Time-Varying Term Structure of Oil Risk Premiums

GONZALO CORTAZAR
Pontificia Universidad Católica de Chile
gcortaza@ing.puc.cl

PHILIP LIEDTKE
Pontificia Universidad Católica de Chile
pliedtke@uc.cl

HECTOR ORTEGA
Pontificia Universidad Católica de Chile
hiortega@uc.cl

EDUARDO S. SCHWARTZ
Beedie School of Business, SFU
UCLA Anderson School of Management
and NBER
eduardos@sfu.ca

Time-Varying Term Structure of Oil Risk Premiums

Abstract

This paper proposes to extract time-varying commodity risk premiums from multi-factor models using futures prices and analyst´s forecasts of future prices. The model is calibrated for oil using a 3-factor stochastic commodity-pricing model with an affine risk-premium specification. WTI futures price data is from NYMEX and analyst´s forecasts from Bloomberg and the U.S Energy Information Administration. Weekly estimations for short, medium and long-term risk premiums between 2010 and 2017 are obtained. Results from the model calibration show that risk premiums are clearly stochastic, that short-term risk premiums tend to be higher than long-term ones and that risk premium 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 oil risk-premiums.

1. Introduction

Even though commodity risk premium is an important topic in financial economics, there is no consensus on its magnitude, behavior and appropriate estimation procedure (Baumeister and Kilian, 2016; Bianchi and Piana, 2017; de Roon, Nijman, and Veld, 2000; Melolinna, 2011; Palazzo and Nobili, 2010). Moreover, the recent *financialization* of commodity markets has increased its relevance for investors and strengthened arguments on its time-varying behavior (Hamilton and Wu, 2014; Baker and Routledge, 2017; Ready, 2016).

Understanding the stochastic behavior of commodity risk premiums 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 required for NPV calculations or risk management purposes. Third, it may shed light on some public policy implications by uncovering their macroeconomic determinants.

This paper provides a procedure for estimating the stochastic process of the term structure of commodity risk premiums by calibrating a multifactor model using analysts´ forecasts of future spot prices and futures contracts oil price data. Once time-varying oil risk premiums are obtained, an empirical analysis is performed to explore the main macroeconomic and oil market specific variables that explain their behavior.

There have been various attempts in the literature to estimate commodity risk premiums. Many practitioners and researchers use futures prices as proxies for market expectations (see Baumeister and Kilian (2016), Bianchi and Piana (2016)), implicitly assuming risk premiums are zero. But Keynes (1930) and Hicks (1939) had proposed in their theory of normal backwardation, 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 premium 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 addition, in recent years there has been some discussion on the impact of the post 2005 growth of commodity index fund traders on risk premiums (Hamilton and Wu, 2014; Singleton, 2014; Hong and Yogo, 2012; Stoll and Whaley, 2010; Irwin and Sanders, 2011; Ready, 2016).

Regardless of its increasing importance, there is no current consensus on how to estimate risk premiums and their stochastic behavior. In the last few years, different methods have been developed to extract risk premiums, 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 efforts have been oriented to commodities.

In what follows we present one way of characterizing existing methods for estimating risk premiums in commodity markets by classifying them into three approaches: Econometric, Economic, and Market.

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 premium is obtained by comparing this expected spot price with the futures price for the same maturity. Baumeister and Kilian (2016) extract expected spot prices from the historical payoffs of future contracts. They first calculate the payoff from different futures contracts as the difference between the futures price for a given maturity and the realized spot price at that date. Then, regressing the above payoffs on different set of variables, expected prices are obtained. Their results show that none of the sets of regressors used is capable of getting a lower MSPE than the Hamilton and Wu (2014) model detailed below when performing an out-of-sample analysis. However, given that realized future spot prices and current futures prices with same maturity are compared, the required data-sample gets bigger as longer-term risk premium are estimated.

In what we call the *economic* approach we include Hamilton and Wu (2014), Bianchi and Piana (2016), and Cortazar, Kovacevic, and Schwartz (2015). 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. For example, Hamilton and Wu (2014), following the normal backwardation theory of Keynes (1930), present a model in which hedgers sell futures contracts to hedge their risk and speculators and investors buy those futures contracts in order to maximize their utility function caring about the expected value and the variance of their future income. They find a change in behavior of commodity risk premiums before and after 2005 due to the *financialization* of commodity markets.

Bianchi and Piana (2016) argue against using realized risk premiums as they do not represent the ex-ante premiums if the spot prices are biased from their expectations. To directly capture the ex-ante risk premiums they create a model with adaptive learning to calculate expected spot prices for every date in the sample using only the past spot prices and the aggregate demand as input. Their model is based on the belief that investors learn from their mistakes predicting spot prices and their next predictions are therefore going to be influenced by their past prediction errors. They analyze the behavior of oil, copper, silver and corn, showing strong evidence on risk premia being time-varying.

