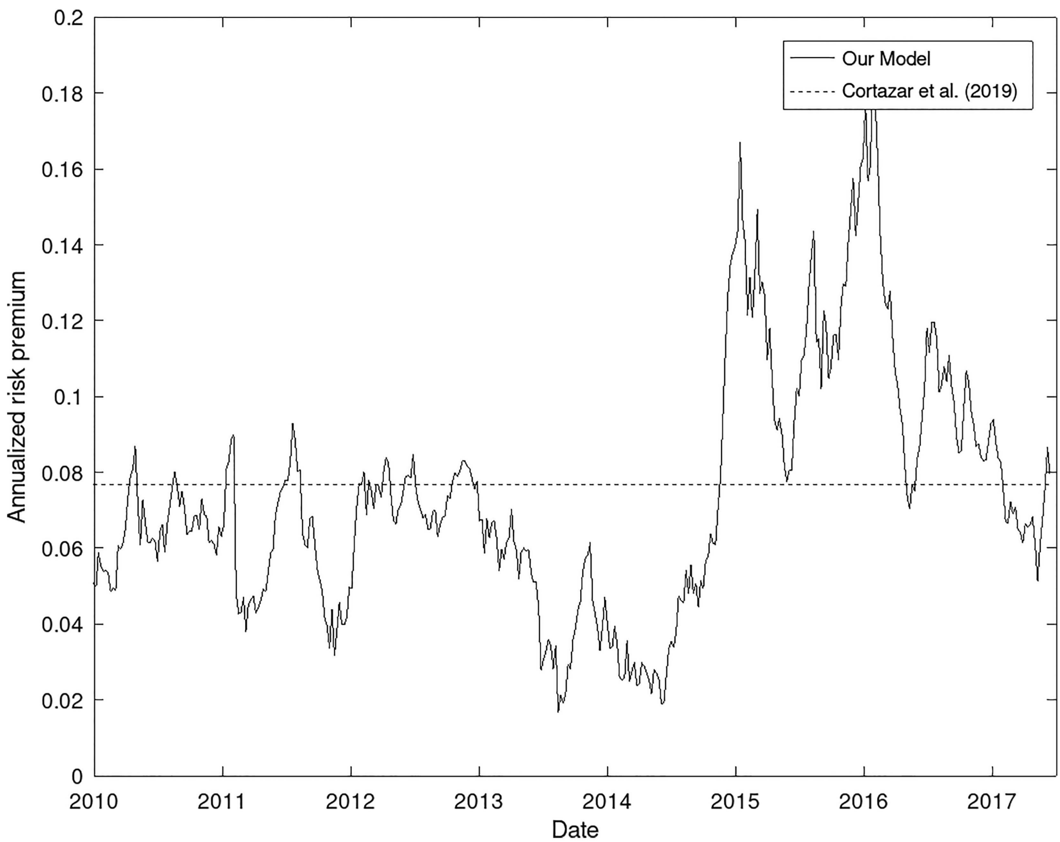


Figure 3: Time series of one-year risk premia given by our model and Cortazar et al. (2019) for comparison. Data between January 2010 and June 2017.



Finally, we analyze the goodness-of-fit of our model to futures and analysts' forecasts data. Table 4 compares the MAPE of our model with Cortazar et al. (2019) and shows that its fit for both data sets is better. The fit for futures prices is only slightly improved, but the fit for expected prices shows a notable improvement. The reason for the minor improvement in futures prices is that Cortazar et al. (2019) must fit both, futures and expected prices, with constant link between them, namely risk premia. As there is a larger number of futures prices observations the model already fits futures prices better than expected prices. When we allow risk premia to be time-varying, futures prices' fit does not improve much as the model did already fit them nicely, however expected prices' fit does improve as the model now allows for more flexibility.

Table 4: Mean Absolute Percentage Error (MAPE) for our time-varying risk premia model and for Cortazar et al. (2019) constant risk premia model. Data between January 2010 and June 2017.

	Our model	Cortazar et. al. (2019)
Futures prices	0.37%	0.39%
Expected prices	7.39%	8.00%

The superiority of the time-varying model becomes apparent when we analyze the market determinants of the risk premia, as we do in the next section.

5. THE DETERMINANTS OF OIL RISK PREMIA

5.1 The Methodology

In this section, we explore for the potential market determinants that may explain the time-series of the estimated oil risk premia. We consider a set of market variables that have been previously reported in the literature as affecting risk premia. We then perform a series of linear regressions to determine which variables are significant in explaining the term structure of oil risk premia.

There are not too many studies which analyze risk premia directly (e.g., Bhar and Lee, 2011; Bianchi and Piana, 2018; Chen and Zhang, 2011; Melolinna, 2011) as most only calculate them as a side result from a price prediction model. However, below we review some of the literature that discusses the impact of different market variables on risk premia.

The potential explanatory variables for oil risk premia that we consider are: the S&P500 Index returns, the NASDAQ Emerging Markets Index returns (EMI), oil inventories percentage variation, oil futures open interest percentage variation, hedging pressure, the interest rate term premium, the default premium and the 5-year treasury bill rate. As we explain below these variables have been shown to include the risk factors present in the oil market.

