Commodity Price Forecasts, Futures Prices and Pricing Models

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

CRISTOBAL MILLARD
Pontificia Universidad Católica de Chile
camillar@uc.cl

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

EDUARDO S. SCHWARTZ
Anderson Graduate School of Management UCLA
and NBER
eduardo.schwartz@anderson.ucla.edu

Commodity Price Forecasts, Futures Prices and Pricing Models

Abstract

Even though commodity-pricing models have been successful in fitting the term structure of futures prices and its dynamics, they do not generate accurate true distributions of spot prices. This paper develops a new approach to calibrate these models using not only observations of oil futures prices, but also analysts´ forecasts of oil spot prices.

We conclude that to obtain reasonable expected spot curves, analysts´ forecasts should be used, either alone or jointly with futures data. The use of both futures and forecasts, instead of using only forecasts, generates expected spot curves that do not differ considerably in the short/medium term, but long term estimations are significantly different. The inclusion of analysts´ forecasts, in addition to futures, instead of only futures prices, does not alter significantly the short/medium part of the futures curve, but does have a significant effect on long-term futures estimations.

1. Introduction

Over the last decades, commodity-pricing models have been very successful in fitting the term structure of futures prices and its dynamics. These models make a wide variety of assumptions about the number of underlying risk factors, and the drift and volatility of these factors. [Gibson, R. & Schwartz, E.S. (1990); Schwartz, E.S. (1997); Schwartz, E.S. & Smith, J. (2000); Cortazar, G. & Schwartz, E.S. (2003); Cortazar, G. & Naranjo, L. (2006); Cassasus, J. & Collin-Dufresne, P. (2005); Cortazar, G., & Eterovic, F. (2010); Heston, S. L. (1993); Duffie, D., J. Pan, & K. Singleton (2000); Trolle, A. B. & Schwartz, E. S. (2009); Chiang, I., Ethan, H., Hughen, W. K., & Sagi, J. S. (2015).]

The performance of commodity pricing models is commonly assessed by how well these models fit derivative prices. It is well known that derivative prices are obtained from the risk neutral or risk adjusted probability distribution (e.g. futures prices are the expected spot prices under the risk neutral probability distribution). These models also provide the true or physical distribution of spot prices, but this has not been stressed in the literature because they have mainly been used to price derivatives. However, as Cortazar, Kovacevic & Schwartz, (2015) point out, the latter is also valuable and is used by practitioners for risk management, NPV valuations, and other purposes[1].

Despite the diversity of commodity-pricing models found in the literature, they all share the characteristic of relying only on market prices (e.g. futures and options) to calibrate all parameters. In most of these models the risk premium parameters are measured with large errors and typically are not statistically significant, making estimations of expected prices (which differ from futures prices on the risk premiums) inaccurate. One exception is Hamilton & Wu (2014) who are able to get significant estimates through the use of a term structure of commodity futures prices model derived from the expected rational behavior of hedgers and speculators in commodity markets. Baumeister & Kilian (2016) show that this model is able to outperform any linear regression in its ability to predict future spot prices in a time horizon up to 1 year. Although their model appears to be the best alternative to forecast oil prices, the model only uses the three closest to maturity futures contracts as data input, which raises questions on how reliable can such a model be for longer maturities if it does not use any information of longer horizon prices.

In contrast to Hamilton & Wu (2014) we propose a term structure model which is capable of combining all available information in futures prices and survey price expectations obtaining statistically significant risk

---

[1] For instance, when mining and oil companies use Real Options to value their mines and oil deposits they need a model that has a good fit to the futures term structure. But these operations can last well beyond the time frame of existing futures contracts (20 and 30 years ahead)

premium parameters and credible short-, medium- and long-term risk premia and therefore reliable expected spot price forecasts. Even though our model's predictive power relies on the accuracy of the survey´s mean forecasts, we argue that they are the only market source available of future spot prices. If only futures prices are used in the estimation, no explicit information on the current risk premiums is incorporated into the model, not allowing it to obtain consistent risk premium estimates.

To solve this problem Cortazar et al. (2015) propose using an equilibrium asset pricing model (e.g. CAPM) to estimate the expected returns on futures contracts from which the risk premium parameters can be obtained, which results in more accurate expected prices. However, these prices depend on the particular asset pricing model chosen[2].