Cortazar, Kovacevic, and Schwartz (2015) follows the extensive literature on no-arbitrage commodity pricing models that uses multifactor models to explain the time-series and cross section of futures prices (Gibson and Schwartz, 1990; Heston, 1993; Schwartz, 1997; Duffie, Pan, and Singleton, 2000; Schwartz and Smith, 2000; Cortazar and Schwartz, 2003; Casassus and Collin-Dufresne, 2005; Cortazar and Naranjo, 2006; Trolle and Schwartz, 2009; Cortazar and Eterovic, 2010; Bhar andLee, 2011; Chiang, Hughen, and Sagi, 2015). They argue that these models, being successful in fitting futures prices, provide very poor risk premium estimates. Therefore, they propose using an asset-pricing model instead of restricting some of their parameters. Asset-pricing models have been extensively applied to estimate commodity risk premiums, diverging on their approach and application, including the definition and number of risk factors, obtaining mixed results (Dusak, 1973; Bodie and Rosansky, 1980; Carter et al., 1983; Chang et al., 1990; Bessembinder and Chan, 1992; Bjorson and Carter, 1997; Erb and Harvey, 2006; Hong and Yogo, 2012; Dhume, 2010).

In what we call the *market* approach we include a recent paper by Cortazar, Millard, Ortega, and Schwartz (2018) 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 premiums 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 previously used in other contexts. For example, Chun (2011) shows that using GDP, inflation and other macroeconomic variables' survey forecasts adds important information, not fully incorporated in market prices, to interest rates prediction models and gives them a higher accuracy. Altavilla, Giacomini, and Ragusa (2016) develop a method in which interest rate predictions become more accurate using interest rate surveys.

This paper proposes to extract time-varying risk premium observations using the *market* approach by extending Cortazar et al. (2018) to allow for a stochastic specification of risk premiums. We propose a 3-factor model based on Cortazar and Naranjo (2006) and Dai and Singleton (2000), and consider an affine risk premium specification following Duffee (2002). The model is estimated with the Kalman Filter using WTI oil analysts´ forecasts of spot prices and futures contracts price data between 2010 and 2017. Analysts´ forecasts are provided by Bloomberg and the U.S. Energy Information Administration (EIA) for up to 25 years, and oil futures price data is obtained from the New York Mercantile Exchange (NYMEX) for maturities up to 10 years. This allows us to obtain weekly estimates for short, medium and long-term oil risk premiums and to analyze the market determinants of these premiums comparing them with previous findings in the literature. This analysis requires having time-varying risk premium estimates provided by our procedure and which were not available in the previous literature.

Once the term structures for oil risk premiums between 2010 and 2017 are computed, we explore the market determinants of those premiums. Following Bhar and Lee (2011) among others, we perform several regressions on different market variables that have been previously proposed in the literature. In this way, we provide some light on the determinants of risk premium variations and propose an adjustment to futures prices as a new simple way to estimate market expected prices.

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

2. The Model to Estimate Risk Premiums

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 premiums as in Duffee (2002). We propose calibrating this model using both futures prices and analyst´s forecasts to obtain a time-varying term structure of risk premiums[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 a time-varying risk premium in our constant-volatility model we resort to Duffee (2002) who shows how to use affine risk premiums in all types of Dai and Singleton (2000) canonical models, including the ones with non-stochastic volatility. Let RP_t be the commodity risk premium and assume that:

[1]This paper builds on Cortazar, Millard, Ortega, Schwartz (2018) which also used futures and analysts´ forecasts, but assumed constant risk premiums. That paper used the Cortazar and Naranjo (2006) N-factor model. In Appendix 1 we show that our proposed model is a rotated version of the Cortazar and Naranjo (2006) model.

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

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

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

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

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.

Notice 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+Λ matrix, while the latter only with matrix A. However, if the Λ matrix were set to zero, risk premiums would be a constant and not time-varying.

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

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

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

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

with[3]

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

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

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

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

[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 premiums.

[3]See Appendix 2

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

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

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

It can be shown[4] that Equations 9 and 10 have a closed form solution if matrix A+Λ 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].

Risk premiums may be defined as the return of the expected spot price over the future price. Let, $\pi_t(T-t)$ be the instantaneous risk premium at time t for T-t years ahead:

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

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

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

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 allow for measurement errors.

[4] See 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+Λ in eigenvalues and eigenvectors. See Appendix 3.

[6] See Appendix 4

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., 2018):

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

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

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. (2018) 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 premiums, futures prices and analyst´s forecasts for different dates and maturities are required. This section describes the data used.

3.1 Futures Contracts

WTI crude oil futures prices are obtained from the *New York Mercantile Exchange*. 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, separated in one-year buckets.