The **S&P500 index returns** is used in some studies (De Roon, Nijman, and Veld, 2000; Bianchi and Piana, 2018) as a proxy for the state of the US economy which is a big player in the oil market. Daily returns are available in Bloomberg since 1950.

The **NASDAQ Emerging Markets Index (EMI)** represents the state of the emerging markets' economy. Many big emerging economies, such as Russia and China, are important oil market players, hence their economic performance could directly affect oil prices and premia. EMI daily returns are available from the NASDAQ database since 2001.

Oil inventories percentage variation is a commonly used regressor in oil studies (Gorton, Hayashi, and Rouwenhorst, 2013; Melolinna 2011) since it directly affects the supply of oil and therefore its price. The theoretical relationship between available stocks and risk premia was first introduced by Kaldor (1939) in his *Theory of Storage*, in which he proposes the existence of a convenience yield to explain differences between current spot and futures prices. Gorton et al. (2013) develop a model, based on Kaldor (1939) *Theory of Storage,* which implies that a rise in inventories should lead to a decrease in the overall risk premia, and they obtain empirical results supporting their model. Weekly US WTI inventories starting at 1983 are available from the EIA. Their percentage differences were used in order to obtain a stationary time series.

Open Interest (OI) and **Hedging Pressure (HP)** are the usual measures of size and behavior of the market for a futures contract (in our case WTI futures). OI is measured as the total number of outstanding contracts, and therefore represents the market size. Larger amounts of outstanding contracts could affect the liquidity of the market and therefore the premia. Kang, Rouwenhorst, and Tang (2019) argue that there exists a liquidity premium on commodity futures markets. OI is often used as an explanatory variable for commodity related studies (Bianchi and Piana, 2018; Hong and Yogo, 2012).

HP is measured as the net positions of hedgers in a specific market, and represents the difference between hedgers' and speculators' positions, which according to Keynes (1930) and Hicks (1939) should have a strong relation with risk premia. If hedgers want to hedge their risk by selling futures contracts, the buyers of those contracts should get a compensation for taking on that risk. As HP rises, risk premia will rise, because speculators will be willing to accept a greater amount of risk only if premia are big enough. The relation between HP and prices or premia has been empirically tested by different studies (Bianchi and Piana, 2018; De Roon, Nijman, and Veld, 2000; Gorton, Hayashi, and Rouwenhorst, 2013; Kang, Rouwenhorst, and Tang, 2019; among others) generally supporting Keynes (1930). OI and HP weekly data was obtained from reports from the Commodity Futures Trading Commission (CFTC), which is available since 2007. OI is directly available in the reports and their weekly percentage variations were used in the analysis. HP was computed as the short minus long commercial positions, divided by the total amount of outstanding contracts:

$$HP_t = \frac{CS_t - CL_t}{OI_t} \tag{22}$$

where CS_t and CL_t stand for short and long commercial positions, respectively.

The **term premium (TRM)** and the **default premium (DEF)** have shown to predict market excess returns in stocks and bonds (Fama and French, 1989; Keim and Stambaugh, 1986), and could, therefore, also affect oil risk premia. TRM is defined as the difference between the 10-year treasury bill rate and the 3-month treasury bond yield, and DEF as the difference between the BAA-rated and the AAA-rated corporate bond yields. Daily treasury bill rates are available at the Federal Reserve while corporate bond yields were obtained from the Federal Reserve Bank of St. Louis. Finally, the **5-year treasury bill rate (**5Y T-Bill) is also included as a proxy for the state of the economy.

Once the potential independent variables were chosen a set of multivariate OLS regressions were conducted:

$$RP_{it} = \beta_{0i} + \beta_{1i} X_t + \varepsilon_{it} \tag{23}$$

where RP_{it} is the risk premia for maturity i and date t, X_t is the set of regressors described previously which are independent of the maturity, β_{0i} and β_{1i} are the estimators for each maturity i, and ε_{it} is the regression error for maturity i and date t.

We conduct our analysis in two steps. In the first step, we run a multi-variate regression for all the independent variables to check which ones can explain risk premia. Then, we run a second multivariate regression using only the variables that were significant in the first regression. We run risk premia regressions for 3, 6, 12, 18 months and 2, 5 and 10-year maturities. An independent regression is performed for every different time horizon, both in the first and second multivariate regressions. Robust standard errors were used in order to account for possible heteroscedasticity.

5.2 The Results

Table 5 shows the results of the first multivariate regression with all the variables included and for different maturities from 3 months (M3) to 10 years (Y10). Inventories, HP, TRM, 5Y T-Bill and DEF have significant (p-value) coefficients so they are candidates for inclusion in the second regression, while the other variables are not.

Table 5: Regression analysis for all the chosen independent variables and for each different maturity. Data between January 2010 and June 2017.