This paper develops an alternative way to estimate risk-adjusted and true distributions that does not rely on any particular equilibrium asset pricing model. The idea is to use forecasts of future spot prices provided by analysts and institutions who periodically forecast these prices, such as those available from Bloomberg and other sources. Thus, by calibrating the commodity pricing model with both futures prices and analysts' forecasts, two different data sets are jointly used to calibrate the model.

The use of survey forecasts as input for a no-arbitrage term structure model is not standard among researchers but seems to be promising. Even though there is no consensus on the amount of new information not captured already in market prices, their key feature is that, if accurate, they would be the most straightforward way of getting real time market's expectations.

It is well documented that surveys have been successful in predicting short-term macroeconomic variables, such as GDP, inflation and yields (Altavilla, Giacomini, & Ragusa (2016), Stark (2010), and Chun (2011)). Survey forecasts have also been used for predicting yield curves with promising results (Altavilla, Giacomini & Constantini (2014)), Chun (2011), Chernov & Mueller (2012), Dijk, Koopman, & Wel (2012), and Kim & Orphanides (2012)).

The use of survey forecasts in the oil market is somewhat more unusual (Alquist, Kilian, & Vigfusson (2013), Baumeister, Kilian, & Lee (2014), Sanders, Manfredo, & Boris (2009)). There are, however, several reputed sources of market's expectations' data available such as the Bloomberg's oil survey forecasts, the Energy Information Administration's (EIA), International Monetary Fund's (IMF) and World Bank's (WB), all of which we propose using in this study.

[2] Another approach is to link the risk premia with macroeconomic variables (Baker & Routledge (2011), Ready (2016)).

The Bloomberg's oil survey forecasts summarize a set of predictions made by different professional analysts who are specialized in commodity markets and hence have deep knowledge on the behavior of prices. This reduces potential biases, and addresses quality homogeneity issues and limited information processing (Bianchi & Piana (2016), Cutler, Poterba, & Summers (1990), Greenwood & Shleifer (2014), and Koijen, Schmeling, & Vrugt (2015)). Furthermore, the analysts incomes are generally directly related to the accuracy of their predictions, which also limits the incentive for hiding information for personal benefits (Bianchi & Piana (2016)). Although Bloomberg releases forecasts frequently (daily), the longest predicted price is for only 5 years ahead.

The U.S. Energy Information Administration (EIA) forecasts are generated using a model (NEMS) that captures various interactions that shape energy supply, demand and prices. It generates long-term estimations (over 25 years ahead) of the future states of several indicators of critical importance to the financial markets, such as WTI oil price forecasts, given a set of assumptions over the driving forces of these variables in different scenarios. Among these scenarios, the reference case scenario reflects the current central opinion of the leading economic forecasters and modelers. The reference case is widely used by practitioners and market players as a "best estimate" for market expectations, as explained by Bollinger et al. (2006) in the case of natural gas price forecasts[3]. These authors argue that, although these estimates may be subject to large errors, a significant and important set of market players, such as the natural gas utilities, rely on them to construct their resource planning models. Baumeister & Kilian (2014) reinforce this idea explaining that, even though they are not easy to replicate and justify given their long-term nature, they are widely used by practitioners and thus may be used as a good proxy for market´s expectations of long-term forecasts. Moreover, Auffhammer (2007) describes the EIA as the most important energy data source for the US, and that policymakers, industry and modelers extensively use them. Using similar arguments several other authors (Haugom et al (2016), Bianchi & Piana (2016), Berber & Piana (2006), Lee & Huh (2017)) have used these forecasts in their analyses.

The EIA source of information is unique, due to the very-long-term maturity of its forecasts. Since the forecasts are estimated for years in which there is not enough historic data to contrast to, it is difficult to verify their precision with high confidence. However, Bernard et al. (2015) show that EIA's forecasts are indeed accurate, especially for long-term predictions and, that combining EIA's forecasts with other forecasts improves the accuracy of them, for almost every forecast horizon.

[3] Natural gas price forecasts are generated using the same model (NEMS) as the WTI price forecasts submitted by the EIA.

Finally, the International Monetary Fund (IMF) and World Band (WB) price forecasts are made for up to 10 years ahead. These predictions rely on different macroeconomic models, datasets and approaches to market behavior.