Table 1: Futures price observations between January 2010 and June 2017 by 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	77.8762	22.2808	113.7	26.55	0.4472	968
1-2	78.2296	19.4252	112.83	35.36	1.4795	795
2-3	77.5466	17.5891	109.33	38.66	2.4947	821
3-4	77.2896	16.4093	107.14	41.34	3.5103	783
4-5	77.39	15.7564	105.8	43.24	4.4861	786
5-6	77.4764	15.4024	105.56	44.42	5.4722	809
6-7	78.0038	15.2228	105.88	45.77	6.5043	767
7-8	78.1963	15.2019	106.3	46.5	7.4942	774
8-9	78.2701	15.7178	106.95	46.99	8.4316	635
9-10	77.1498	13.7851	95.16	55.08	9.0582	44

3.2 Survey Based Expected Prices

Since we assume that analysts' forecasts are noisy proxies for expected future spot prices, WTI's expected prices were collected from Bloomberg's analysts' predictions, a list of surveys done to professional analysts on the expected future commodity prices. The expectations are given quarterly 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 be available any day of the week. Each prediction is grouped on the oncoming 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 left out. Table 2 describes the forecast data used. The 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 separated by maturity bucket.

Maturity Bucket (years)	Mean Price ($/bbl.)	Price SD	Max Price ($/bbl.)	Min Price ($/bbl.)	Mean Maturity (years)	Number of Observations
0-1	81.0201	22.2566	122	35	0.5314	1118
1-2	85.619	21.2545	135	40	1.4296	808
2-3	89.0411	23.4254	189	44	2.4752	289
3-4	88.3383	23.0256	154	40	3.4448	239
4-5	86.2134	22.6815	150	38.5	4.4235	179
5-10	101.217	22.0268	152.96	60	6.2903	79
10-34	171.5592	34.2276	265.2	104.678	18.4838	134

4.- Results

This section presents the results of using WTI oil weekly data between January 2010 and June 2017 to calibrate the N-factor term structure 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.

Table 3: Parameter estimates for the 3-factor model. Data between January 2010 and June 2017. Significance levels are given by ***1%, **5% and *10%.

	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	-10.2143	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

Figure 1 shows the term structure (from 1 month to 10 years) of annualized risk premiums over the whole sample period (01/2010 to 06/2017). Three things are worth noting. First, risk premiums are clearly stochastic. Second, short-term risk premiums tend to be higher that long-term ones. Third, risk premium volatility is much higher for short maturities.

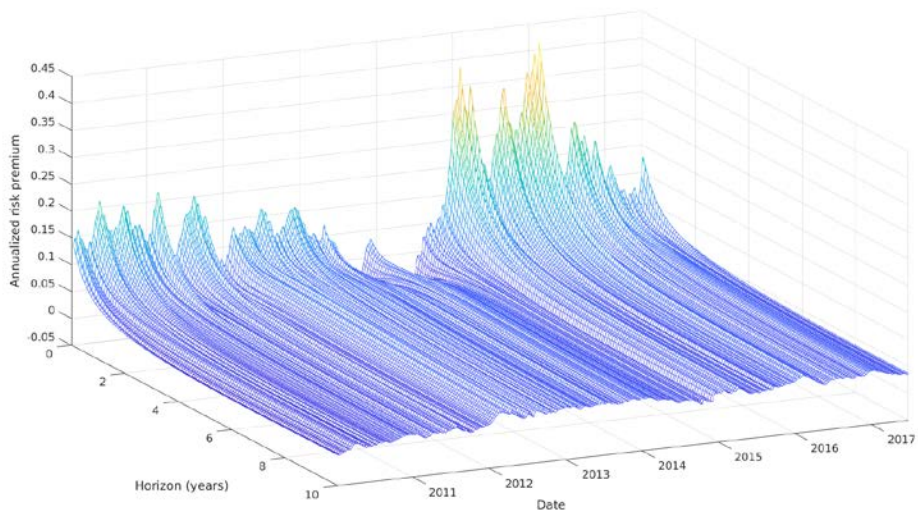
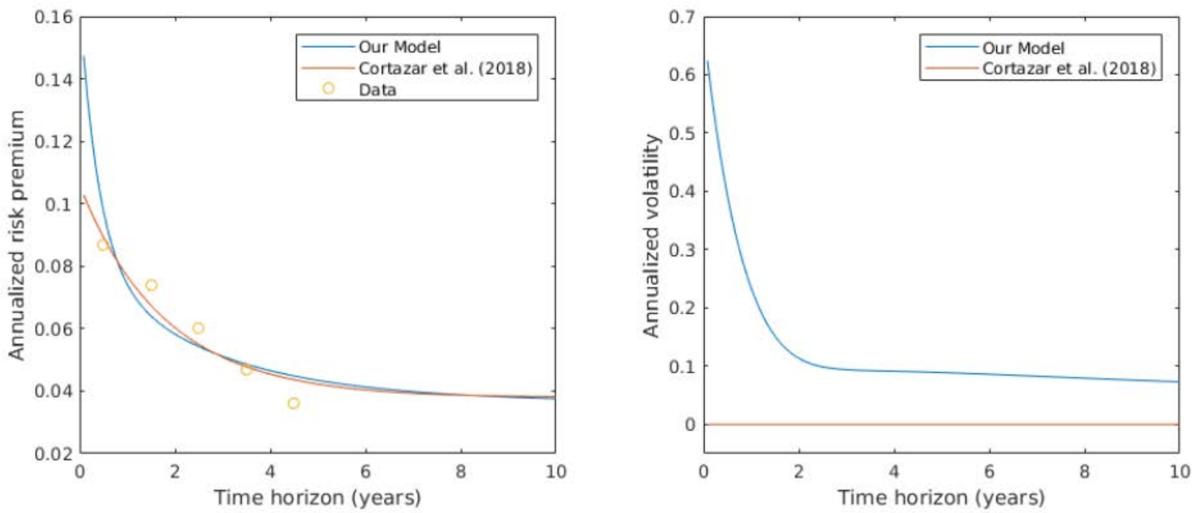