	M3	M6	Y1	M18	Y2	Y5	Y10
Intercept	0.0108	0.0187	0.0353***	0.0461***	0.0495***	0.0438***	0.0385***
S&P500	–0.1666	–0.1147	–0.043	–0.0095	0.0057	0.0167	0.0126
EMI	–0.1834	–0.1403	–0.0874	–0.0531	–0.0334	–0.0032	–0.0018
Inventories	2.4923***	2.0111***	1.3525***	1.0618***	0.8731***	0.4878***	0.387***
OI	–0.1224	–0.0916	–0.0523	–0.0291	–0.0168	–0.0006	–0.0005
HP	–0.0021	0.0099	0.0267**	0.0399***	0.0497***	0.0595***	0.0488***
TRM	–0.0762***	–0.0601***	–0.038***	–0.0249***	–0.017***	–0.0055***	–0.0057***
5Y T-Bill	0.0918***	0.0698***	0.0402***	0.023***	0.0134***	0.0007	0.002*
DEF	0.1472***	0.1093***	0.0605***	0.0335***	0.0192***	0.001	0.0008
R2	0.4444	0.4537	0.4774	0.5201	0.5624	0.5276	0.5507

* Significant at the 10% level; ** significant at the 5% level; *** significant at the 1% level.

Table 6 shows the results of the second multivariate regression for each maturity using only the variables that were significant in the previous analysis. It can be noted that the R-Squared of the regressions vary between 43.74% and 55.48%, and all variables are significant for most of the maturities.

Table 6: Regression coefficients for each maturity for the variables that showed statistical significance in the first regression analysis. Data between January 2010 and June 2017.

	M3	M6	Y1	M18	Y2	Y5	Y10
Intercept	0.0052	0.0151	0.0345***	0.0458***	0.0494***	0.0437***	0.0385***
Inventories	2.297***	1.8561***	1.2818***	1.0172***	0.8458***	0.4883***	0.3884***
HP	0.0032	0.0132	0.0282**	0.0406***	0.0502***	0.0598***	0.0489***
TRM	–0.0755***	–0.0596***	–0.0379***	–0.0249***	–0.0169***	–0.0054***	–0.0057***
5Y T-Bill	0.0922***	0.07***	0.0401***	0.0229***	0.0133***	0.0007	0.0021*
DEF	0.1496***	0.1109***	0.0608***	0.0336***	0.0192***	0.001	0.0008
R2	0.4374	0.4465	0.4722	0.5177	0.5609	0.5318	0.5548

* Significant at the 10% level; ** significant at the 5% level; *** significant at the 1% level.

The key results from Table 6 are as follows. First, we find a statistically significant and maturity-independent positive relation between inventories and risk premia. This is similar to Dincerler, Khokher, and Simin (2020) and Khan, Khokher and Simin (2008), but not consistent with the results in Gorton et al. (2013).

Second, our positive and significant value for the HP estimator is consistent with Keynes (1930) theory of normal backwardation, as a larger number of hedgers wanting to hedge their risk produces a larger HP which should by related to speculators demanding larger premia to take on that risk. Basu and Miffre (2013), De Roon, Nijman, and Veld (2000) and Bianchi and Piana (2018), among others, obtain similar results. HP is only significant for maturities of over one year, which could be explained by hedgers normally buying or selling contracts to hedge their medium- and long-term risk to secure the continuity of their operations on the long run.

Third, TRM is negative and significant for all maturities. These results are consistent with a negative slope of the yield curve predicting a decrease in GDP (Estrella and Hardouvelis, 1991; Harvey, 1988) which could lead to an inverse relation with premia.

Finally, the 5Y T-Bill and DEF have a positive effect on risk premia, however only for maturities up to two years. The DEF is considered as a good measure of market uncertainty, especially at the short-term. Higher short-term uncertainty should induce the average investor to demand larger premia especially for short term investments which is consistent with DEF affecting only short-term premia. If the treasury bill yield serves as a proxy for the current state of the economy, being higher when the economy grows and lower on slow economic periods, we would expect to get a negative effect of it on risk premia, such as in Bhar and Lee (2011). However, interest rates were unusually and constantly low during our sample period, which might alter the way in which treasury bill yields represent the state of economy.

These results suggest that these 5 market variables are able to explain half of the variation of oil risk premia in our model for all studied maturities. In addition to the economic insight the regression results provide, they can also be used to obtain estimates of risk premia and therefore expectations of future spot prices. For example, many practitioners who currently use futures prices as a proxy for the market's spot price expectations, could infer them directly from our market variables.

Figure 4 shows expected spot price estimations for two different maturities obtained by adding the expected risk premia from our regression analysis to the observed futures prices, along with analysts' forecasts and futures prices observations. The figure shows that by adding the risk premia to futures prices we obtain a less volatile version of the analysts' forecasts, that represents an estimation of the expected prices. In addition, Table 7 shows that the new expected spot price estimations fit the analysts' forecast better than futures for all maturities, meaning that there indeed is information hidden in our identified market variables that can be extracted to create simple risk premia estimations. In this way it appears to be possible to make a reasonable estimation of expected prices directly from market variables that match the analysts' predictions, however deeper analysis would be needed to come up with a robust methodology.