Survey data, taken individually, may exhibit high prediction errors. By using data form different sources, some error diversification should occur. Our proposed model could be used with any set of forecasts, so as new research is able to confirm or reject the reliability of any given data source, the model could be updated easily to adjust to the new information.

In this paper, by proposing to use both market data (futures prices) and analysts' forecasts (proxy for expected spot prices) to calibrate a commodity-pricing model, several related objectives are pursued. The first one is to formulate a procedure using the Kalman filter methodology which includes both sets of data.

Acknowledging that analysts' price forecasts are very volatile, both because at any point in time there is great disagreement between them, and also because their opinions change greatly over time, our second objective is to build an analysts' consensus curve that optimally aggregates and updates all their opinions.

Our third objective is to improve estimations for long-term futures prices. This is motivated by current practice, which consists in calibrating commodity-pricing models using futures with maturities only up to a few years and then is silent about whether the model will behave well for longer maturities. However, there is evidence that extrapolating a model calibrated only with short/medium term prices to estimate long term ones is unreliable [Cortazar, Milla, & Severino (2008)]. In this paper, long-term futures price estimations will be obtained by using also information from long-term analysts' forecasts.

Finally, the fourth objective is to estimate the term structure of the commodity risk premiums. This can be done by comparing the term structure of expected spot and futures prices.

The paper is organized as follows. To motivate the proposed approach, Section 2 provides empirical illustrations of some of the weaknesses of current approaches. Section 3 describes the model and parameter estimation technique used, while Section 4 describes the data set. The main results of the paper are presented in Section 5. Section 6 concludes.

2. The Issues

In what follows, some of the issues that will be addressed in this paper are described. The first issue, already pointed out in Cortazar et al. (2015), is that expected prices under the true distribution are unreliable when calibrating a commodity-pricing model using only futures contract prices. As an illustration, **Figure 1** shows the futures and expected oil prices for 02-05-2014 using the Schwartz and Smith (2000) two-factor model. It can be seen that while the 4.5 year maturity futures price is 77.9 US$/bbl., the model's expected price, for the same maturity, is 365.8 US$/bbl. To justify that this expected price is unreasonable, the Bloomberg's Analysts´ Median Composite Forecast for 2018, which amounts to only 96.5 US$/bbl., is also plotted. While the model fits extremely well the term structure of futures prices, the expected spot prices it generates are clearly unreasonable.

Figure 2 shows the model expected spot prices, futures prices and analysts' forecasts for a contract maturing around 07-01-2018 during the year 2014. It can be seen that the model expected spot prices are for the whole year around three times higher than the futures prices and analysts' forecasts.

Given that we will make use of a diverse set of analysts' forecasts, a second issue is how to optimally generate and update an analysts' consensus curve, as new information arrives. **Figure 2** illustrates how the mean price forecasts for 2018 changes every week as new analysts provide their forecasts during 2014. It also shows that these forecasts are close to the corresponding futures prices, but the expected prices from the two-factor commodity model, when estimated using only futures, are much higher. Some efforts to provide an analysts' consensus curve have already been made (the Bloomberg Median Composite, also plotted in **Figure 2**), but in general they are computed using only simple moving averages of previous forecasts.