Figure 1: Annualized risk premium term structure from 1 month to 10 years. Data between January 2010 and June 2017.

Figure 2 analyzes the term structure mean and volatility of risk premiums. Figure 2a) compares our model´s mean risk premiums to those of Cortazar et al. (2018) constant risk premium model (with our same data) and to the data means. It can be noted that our model's mean risk premium level is similar to that of Cortazar et al. (2018) and both fit the data risk premiums well. Additionally, both premiums decrease with maturity.

Where both models diverge is Figure 2b) that shows the volatility term structure because by construction Cortazar et al. (2018)´s assumes constant risk premiums while an essential element of our model are time-varying risk premiums.

Finally, we analyze the goodness-of-fit of our model to futures and analysts´ forecasts data. Table 4 presents the mean absolute percentage error (MAPE) of our model and shows that its fit for both data sets is better than for the constant risk premium model in Cortazar et al. (2018).



(a) Mean risk premiums (b) Risk premium volatility

Figure 2: Mean risk premiums (a) and risk premium volatility (b) for our model and for the constant volatility model in Cortazar et al. (2018). Data mean risk premiums are also included in Figure (a). Data between January 2010 and June 2017.

Table 4: Mean Absolute Percentage Error (MAPE) for our time-varying risk premium model and for Cortazar et al. (2018) constant risk premium model. Data between January 2010 and June 2017.

	Our model	Cortazar et. al. (2018)
Futures prices	0.37%	0.39%
Expected prices	7.39%	8.00%

5.- The Determinants of Oil Risk Premiums

5.1 The Methodology

In this section, we explore the market determinants that may explain the variations of the estimated oil risk premiums. To do this we gather a set of market variables that have been previously reported in the literature as candidates for being related to risk premiums. We then perform a series of linear regressions in order to find which variables are the most significant in explaining the term structure of oil risk premiums.

There are few studies which analyze risk premiums directly (e.g. Bhar and Lee, 2011; Bianchi and Piana, 2016; Chen and Zhang, 2011; Melolinna, 2011) as most investigations only calculate them as a side result from a price prediction model. However, there is some literature that discusses the impact of different market variables on risk premiums which we review below.

The potential explanatory variables for the oils' risk premiums 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 term premium, the default premium and the 5-year treasury bill rate. These variables have been shown to include most of the risk factors taking part in the oil market as we explain below.

The **S&P500 index returns** is used in some studies (de Roon, Nijman, and Veld, 2000; Bianchi and Piana, 2016) as a proxy for the state of the US' economy which could affect oil risk premiums. Daily data is available in Bloomberg since 1950.

The **NASDAQ Emerging Markets Index (EMI)** represents the state of the emerging markets' economy. It is known that many big emerging economies, such as Russia or China, are important oil market players, hence their economic performance could directly affect oil prices and premiums. 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 premiums 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)'s *Theory of Storage,* which under a few assumptions implies that a rise in inventories should lead to a decrease in the overall risk premiums, and they find empirical results supporting their model. Weekly US WTI inventories starting at 1983 are available from the EIA and their percentage differences were calculated in order to obtain a stationary time series.

Open Interest (OI) and **Hedging Pressure (HP)** are the usual measures to represent the size and behavior of an instrument's market (in our case WTI futures). OI is measured as the total number of outstanding contracts, and therefore represents the market's size. It could be linked to the risk premiums as a larger amount of outstanding contracts could affect market's liquidity and therefore its premium. Kang, Rouwenhorst, and Tang (2017) propose 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, 2016; 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)'s and Hicks (1939)'s theories should have a strong correlation with risk premiums. According to them 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 premium will rise, because speculators will be willing to accept a greater amount of risk only if the premium is big enough. The relation between HP and prices or premiums has been empirically tested by different studies (Bianchi and Piana, 2016; de Roon, Nijman, and Veld, 2000; Gorton, Hayashi, and Rouwenhorst, 2013; Kang, Rouwenhorst, and Tang, 2017; 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, affect oil risk premiums. 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 yield. Daily treasury bill rates are available at the Federal Reserve while corporate bond yields were obtained from the Federal Reserve Bank of St. Louis.