6. CONCLUSIONS

This paper develops a framework to extract time-varying commodity risk premia from multi-factor models using futures prices and analysts' forecasts of future spot prices. The model is calibrated for oil using a 3-factor stochastic commodity-pricing model with an affine risk premia specification with WTI futures data from NYMEX and analysts' forecasts from Bloomberg and the U.S Energy Information Administration from 2010 to 2017.

Figure 4a: 2 year ahead expected prices obtained adding the regression estimated risk premia to the observed futures prices (dotted line) in comparison with analysts' forecasts (dots) and futures prices (solid line).

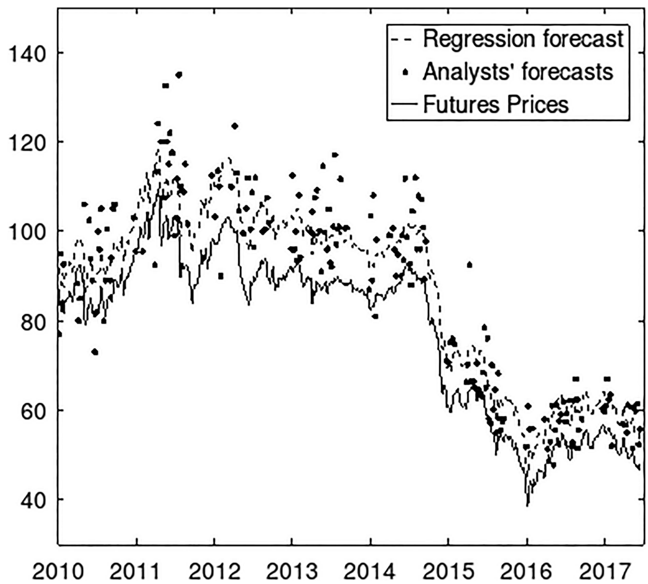
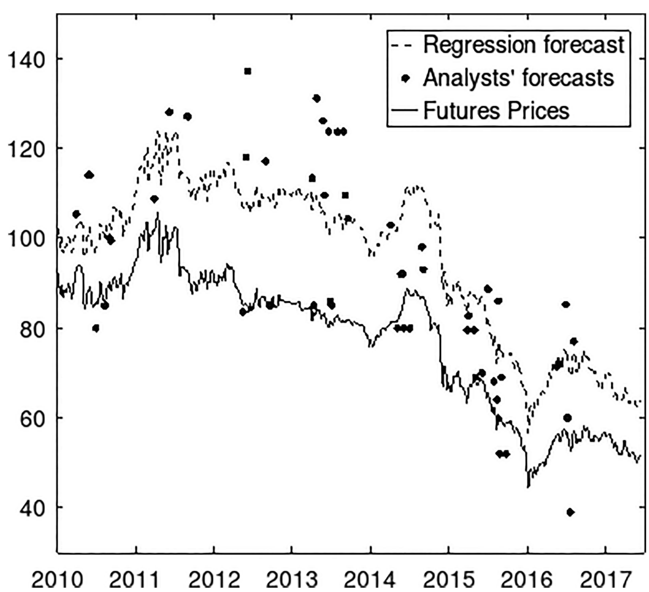


Figure 4b: 5 year ahead expected prices obtained adding the regression estimated risk premia to the observed futures prices (dotted line) in comparison with analysts' forecasts (dots) and futures prices (solid line).



Results from the model calibration show that risk premia are statistically significant and time varying, that short-term risk premia tend to be higher than long-term ones and that risk premia volatility is much higher for short maturities. These three findings give us a sense on how the market price expectations evolve over time and how they differ by maturity.

Table 7: MAPE between analysts' forecasts and two different expected price approaches: Futures and Futures plus Regression Market Risk Premia. Data between January 2010 and June 2017.

	M3	M6	M9	Y1	M18	Y2	Y5
Regression implied expectations	0.0561	0.0565	0.0608	0.0611	0.0684	0.077	0.1616
Futures prices	0.0601	0.0631	0.0771	0.0879	0.1124	0.1204	0.1772

We then use the term structures of oil risk premia obtained to perform an empirical analysis to explore the macroeconomic and oil market specific variables that may explain their stochastic behavior. We find that inventories, hedging pressure, term premium, default premium and the level of interest rates all play a significant role in explaining the risk premia and thus could be used also for estimating expected commodity prices when reliable analysts' forecasts are not available.

The macroeconomic determinants that were found through the regression analysis, may help countries that are heavily dependent on commodity prices define their public policies. For instance, the regression can be used to estimate simple commodity price forecasts that can be used as an input to the country's budget estimations.

REFERENCES

Achraya, Viral V., Lars A. Lochester, and Terun Ramadorai (2013). "Limits to arbitrage and hedging: Evidence from commodity markets." *Journal of Financial Economics* 109(2): 441–465. https://doi.org/10.1016/j.jfineco.2013.03.003.