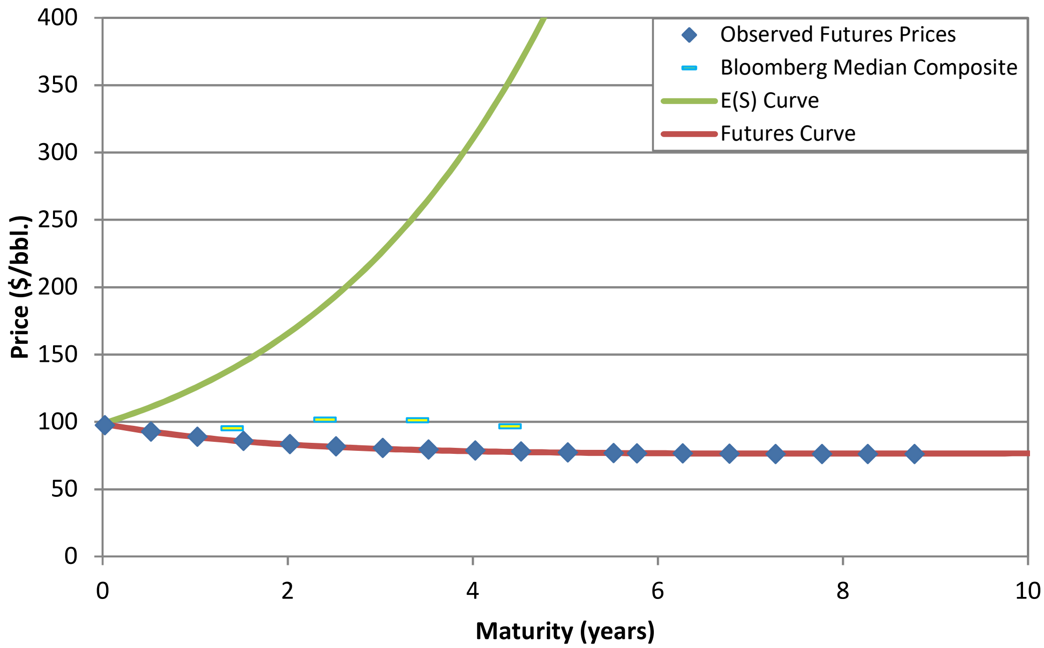

Fig. 1: Oil futures and expected spot curves under the Schwartz and Smith (2000) two-factor model, oil futures prices and Bloomberg's Median Composite for oil price forecasts, for 02-05-2014. The model is calibrated using weekly futures prices (01/2014 to 12/2014).

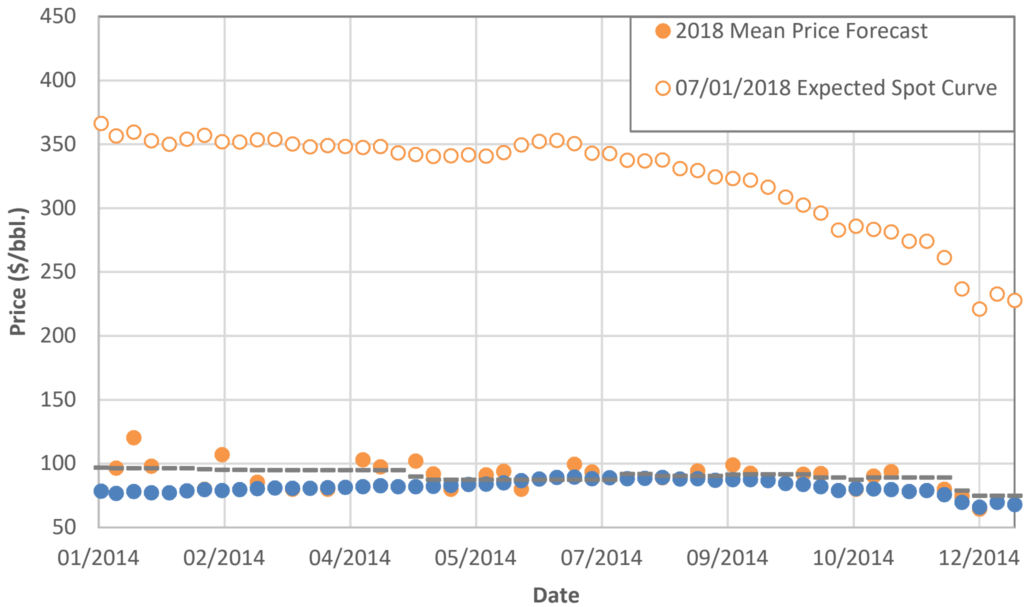

Fig. 2: Analysts' 2018 Oil Price Forecasts, Bloomberg Median Composite Forecast for 2018, Oil futures prices of contracts maturing close to 07-01-2018, and the Schwartz and Smith (2000) two-factor model expected spot at a 07-01-2018 maturity. The model is calibrated using weekly futures prices (01/2014 to 12/2014).

Another and related issue is how to obtain credible estimations of commodity risk premiums. When expected spot prices are unreliable, risk premiums are also unreliable.

The final issue that will be addressed is how to obtain long-term futures price estimations that exceed the longest maturity contract traded in the market, using the information contained in long-term analysts´ forecasts. Cortazar et al. (2008) already showed that extrapolations are unreliable: even if commodity-pricing models fit well existing data, contracts with longer maturities are estimated with large errors.