The **5-year treasury bill rate (**5Y T-Bill) was used directly as it represents a good approach for a medium-term interest rate. Daily rates are available at the Federal Reserve.

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

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

where RP_{it} is the risk premium 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, a univariate regression is done for each independent regressor to check whether it is able to explain risk premiums in a statistically significant way. Then a multivariate regression analysis is performed using only the variables that were significang[7] in the univariate regressions. We run risk premiums regressions for 3, 6, 12, 18 months and 2, 5 and 10 years maturities. An independent regression is performed for every different time horizon, both in the univariate and 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 univariate regressions for each of the independent variables and maturities chosen. Inventories, HP, TRM, DEF and 5Y T-Bill have reasonable significance (p-value) to explain changes in oil risk premiums and are candidates for inclusion in the multivariate analysis, while the others are not.

Table 6 shows the results of multivariate regressions for each maturity using only the above variables. It can be noted that the R-Squared of the regressions vary between 47.61% and 60.10%, and all variables are significant for most of the maturities.

From the above tables several results are worth discussing.

First, we find a statistically significant and maturity-independent positive relation between inventories and risk premiums, similar to Dincerler, Khokher, and Simin (2005) and Khan, Khokher and Simin (2008). Our results are, however, contrary to Gorton et al. (2013)'s model which could be due to their assumptions not holding for our sample period.

Second, our statistical significance and positive value of the HP estimator over all studied maturities is backed up by Keynes (1930)' theory of normal backwardation, as a larger number of hedgers wanting to hedge their risk produces a greater HP which should by related to speculators demanding a larger premium to take on that risk. Basu and Miffre (2013), de Roon, Nijman, and Veld (2000) and Bianchi and Piana (2016), among others, obtain similar results.

[7]Meaning the variables that showed p-values under **5% or R squared of over 30%** for most maturities.

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

Finally, the 5Y T-Bill and DEF have a positive effect on risk premiums, however only for maturities up to two years. The relation between DEF and risk premiums was expected to be positive as the first one is highly correlated with the short-term market uncertainty, and should therefore affect risk premiums in a positive way. Higher short-term uncertainty should induce the average investor to demand a larger premium specially for short term investments which is consistent with DEF affecting only short term premiums in a significant way. 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 premiums, 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 premiums in our model for all studied maturities. In addition to the economic insight the regression results provide, they could also be used to obtain estimates of risk premiums 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.

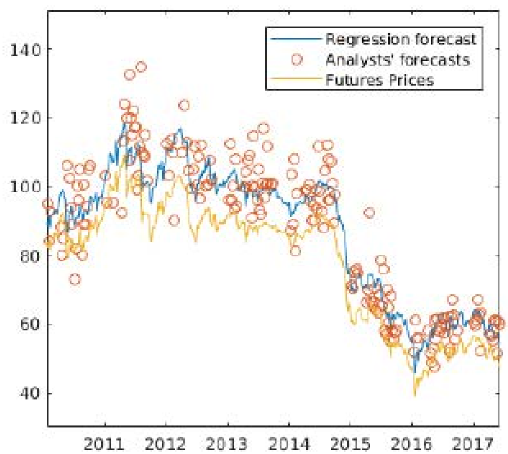
Figure 6 shows expected spot price estimations for two different maturities obtained by adding the expected risk premium from our regression analysis to the observed futures prices, along with analysts' forecasts and futures prices observations. The figure shows that by adding the risk premium to futures prices a less volatile estimate of expected prices is obtained. In addition, as Table 7 shows, this also increases its fit to analysts´ forecasts, reducing estimation errors.

Table 5: Univariate regression analysis for each of the chosen independent variables and for each different maturity. Monthly maturities are written as “Mn” and yearly maturities as “Yn”. Data between January 2010 and June 2017.