Altavilla, C., R. Giacomini, and R Costantini (2014). "Bond returns and market expectations." *Journal of Financial Econometrics* 12(4): 708–729. https://doi.org/10.1093/jjfinec/nbu012.

Altavilla, Carlo, Raffaella Giacomini, and Guiseppe Ragusa (2017). "Anchoring the yield curve using survey expectations." *Journal of Applied Econometrics* 32(6): 1055–1068. https://doi.org/10.1002/jae.2588.

Baker, S.D. and B.R. Routledge (2017). "The price of oil risk." *Working Paper.*

Basu, Devraj and Joëlle Miffre (2013). "Capturing the Risk Premium of Commodity Futures: The role of Hedging Pressure." *Journal of Banking and Finance* 37(7): 2652–2664. https://doi.org/10.1016/j.jbankfin.2013.02.031.

Baumeister, Christiane and Lutz Kilian (2016). "A General Approach to Recovering Market Expectations from Future Prices With an Application to Crude Oil." *Working paper.*

Bernard, Jean-Thomas, Lynda Khalaf, Maral Kichian, and Clement Yelou (2018). "Oil Price Forecasts for the Long-Term: Expert Outlooks, Models, or Both?" *Macroeconomic Dynamics* 22(3): 581–599. https://doi.org/10.1017/S1365100516001279.

Bhar, Ramaprasad and Damien Lee (2011). "Time-Varying Market Price of Risk in the Crude Oil Futures Market." *The Journal of Futures Markets* 31(8): 779–807. https://doi.org/10.1002/fut.20493.

Bianchi, Daniele and Jacopo Piana (2018). "Adaptive Expectations and Commodity Risk Premia." *Working paper.* https://doi.org/10.2139/ssrn.2785563.

Casassus, Jaime and Pierre Collin-Dufresne (2005). "Stochastic Convenience Yield Implied from Commodity Futures and Interest Rates." *The Journal of Finance* 60(5): 2283–2331. https://doi.org/10.1111/j.1540-6261.2005.00799.x.

Chen, Long and Lu Zhang (2011). "Do time-varying risk premiums explain labor market performance." *Journal of Financial Economics* 99(2): 385–399. https://doi.org/10.1016/j.jfineco.2010.09.002.

Chernov, Mikhail and Philippe Mueller (2012). "The term structure of inflation expectations." *Journal of Financial Economics* 106(2): 367–394. https://doi.org/10.1016/j.jfineco.2012.06.004.

Chun, Albert Lee (2011). "Expectations, Bond Yields, and Monetary Policy." *Rewiev of Financial Studies* 24(1): 208–247. https://doi.org/10.1093/rfs/hhq090.

Cortazar, Gonzalo and Eduardo S. Schwartz (2003). "Implementing a stochastic model for oil futures prices." *Energy Economics* 25(3): 215–238. https://doi.org/10.1016/S0140-9883(02)00096-8.

Cortazar, Gonzalo and Lorenzo Naranjo (2006). "An N-factor Gaussian Model of Oil Futures Prices." *The Journal of Futures Markets* 26(3): 243–268. https://doi.org/10.1002/fut.20198.

Cortazar, Gonzalo, Cristobal Millard, Hector Ortega, and Eduardo S. Schwartz (2019). "Commodity Price Forecasts, Futures Prices and Pricing Models." *Management Science* 65(9): 4141–4155. https://doi.org/10.1287/mnsc.2018.3035.

Cortazar, Gonzalo, Ivo Kovacevic, and Eduardo S. Schwartz (2015). "Expected commodity returns and pricing models." *Energy Economics* 49: 60–71. https://doi.org/10.1016/j.eneco.2015.01.015.

Cortazar, Gonzalo, Lorenzo Naranjo, and Eduardo S. Schwartz (2007). "Term-structure Estimation in Markets with Infrequent Trading." *International Journal of Financial Economics* 12(4): 353–369. https://doi.org/10.1002/ijfe.317.

Cortazar, Gonzalo, Simon Gutierrez, and Hector Ortega (2015). "Empirical Performance of Commodity Pricing Models: When is it worthwhile to use a stochastic volatility specification." *The Journal of Futures Markets* 36(5): 457–487. https://doi.org/10.1002/fut.21740.

Cox, John C., Jonathan E. Ingersoll, and Stephen A. Ross (1981). "The relation between forward prices and future prices." *Journal of Financial Economics* 9(4): 321–346. https://doi.org/10.1016/0304-405X(81)90002-7.

Dai, Qiang and Kenneth J. Singleton (2002). "Expectation puzzles, time-varying risk premia, and affine models of the term structure." *Journal of Financial Economics* 63(3): 415–441. https://doi.org/10.1016/S0304-405X(02)00067-3.

Dai, Qiang and Kenneth J. Singleton (2000). "Specification analysis of affine term structure models." *Journal of Finance* 55(1): 1943–1978. https://doi.org/10.1111/0022-1082.00278.

De Roon, Frans A., Theo E. Nijman, and Chris Veld (2000). "Hedging Pressure Effects in Futures Markets." *The Journal of Finance* 55(3): 1437–1456. https://doi.org/10.1111/0022-1082.00253.