To illustrate the point discussed above, the Schwartz and Smith (2000) two-factor model is calibrated using three alternative data panels of oil futures: all futures including maturities up to 9 years, futures only up to 4.5 years, and futures only up to 2.25 years. For each data panel pricing errors for the longest observed futures price (around 9 years) are computed, finding that the longer the extrapolation, the higher the errors[4].

[4] Mean Absolute Errors were 0.9, 2.1 and 18.5$/bbl., respectively. Differences are significant at the 99% confidence level.

3. The Model

3.1. The N-Factor Gaussian Model

The Cortazar and Naranjo (2006) N-factor model[5] is used to illustrate the benefits of including analysts' forecasts, in addition to futures prices. This model nests several well-known commodity-pricing models (e.g. Brennan and Schwartz (1985), Gibson and Schwartz (1990), Schwartz (1997), Schwartz and Smith (2000), Cortazar and Schwartz (2003)). In the following sections the model will be implemented and results will be reported only for the 3-factor specification, though this model can be implemented for any number of factors.

Following Cortazar and Naranjo (2006), the stochastic process of the (log) spot price (S_t) of a commodity is assumed to be given by:

$$\log S_t = \mathbf{1}'\boldsymbol{x_t} + \mu t \quad (1)$$

where $\boldsymbol{x_t}$ is the $(1\ x\ n)$ vector of state variables and μ is the log-term price growth rate, assumed constant. The vector of state variables is assumed to follow the stochastic process:

$$d\boldsymbol{x_t} = -\boldsymbol{K}\boldsymbol{x_t}dt + \boldsymbol{\Sigma} d\boldsymbol{w_t} \quad (2)$$

where $\boldsymbol{K}$ and $\boldsymbol{\Sigma}$ are $(n\ x\ n)$ diagonal matrices containing positive constants (with the first element of $\boldsymbol{K}$, $\kappa_1 = 0$), and $d\boldsymbol{w_t}$ is a set of correlated Brownian motions such that $(d\boldsymbol{w_t})'(d\boldsymbol{w_t}) = \boldsymbol{\Omega} dt$, with each element of $\boldsymbol{\Omega}$ being $\rho_{ij} \in [-1,1]$. The risk adjusted process followed by the state variables is:

$$d\boldsymbol{x_t} = -(\boldsymbol{\lambda} + \boldsymbol{K}\boldsymbol{x_t})dt + \boldsymbol{\Sigma} d\boldsymbol{w_t}^Q \quad (3)$$

where $\boldsymbol{\lambda}$ is a $(1\ x\ n)$ vector containing the risk premium parameters corresponding to each risk factor, all assumed to be constants.

Under the N-Factor model, the futures price at time t, of a contract maturing at T, can be obtained by computing the conditional expected value of the spot price, under the risk-adjusted measure:

$$F(\boldsymbol{x_t}, t, T) = E_t^Q(S(\boldsymbol{x_t}, T)) \quad (4)$$

As shown in Cortazar and Naranjo (2006), this boils down to:

$$F(\boldsymbol{x_t}, t, T) = \exp(\boldsymbol{u}(t,T)'\boldsymbol{x_t} + v_F(t,T)) \quad (5)$$

---

[5] As shown in Cortazar and Naranjo (2006) the two-factor specification of this model is equivalent to the Schwartz and Smith (2000) model, but may easily be extended to N-factors.

where,

$$u_i(t,T) = e^{-\kappa_i(T-t)} \tag{6}$$

$$\begin{aligned} v_F(t,T) = \; &\mu t + \left(\mu - \lambda_1 + \frac{1}{2}\sigma_1{}^2\right)(T-t) - \sum_{i=2}^{n}\left(\frac{1-e^{-\kappa_i(T-t)}}{\kappa_i}\lambda_i\right) \\ &+ \frac{1}{2}\sum_{i\cdot j\neq 1}^{n}\left(\sigma_i\sigma_j\rho_{ij}\frac{1-e^{-(\kappa_i+\kappa_j)(T-t)}}{\kappa_i+\kappa_j}\right) \end{aligned} \tag{7}$$