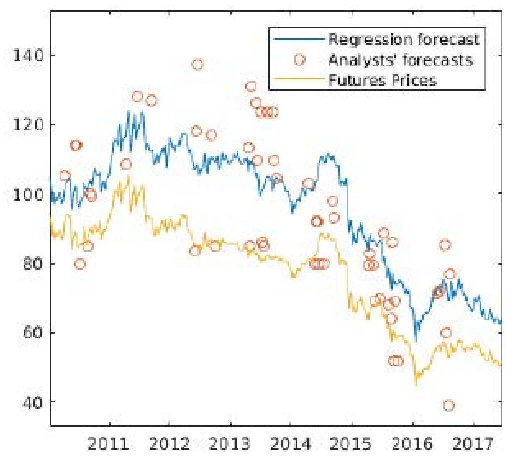
		M3	M6	Y1	M18	Y2	Y5	Y10
S&P500	Estimate	-0.3639	-0.2689	-0.1437	-0.0723	-0.0312	0.0224	0.0175
	p-value	0.0518	0.0586	0.0893	0.1859	0.4446	0.4826	0.5018
	r2	0.0072	0.0066	0.0049	0.0019	-0.0011	-0.0013	-0.0014
EMI	Estimate	-0.2225	-0.1641	-0.0877	-0.0446	-0.0201	0.0109	0.0095
	p-value	0.1315	0.1433	0.1888	0.3009	0.5319	0.6628	0.6434
	r2	0.0033	0.0029	0.0019	0.0002	-0.0016	-0.0021	-0.002
Inventories	Estimate	4.5019	3.5144	2.225	1.4986	1.0845	0.5104	0.4163
	p-value	0	0	0	0	0	0.0049	0.0051
	r2	0.043	0.0456	0.052	0.0568	0.0533	0.0177	0.0175
OI	Estimate	-0.051	-0.041	-0.0283	-0.0213	-0.0175	-0.0122	-0.0096
	p-value	0.7157	0.6998	0.6549	0.6013	0.5651	0.608	0.6242
	r2	-0.0022	-0.0022	-0.0021	-0.0019	-0.0017	-0.0019	-0.002
HP	Estimate	-0.0269	-0.0025	0.0299	0.0485	0.0592	0.0689	0.0567
	p-value	0.3916	0.9149	0.0347	0	0	0	0
	r2	-0.0007	-0.0025	0.0089	0.0701	0.1917	0.4299	0.4325
TRM	Estimate	-0.0363	-0.0305	-0.0228	-0.0182	-0.0156	-0.011	-0.0094
	p-value	0	0	0	0	0	0	0
	r2	0.0984	0.1208	0.1919	0.2977	0.3895	0.3174	0.3463
5Y T-Bill	Estimate	-0.0005	-0.003	-0.0066	-0.009	-0.0105	-0.0122	-0.0096
	p-value	0.9503	0.6117	0.0578	0.0001	0	0	0
	r2	-0.0026	-0.0019	0.0067	0.0383	0.0977	0.2201	0.1993
DP	Estimate	0.1332	0.0997	0.0559	0.0313	0.0174	-0.0006	-0.0009
	p-value	0	0	0	0	0	0.8105	0.6747
	r2	0.2051	0.1989	0.1768	0.1324	0.0721	-0.0024	-0.0021

Table 6: Multivariate regression coefficients for each maturity. Data between January 2010 and June 2017. Significance levels are given by ***1%, **5% and *10%.

	M3	M6	Y1	M18	Y2	Y5	Y10
Intercept	-0.0774***	-0.0506***	-0.0065	0.0212***	0.0358***	0.0461***	0.0395***
Inventories	2.6957***	2.1408***	1.4595***	1.0931***	0.8667***	0.4687***	0.3731***
HP	0.0784***	0.072***	0.0641***	0.0608***	0.0606***	0.0568***	0.0473***
TRM	-0.0693***	-0.0549***	-0.0355***	-0.0238***	-0.0167***	-0.0059***	-0.006***
5Y T-Bill	0.1083***	0.0828***	0.0485***	0.0284***	0.0166***	0.0002	0.0019
DEF	0.1827***	0.1377***	0.0784***	0.0448***	0.0259***	0.0008	0.001
R2	0.4761	0.49	0.5287	0.5727	0.601	0.534	0.5541



(a) 2 year ahead expected prices



(b) 5 year ahead expected prices

Figure 6: Expected prices obtained adding the regression estimated risk premiums to the observed futures prices (blue line) in comparison with analysts' forecasts (red dots) and futures prices (yellow line)

Table 7: MAPE between analysts' forecasts and two different expected price approaches: Futures and Futures plus Regression Market Risk Premium. 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

5.- Conclusions

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

Results from the model calibration show that risk premiums are clearly stochastic, that short-term risk premiums tend to be higher than long-term ones and that risk premium volatility is much higher for short maturities.

Once weekly term structures of oil risk premiums are obtained an empirical analysis to explore the macroeconomic and oil market specific variables that may explain their stochastic behavior is performed. 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 premium and thus could be used also for estimating expected commodity prices when reliable analyst's forecasts are not available.

References

Achraya, Viral V., Lars A. Lochstoer, and Tarun Ramadorai, 2013, Limits to arbitrage and hedging: Evidence from commodity markets, Journal of Financial Economics 109, 441–465.

Altavilla, Carlo, Rafaella Giacomini, and Riccardo Costantini, 2014, Bond returns and market expectations, Working paper –.

Altavilla, Carlo, Raffaella Giacomini, and Giuseppe Ragusa, 2016, Anchoring the yield curve using survey expectations, Working paper –.

Baker, Steven D., and Bryan R. Routledge, 2017, The Price of Oil Risk, Working paper.

Basu, Devraj, and Jolle Miffre, 2013, Capturing the Risk Premium of Commodity Futures: The Role of Hedging Pressure, Journal of Banking Finance 37, 2652–2664.

Baumeister, Christiane, and Lutz Kilian, 2016, A General Approach to Recovering Market Expectations from Future Prices With an Application to Crude Oil,

Bernard, Jean-Thomas, Lynda Khalaf, Maral Kichian, and Clement Yelou, 2015, Oil Price Forecast for the Long-Term: Expert Outlooks, Models, or Both?, Working paper .