Dhume, Deepa (2010). "Using durable consumption risk to explain commodities returns." *Working paper.*

Diebold, Francis X. and Canlin Li (2006). "Forecasting the term structure of government bond yields." *Journal of Econometrics* 130(2): 337–364. https://doi.org/10.1016/j.jeconom.2005.03.005.

Dincerler, Cantekin, Zeigham Khokher, and Timothy Simin (2020). "An empirical analysis of commodity convenience yields." *Quarterly Journal of Finance* 10(2): 2050009. https://doi.org/10.1142/S2010139220500093.

Duarte, Jefferson (2004). "Evaluating an alternative risk preference in affine term structure models." *Review of Financial Studies* 17(2): 379–404. https://doi.org/10.1093/rfs/hhg046.

Duffee, Gregory R. (2002). "Term premia and interest rate forecasts in affine models." *The Journal of Finance* 57(1): 405–443. https://doi.org/10.1111/1540-6261.00426.

Erb, Claude B. and Campbell R. Harvey (2006). "The strategic and tactical value of commodity futures." *Financial Analysts Journal* 62(2): 69–97. https://doi.org/10.2469/faj.v62.n2.4084.

Estrella, Arturo and Gikas A. Hardouvelis (1991). "The term structureas a predictor of real economic activity." *The Journal of Finance* 46(2): 555–576. https://doi.org/10.1111/j.1540-6261.1991.tb02674.x.

Etula, Erkko (2013). "Broker-dealer risk appetite and commodity returns." *Journal of Financial Econometrics* 11(3): 486–521. https://doi.org/10.1093/jjfinec/nbs024.

Fama, Eugene F. and Kenneth R. French (1989). "Business condition and expected return on stocks and bonds." *Journal of financial economics* 25(1): 23–49. https://doi.org/10.1016/0304-405X(89)90095-0.

Fattouh, Bassam A, Lutz Kilian, and Lavan Mahadeva (2013). "The Role of Speculation in Oil Markets: What Have We Learned So Far?" *The Energy Journal* 34(3): 7–33. https://doi.org/10.5547/01956574.34.3.2.

Gorton, Gary B., Fumio Hayashi, and K. Geert Rouwenhorst (2013). "The fundamentals of commodity futures returns." *Review of Finance* 17(1): 35–105. https://doi.org/10.1093/rof/rfs019 .

Hamilton, James D. and Jing Cynthia Wu (2014). "Risk premia in crude oil futures prices." *Journal of International Money and Finance* 42: 9–37. https://doi.org/10.1016/j.jimonfin.2013.08.003.

Harvey, Campbell R. (1988). "The real term structure and consumption growth." *Journal of Financial Economics* 22(2): 305–333. https://doi.org/10.1016/0304-405X(88)90073-6.

Hicks, John Richard (1939). "Value and capital: an inquiry into some fundamental principles of economic theory." *Oxford University Press.*

Hong, Harrison and Motohiro Yogo (2012). "What does futures market interest tell us about the macroeconomy and asset prices?" *Journal of Financial Economics* 105(3): 473–490. https://doi.org/10.1016/j.jfineco.2012.04.005.

Kaldor, Nicholas (1939). "Speculation and economic stability." *The rReview of Economic Studies* 7(1): 1–27. https://doi.org/10.2307/2967593.

Kalman, Rudolf E. (1960). "A new approach to linear filtering and prediction problems." *Journal of Basic Engineering* 82: 35–45. https://doi.org/10.1115/1.3662552.

Kang, Wenjin, K. Greet Rouwenhorst, and Ke Tang (2019). "A tale of two premiums: the role of hedgers and speculators in commodity futures markets." *The Journal of Finance* 75(1): 377–417. https://doi.org/10.1111/jofi.12845.

Keim, Donald B. and Robert F. Stambaugh (1986). "Predicting returns in the stock and bond markets." *Journal of Financial Economics* 17(2): 357–390. https://doi.org/10.1016/0304-405X(86)90070-X.

Keynes, John M. (1930). "A treatise on money." Macmillan, London 2.

Khan, Saqib, Zeigham Khokher, and Timothy Simin (2008). "Expected commodity futures returns." Available at SSRN 1107377. https://doi.org/10.2139/ssrn.1107377.

Kim, Don H. and Athanasios Orphanides (2012). "Term structure estimation with survey data on interest rate forecasts." *The Journal of Financial and Quantitative Analysis* 47(1): 241–272. https://doi.org/10.1017/S0022109011000627.

Manoliu, Mihaela and Stathis Tompaidis (2002). "Energy futures prices: term structure models with kalman filter estimation." *Applied Mathematical Finance* 9(1): 21–43. https://doi.org/10.1080/13504860210126227.

Melolinna, Marko (2011). "What explains risk premiums in crude oil futures." *OPEC Energy Review* 35(4): 287–307. https://doi.org/10.1111/j.1753-0237.2011.00201.x.