Similarly, it can be shown that the expected spot price for time T at time t, is given by:

$$E_t(S(\boldsymbol{x_t},T)) = \exp(\boldsymbol{u}(t,T)'\boldsymbol{x_t} + \boldsymbol{v_E}(t,T)) \tag{8}$$

where,

$$v_E(t,T) = \mu T + \frac{1}{2}\sigma_1{}^2(T-t) + \frac{1}{2}\sum_{i\cdot j\neq 1}^{n}\left(\sigma_i\sigma_j\rho_{ij}\frac{1-e^{-(\kappa_i+\kappa_j)(T-t)}}{\kappa_i+\kappa_j}\right) \tag{9}$$

Note that the only differences between the futures and expected spot dynamics are the risk premium parameters. In addition, if these parameters were zero, the futures and expected spot prices would be equal. Define:

$$E_t(S(\boldsymbol{x_t},T)) = F(\boldsymbol{x_t},t,T) * e^{\pi_F\,(T-t)} \tag{10}$$

where π_F is the futures' risk premium, given by:

$$\pi_F = \lambda_1 + \sum_{i=2}^{n}\left(\frac{1-e^{-\kappa_i(T-t)}}{\kappa_i\,(T-t)}\lambda_i\right) \tag{11}$$

Finally, the model implied volatility (assumed constant in the time-series) is given by:

$$\sigma_F{}^2(\tau) = \sum_{i=1}^{n}\sum_{j=1}^{n}\sigma_i\sigma_j\rho_{ij}e^{-(\kappa_i+\kappa_j)\tau} \tag{12}$$

As mentioned earlier, in this paper analysts' forecasts are assumed to be noisy proxies for expected future spot prices.

3.2. Parameter Estimation

A Kalman filter that incorporates both futures prices and analysts' forecasts into the process of estimating all parameters is implemented. The Kalman Filter has been successfully used with incomplete data panels in commodities (Cortazar and Naranjo (2006)) and bond yields (Cortazar et al. (2007)), among others. Let's define m_t as the time-variant number of observations available at time t.

The application of the Kalman Filter requires two equations to be defined:

- The transition equation, which describes the true evolution of the $n \ x \ 1$ vector of state variables ($\boldsymbol{x_t}$) over each time step (Δt):

$$\begin{gathered}\boldsymbol{x_t} = \boldsymbol{A_t x_{t-1}} + \boldsymbol{c_t} + \boldsymbol{\varepsilon_t} \\ \boldsymbol{\varepsilon_t} \sim N(\boldsymbol{0}, \boldsymbol{Q_t})\end{gathered} \tag{13}$$

where $\boldsymbol{A_t}$ is a $n \ x \ n$ matrix, $\boldsymbol{c_t}$ is a $n \ x \ 1$ vector and $\boldsymbol{\varepsilon_t}$ is an $n \ x \ 1$ vector of disturbances with mean 0 and covariance matrix $\boldsymbol{Q_t}$.

- The measurement equation, which relates the state variables to the log of observed futures prices and analysts' forecasts:

$$\begin{gathered}\boldsymbol{z_t} = \boldsymbol{H_t x_t} + \boldsymbol{d_t} + \boldsymbol{v_t} \\ \boldsymbol{v_t} \sim N(\boldsymbol{0}, \boldsymbol{R_t})\end{gathered} \tag{14}$$

where $\boldsymbol{z_t}$ is a $m_t \ x \ 1$ vector, $\boldsymbol{H_t}$ is a $m_t \ x \ n$ matrix, $\boldsymbol{d_t}$ is a $m_t \ x \ 1$ vector and $\boldsymbol{v_t}$ is a $m_t \ x \ 1$ vector of disturbances with mean 0 and covariance matrix $\boldsymbol{R_t}$.

An additional complication is that analysts provide their price forecasts as an annual average, instead of a price for every maturity, as is the case for futures. Thus, Equations (5) and (8) become

$$\log F(\boldsymbol{x_t}, t, T) = \boldsymbol{u}(t, T)' \boldsymbol{x_t} + v_F(t, T) \tag{15}$$

$$\log E_t(S(\boldsymbol{x_t}, T)) = \log\left(\frac{1}{N_P} \sum_{i=1}^{N_P} \exp(\boldsymbol{u}(t, T)' \boldsymbol{x_t} + v_E(t, T))\right) \tag{16}$$

Notice that in order to measure the analysts' forecast observations we numerically approximate the mean annual price as the mean of N_P observations evenly spaced over the same year of the estimation. As can be observed, unlike futures prices, price forecasts are not a linear function of the state variables.