Bessembinder, Hendrik, and Kalok Chan, 1992, Time-varying risk premia and forecastable returns in futures markets, Journal of Financial Economics 32, 169–193.

Bhar, Ramaprasad, and Damien Lee, 2011, Time-Varying Market Price of Risk in the Crude Oil Futures Market, The Journal of Futures Markets 31, 779–807.

Bianchi, Daniele, and Jacopo Piana, 2017, Expected Spot Prices and the Dynamics of Commodity Risk Premia, Working paper.

Bjornson, Bruce, and Colin Carter, 1997, New Evidence on Agricultural Commodity Return Performance under Time-Varying Risk, American Journal of Agricultural Economics 79, 918–930.

Bodie, Zvi, and Victor I. Rosansky, 1980, Risk and Return in Commodity Futures, Financial Analysts Journal 36, 27–39.

Carter, Colin A., Gordon C. Rausser, and Andrew Schmitz, 1983, Efficient Asset Portfolios and the Theory of Normal Backwardation, Journal of Political Economy 91, 319–331.

Casassus, Jaime, and Pierre Collin-Dufresne, 2005, Stochastic Convenience Yield Implied from Commodity Futures and Interest Rates, The Journal of Finance 60, 2283–2331.

Chang, Eric C., Chao Chen, and Son-Nan Chen, 1990, Risk and return in copper, platinum, and silver futures, The Journal of Futures Markets 10, 29–39.

Chen, Long, and Lu Zhang, 2011, Do time-varying risk premiums explain labor market performance, Journal of Financial Economics 99, 385–399.

Chernov, Mikhail, and Philippe Mueller, 2012, The term structure of inflation expectations, Journal of Financial Economics 106, 367–394.

Chiang, I-Hsuan Ethan, Keener Hughen, and Jacob S. Sagi, 2015, Estimating Oil Risk Factors Using Information from Equity and Derivatives Markets, The Journal of Finance 70.

Chun, Albert Lee, 2011, Expectations, Bond Yields, and Monetary Policy, Review of Financial Studies 24, 208–247.

Cortazar, Gonzalo, and Francisco Eterovic, 2010, Can oil prices help estimate commodity futures prices? The cases of copper and silver, Resources Policy 35, 283–291.

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 Future Markets 00, 1–31.

Cortazar, Gonzalo, Ivo Kovacevic, and Eduardo S. Schwartz, 2015, Expected commodity returns and pricing models, Energy Economics 49, 60–71.

Cortazar, Gonzalo, Cristobal Millard, Hector Ortega, and Eduardo S. Schwartz, 2018, Commodity Price Forecasts, Futures Prices and Pricing Models, Management Science, forthcoming.

Cortazar, Gonzalo, and Lorenzo Naranjo, 2006, An N-Factor Gaussian Model of Oil Futures Prices, The Journal of Futures Markets 26, 243–268.

Cortazar, Gonzalo, Lorenzo F. Naranjo, and Eduardo S. Schwartz, 2007, Term-structure estimation in markets with infrequent trading, International Journal of Finance and Economics 12, 353–369.

Cortazar, Gonzalo, and Eduardo S. Schwartz, 2003, Implementing a stochastic model for oil futures prices, Energy Economics 25, 215–238.

Cox, John C., Jonathan E. Ingersoll, and Stephen A. Ross, 1981, The relation between forward prices and futures prices, Journal of Financial Economics 9, 321–346.

Dai, Qiang, and Kenneth J. Singleton, 2000, Specification Analysis of Affine Term Structure Models, The Journal of Finance 55, 1943–1987.

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, 415–441.

de Roon, Frans A., Theo E. Nijman, and Chris Veld, 2000, Hedging Pressure Effects in Futures Markets, American Finance Association 55, 1437–1456.

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, 337–364.

Dinceler, Cantekin, Zeigham Khoker, and Timothy Simin (2005), “An Empirical Analysis of Commodity Convenience Yields,” University of Western Ontario, Working Paper.

Duarte, Jefferson, 2004, Evaluating an Alternative Risk Preference in Affine Term Structure Models, Review of Financial Studies 17, 379–404.

Duffee, Gregory R., 2002, Term Premia and Interest Rate Forecasts in Affine Models, The Journal of Finance 57, 405–443.

Duffie, Darrell, Jun Pan, and Kenneth Singleton, 2000, Transform Analysis and Asset Pricing for Affine Jump-Diffusions, Econometrica 68, 1343–1376.

Dusak, Katherine, 1973, Futures Trading and Investor Returns: An Investigation of Commodity Market Risk Premiums, Journal of Political Economy 81, 1387–1406.

Erb, Claude B., and Campbell R. Harvey, 2006, The Strategic and Tactical Value of Commodity Futures, Financial Analysts Journal 62, 69–97.

Estrella, Arturo, and Gikas A. Hardouvelis, 1991, The Term Structure as a Predictor of Real Economic Activity, The Journal of Finance 46, 555–576.