Pagano, Patrizio and Massimiliano Pisani (2009). "Risk-adjusted forecasts of oil prices." *The B.E. Journal of Macroeconomics* 9(1): 67–82. https://doi.org/10.2202/1935-1690.1626.

Palazzo, Gerardo and Stefano Nobili (2010). "Explaining and fortecasting bond risk premiums." *Financial Analysts Journal* 66(4): 67–82. https://doi.org/10.2469/faj.v66.n4.5.

Pashke, Raphael and Marcel Prokopczuk (2009). "Integrating multiple commodities in a model of stochastic price dynamics." *Journal of Energy Markets* 2(3): 47–82. https://doi.org/10.21314/JEM.2009.025.

Ready, Robert C. (2018). "Oil Consumption, Economic Growth, and Oil Futures: The Impact of Long-Run Oil Supply Uncertainty on Asset Prices." *Journal of Monetary Economics* 94: 1–26. https://doi.org/10.1016/j.jmoneco.2017.09.010.

Sadorsky, Perry (2002). "Time-varying risk premiums in petroleum futures prices." *Energy Economics* 24(6): 539–556. https://doi.org/10.1016/S0140-9883(02)00062-2.

Singleton, Kenneth J. (2014). "Investor Flows and the 2008 Boom/Bust in oil prices." *Management Science* 60(2): 300–318. https://doi.org/10.1287/mnsc.2013.1756.

APPENDIX 1: ROTATION OF CORTAZAR AND NARANJO (2006)'S MODEL INTO OURS

Given the state space model of the form

$$Y_t = 1'x_t$$

$$dx_t = \left(-Ax_t + b\right)dt + \Sigma dw_t$$

where A and Σ are n×n diagonal matrices, b is a n×1 vector whose elements are zero excepting its first one and dw_t is an n×1 vector of correlated Brownian motions such that $dw_t dw_t' = \Theta dt$. The covariance matrix $\Sigma\Theta\Sigma'$ is positive definite and therefore admits a Cholesky decomposition. Let's define the matrix M as

$$M = \begin{bmatrix} 0 & \cdots & 0 & 1 \\ 0 & \cdots & 1 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 1 & \cdots & 0 & 0 \end{bmatrix}$$

where $M^{-1} = M$, then the matrix $M\Sigma\Theta\Sigma'M$ is still positive definite and still admits a Cholesky decomposition (L) so that

$$M\Sigma\Theta\Sigma'M = \mathrm{LL}'$$

then applying the transformation $\xi_t = ML^{-1}Mx_t$ where $ML^{-1}M$ is an upper triangular matrix

$$Y_t = \left(\vec{1}'MLM\right)\left(ML^{-1}Mx_t\right) = h'\xi_t$$

$$d\xi_t = \left(-\left(ML^{-1}MAMLM\right)\left(ML^{-1}Mx_t\right) + ML^{-1}Mb\right)dt + ML^{-1}M\Sigma dw_t = \left(-\hat{A}\xi_t + \hat{b}\right)dt + d\hat{w}_t$$

where h is an n×1 vector, $\hat{A}$ is an n×n upper triangular matrix whose first eigenvalue is zero, $\hat{b}$ is an n×1 vector with zeros in all its entries except for the first one and $d\hat{w}_t$ is an n×1 vector of uncorrelated brownian motions. This formulation is the one used by Dai and Singleton (2000) modified to hold for a matrix A with one zero valued eigenvalue by adding the $\hat{b}$ vector.

APPENDIX 2: EXPECTED VALUE AND COVARIANCES OF STATE VARIABLES

In this section we show how to get the expected value and covariances of the state variables of any model of the type

$$dx_t = \left(-Ax_t + b\right)dt + \Sigma dw_t$$

$$dw_t dw_t' = \Theta dt$$

Where dw_t are correlated Brownian motions with a correlation matrix given by $dw_t dw_t' = \Theta dt$. First we define the following state space vector

$$y_t = e^{At} x_t$$

and applying Ito's Lemma

$$dy_t = e^{At} dx_t + Ae^{At} x_t dt$$

$$dy_t = e^{At}\left(\left(-Ax_t + b\right)dt + \Sigma dw_t\right) + Ae^{At} x_t dt$$

$$dy_t = e^{At} b dt + e^{At} \Sigma dw_t$$

This last equation can be integrated as follows

$$\int_t^T y_s = \int_t^T e^{As} b ds + \int_t^T e^{As} \Sigma dw_s$$

$$y_T - y_t = \left(\int_t^T e^{As} ds\right) b + \int_t^T e^{As} \Sigma dw_s$$

$$x_T = e^{-A(T-t)} x_t + e^{-AT}\left(\int_t^T e^{As} ds\right) b + e^{-AT} \int_t^T e^{As} \Sigma dw_s$$

Now it is straightforward to obtain the expected value and the variance of the state space variables

$$E_t\left(x_T\right) = e^{-A(T-t)} x_t + \left(\int_0^{T-t} e^{-A\tau} d\tau\right) b$$

$$Cov_t\left(x_T\right) = \int_0^{T-t} e^{-A\tau} \Sigma \Theta \Sigma' \left(e^{-A\tau}\right)' d\tau$$

APPENDIX 3: METHOD TO AVOID NUMERICAL INTEGRATION

To get the expected values and covariances of the state variables as shown in Appendix 2 numerical integration seems to be necessary. Nevertheless, there is an alternative method shown by Pashke and Prokopczuk (2009) which does not need numerical integration but uses eigenvalues and eigenvectors of some matrices.