In order for expected spot prices to be normally distributed, under the N-Factor model, the $\log E(S)$ must be represented by a linear combination of the state variables. This can be achieved by linearizing the measured $\log E_t(S(\boldsymbol{x_t}, T))$ when computing each measurement step of the Kalman Filter[6].

If m_t^F and m_t^E are the number of observations of futures prices and analysts' forecasts at time t, the matrices corresponding to the measurement equation are:

$$\mathbf{z_t} = \begin{pmatrix} \mathbf{z_t}^{\boldsymbol{F}} \\ \mathbf{z_t}^{\boldsymbol{E}} \end{pmatrix} \tag{17}$$

where $\mathbf{z_t}^{\boldsymbol{F}}$ is a $m_t^F \times 1$ vector containing the futures observations and $\mathbf{z_t}^{\boldsymbol{E}}$ is a $m_t^E \times 1$ vector containing the price forecasts observations.

Let

$$\mathbf{H_t} = \begin{pmatrix} \mathbf{H_t}^{\boldsymbol{F}} \\ \mathbf{H_t}^{\boldsymbol{E}} \end{pmatrix} \tag{18}$$

and

$$\mathbf{d_t} = \begin{pmatrix} \mathbf{d_t}^{\boldsymbol{F}} \\ \mathbf{d_t}^{\boldsymbol{E}} \end{pmatrix} \tag{19}$$

where $\mathbf{H_t}^{\boldsymbol{F}}$ is a $m_t^F \times n$ matrix and $\mathbf{d_t}^{\boldsymbol{F}}$ is a $m_t^F \times 1$ vector containing the measurement equations for the futures data and $\mathbf{H_t}^{\boldsymbol{E}}$ is a $m_t^E \times n$ matrix and $\mathbf{d_t}^{\boldsymbol{E}}$ is a $m_t^E \times 1$ vector containing the linearized measurement equations for the price forecasts data.

Finally,

$$\mathbf{R_t} = \begin{pmatrix} \boldsymbol{R_t}^{\boldsymbol{F}} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{R_t}^{\boldsymbol{E}} \end{pmatrix} \tag{20}$$

where $\boldsymbol{R_t}^{\boldsymbol{F}} = diag_{m_t^F}(\xi^F)$ and $\boldsymbol{R_t}^{\boldsymbol{E}} = diag_{m_t^E}(\xi^E)$ are the diagonal covariance matrices of measurement errors of futures and price forecasts observations.

[6] More information on this methodology can be found in Cortazar, Schwartz & Naranjo (2007).

4. The Data

4.1. Analysts' Price Forecasts Data

As mentioned in the introduction analysts´ price forecasts are obtained from four sources: Bloomberg, World Bank (WB), International Monetary Fund (IMF) and the U.S. Energy Information Administration (EIA).

The first source is the Bloomberg Commodity Price Forecasts. This data base provides information on the mean price of each following year, up to 5 years ahead, made by individual analysts from a wide range of private financial institutions. Even though the data has not been analyzed extensively in the literature, it has been recently recognized as a rich and unexplored source of information [Berber and Piana (2016), Bianchi and Piana (2016)].

The next three sources (WB[7], IMF[8], and EIA[9]), provide periodic (monthly, quarterly or annually) reports with long-term, annual mean price estimations up to 28 years ahead. Most historical data is available since 2010. Among these three sources, the last one has received more attention in the literature. In particular, Berber & Piana (2016) and Bianchi & Piana (2016) use it for oil inventory forecasts, while Bolinger et al. (2006), Auffhammer (2007), Baumeister & Kilian (2015) and Haugom et al. (2016) focus on price forecasts[10]. Finally, Auffhammer (2007) and Baumeister & Kilian (2015) claim this source is widely used by policymakers, industry and modelers.

Figure 3 shows the analysts' price forecasts from all four sources, between 2010 and 2015. It can be seen that short-term forecasts are more frequent, in contrast to long-term forecasts, which are issued in a less recurring, but periodical, basis.

[7] Issued in the Commodity Markets Outlook.

[8] Issued in the Medium Term Commodity Price Baseline.

[9] Issued in the Annual Energy Outlook.