Etula, Erkko, 2013, Broker-Dealer Risk Appetite and Commodity Returns, Journal of Financial Econometrics 11, 486–521.

Fama, Eugene F., and Kenneth R. French, 1989, Business Condition and Expected Return on Stocks and Bonds, Journal of Financial Economics 25, 23–49.

Gibson, Rajna, and Eduardo S. Schwartz, 1990, Stochastic Convenience Yield and the Pricing of Oil Contingent Claims, The Journal of Finance 45, 959–976.

Gorton, Gary B., Fumio Hayashi, and K. Geert Rouwenhorst, 2013, The Fundamentals of Commodity Futures Returns, Review of Finance 17, 35–105.

Hamilton, James D., and Jing Cynthia Wu, 2014, Risk premia in crude oil futures prices, Journal of International Money and Finance 42, 9–37.

Harvey, Campbell R., 1988, The Real Term Structure and Consumption Growth, Journal of Financial Economics 22, 305–333.

Heston, Steven L., 1993, A Closed form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options, The Review of Financial Studies 6, 327–343.

Hicks, John Richard, 1939, Value and Capital: An inquiry into some fundamental principles of economic theory. Oxford, UK: Clarendon 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, 473–490.

Irwin, Scott H., and Dwight R. Sanders, 2011, Index Funds, Financialization, and Commodity Futures Markets, Applied Economic Perspectives and Policy 33, 1–31.

Khan, S., Z. Khokher, and T. Simin (2008), "Expected Commodity Futures Returns", Working Paper.

Kaldor, Nicholas, 1939, Speculation and Economic Stability, The Review of Economic Studies 7, 1–27.

Kalman, Rudolf E., 1960, A New Approach to Linear Filtering and Prediction Problems, Journal of Basic Engineering 82, 35–45.

Kang, Wenjin, K. Geert Rouwenhorst, and Ke Tang, 2017, A Tale of Two Premiums: The Role of Hedgers and Speculators in Commodity Futures Markets, Working Paper.

Keim, Donald B., and Robert F. Stambaugh, 1986, Predicting returns in the stock and bond markets, Journal of Financial Economics 17, 357–390.

Keynes, John M., 1930, A Treatise on Money, Macmillan, London 2.

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, 241–272.

Manoliu, Mihaela, and Stathis Tompaidis, 2002, Energy futures prices: term structure models with Kalman filter estimation, Applied Mathematical Finance 9, 21–43.

Melolinna, Marko, 2011, What Explains Risk Premiums in Crude Oil Futures, OPEC Energy Review 35, 287–307.

Pagano, Patrizio, and Massimiliano Pisani, 2009, Risk-adjusted forecasts of oil prices, B.E. Journal of Macroeconomics 9, 1–25.

Palazzo, Gerardo, and Stefano Nobili, 2010, Explaining and Forecasting Bond Risk Premiums, Financial Analysts Journal 66, 67–82.

Pashke, Raphael, and Marcel Prokopczuk, 2009, Integrating Multiple Commodities in a Model of Stochastic Price Dynamics, Journal of Energy Markets 2, 47–82.

Ready, Robert C., 2016, Oil Consumption, Economic Growth, and Oil Futures: The Impact of Long-Run Oil Supply Uncertainty on Asset Prices, Working paper.

Sadorsky, Perry, 2002, Time-varying risk premiums in petroleum futures prices, Energy Economics 24, 539–556.

Schwartz, Eduardo, and James E. Smith, 2000, Short-Term Variations and Long-Term Dynamics in Commodity Prices, Management Science 46, 893–911.

Schwartz, Eduardo S., 1997, The Stochastic Behavior of Commodity Prices: Implications for Valuation and Hedging, The Journal of Finance 52, 923–973.

Singleton, Kenneth J., 2014, Investor Flows and the 2008 Boom/Bust in Oil Prices, Management Science 60, 300–318.

Stoll, Hans R., and Robert E. Whaley, 2010, Commodity Index Investing and Commodity Futures Prices, Journal of Applied Finance 1, 7–46.

Trolle, Anders B., and Eduardo S. Schwartz, 2009, Unspanned Stochastic Volatility and the Pricing of Commodity Derivatives, The Review of Financial Studies 22, 4423–4461.

Appendix 1. Rotation of Cortazar and Naranjo (2006)'s model into our's

Given the state space model of the form

$$Y_t = 1' x_t$$

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

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. Lets 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

$$M\Sigma\Theta\Sigma' M = '$$

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

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

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

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 excepting 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 an covariances of the state variables of any model of the type

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

$$dw_t dw_t' = \Theta dt$$

Where dw 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 Itos lema

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

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

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

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 straight forward to obtain the expected value and the variance of the state space variables

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

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

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,

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

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

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

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

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

$$= [U^{-1}U'^{-1}]_{ij} \frac{1-e^{-(v_i+v_j)(T-t)}}{v_i+v_j}$$

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

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

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