To solve for the expected value of the state variables of equation 13 first we decompose $A = UVU^{-1}$ where V is a matrix containing all A's eigenvalues in its diagonal and U is a matrix containing all its eigenvectors. It can be shown that $e^{-A\tau} = Ue^{-V\tau}U^{-1}$, where $e^{V(T-t)}$ is a diagonal matrix with $e^{v_i(T-t)}$ (where v_i is the *i*-th eigenvalue of matrix A) in its *i*-th position. It can be shown that

$$\int_0^{T-t} e^{-V\tau} d\tau = \begin{bmatrix} \frac{1-e^{v_1(T-t)}}{v_1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \frac{1-e^{v_n(T-t)}}{v_n} \end{bmatrix} = \phi$$

thus, the expected value of the state variables can be written as

$$E_t(x_T) = Ue^{V(T-t)}U^{-1}x_t + U\phi U^{-1}b$$

The variance shown in equation 14 can be calculated using the same properties as the expected value, so that

$$Cov_t(x_T) = U\int_0^{T-t} e^{-V\tau}U^{-1}U'^{-1}\left(e^{-V\tau}\right)' d\tau U' = UHU'$$

where H represents the integral just for ease of notation. As $e^{-V\tau}$ is a diagonal matrix containing $e^{-v_i\tau}$ in each of its diagonal elements, a closed form solution for the integral H can be obtained element-wise. To obtain the element in the *i*-th row and the *j*-th column of the matrix the next expression has to be evaluated

$$H_{i,j} = \int_0^{T-t} e^{-v_i\tau}\left[U^{-1}U'^{-1}\right]_{ij} e^{-v_j\tau} d\tau = \left[U^{-1}U'^{-1}\right]_{ij}\int_0^{T-t} e^{-(v_i+v_j)\tau} d\tau = \left[U^{-1}U'^{-1}\right]_{ij}\frac{1-e^{-(v_i+v_j)(T-t)}}{v_i+v_j}$$

APPENDIX 4: MODEL IMPLIED VOLATILITIES

First, let D be a function of the state variables and time. Its returns can then be modeled as

$$\frac{dD}{D} = \mu_D dt + \sigma_D dw_D$$

Applying Ito's Lemma we find that

$$\frac{dD}{D} = \frac{1}{D}\nabla D dx + \frac{1}{2}\frac{1}{D}\nabla D dx dx' \nabla D' + \frac{1}{D}\frac{dD}{dt}dt$$

where ∇ represents the jacobian operator. Replacing dx from Equation 2,

$$\frac{dD}{D} = \frac{\nabla D(-Ax+c) + \frac{1}{2}\nabla D \nabla D' + \frac{dD}{dt}}{D}dt + \frac{\nabla D}{D}dw_x$$

Additionally, it can be found that,

$$\left(\frac{dD}{D}\right)\left(\frac{dD}{D}\right) = \sigma_D^2 dt$$

which means that,

$$\sigma_D^2 = \frac{\nabla D \nabla D'}{D^2}$$

Now replacing D by the expected spot prices $E_t(S_T)$ calculated in Section 2.1 the Jacobian results in

$$\nabla E_t(S_T) = h' e^{-A(T-t)} E_t(S_T)$$

so that we can get the following structure for the expected spot's implied volatility

$$\sigma_{E(S)}^2 = h' e^{-A(T-t)} \left(e^{-A(T-t)}\right)' h$$

Following the same procedure for futures prices the Jacobian and the future prices' implied volatility respectively result in

$$\nabla F_t(T) = h' e^{-(A+\Lambda)(T-t)} F_t(T)$$

$$\sigma_F^2 = h' e^{-(A+\Lambda)(T-t)} \left(e^{-(A+\Lambda)(T-t)}\right)' h$$

IAEE

International Association for
ENERGY ECONOMICS

Membership in the International Association for Energy Economics is open to anyone worldwide who has an interest in the fields of energy or energy economics. Our membership consists of those working in both the public and private sectors including government, academic and commercial. Our current member base consists of 3900+ members in over 110 nations, with 28 nations having local affiliate organization.

We are an independent, non-profit, global membership organization for business, government, academic and other professionals concerned with energy and related issues in the international community. We advance the knowledge, understanding and application of economics across all aspects of energy and foster communication amongst energy concerned professionals.

We are proud of our membership benefit offerings, which include access to a rich library of energy economics related publications and proceedings as well as a robust line-up of webinars, podcasts and conferences. Learn more about the benefits of membership at: https://www.iaee.org/en/membership/benefits.aspx

In addition to traditional membership, we offer student and institutional memberships.