[10] It must be noted that some of the forecast analysis is only in-sample.

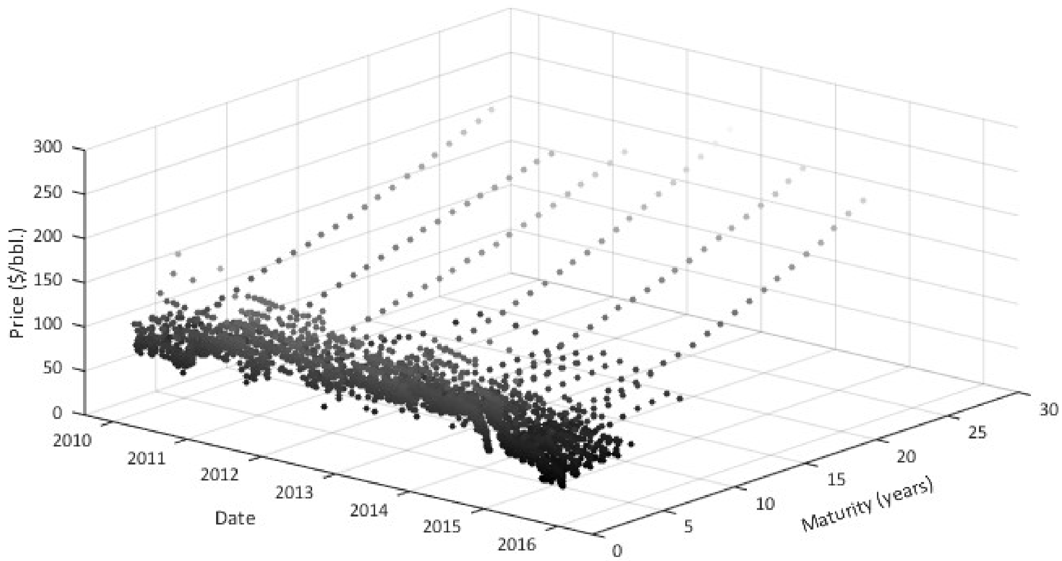

Fig. 3: Oil analysts' price forecasts from 2010 to 2015 provided by Bloomberg's Commodity Price Forecasts, World Bank (WB), International Monetary Fund (IMF) and U.S. Energy Information Administration (EIA).

We use analysts' price forecasts that are made for the average of each year without including the current one. Forecasts for the same year (like quarterly forecasts, for example), which include past information, are discarded as in Bianchi & Piana (2016). Thus, for each forecast its maturity is computed as the difference (in years) between the issue date and the middle of the year of the estimation (July, 1st of each year). Price forecasts are grouped into weeks ending on the following Wednesday, and then averaged[11]. **Table 1** summarizes the data.

4.2. Oil Futures Data

Oil futures price data is obtained from the New York Mercantile Exchange (NYMEX). Weekly futures (Wednesday closing), with maturities for every 6 months, are used. There are from 17 to 19 contracts per week. Futures data is much more frequent than analysts' forecasts, as can be seen by comparing **Figures 3** and **4**. **Table 2** summarizes the futures data by maturity buckets with similar number of observations.

[11] This is similar to what Berber and Piana (2016) or Bianchi and Piana (2016) do when averaging forecasts corresponding to the same period of estimation.

Table 1: Oil analysts' price forecasts from 2010 to 2015 grouped by maturity bucket. Forecasts are aggregated by week ending in the next Wednesday and averaged to obtain the mean price estimate for each following year in the same week.

Maturity Bucket (years)	Mean Price ($/bbl.)	Price S.D.	Mean Maturity (years)	Min. Price ($/bbl.)	Max. Price ($/bbl.)	N° of Observations
0-1	88.4	17.5	0.8	47.2	117.5	149
1-2	93.9	16.6	1.5	52.3	135.0	284
2-3	96.8	19.2	2.5	50.9	189.0	236
3-4	95.5	20.1	3.5	51.5	154.0	190
4-5	93.0	19.7	4.5	52.0	140.0	141
5-10	99.1	18.2	6.7	61.2	153.0	122
10-28	165.9	40.6	16.9	80.0	265.2	110
Total	**100.9**	**29.6**	**4.1**	**47.2**	**265.2**	**1232**

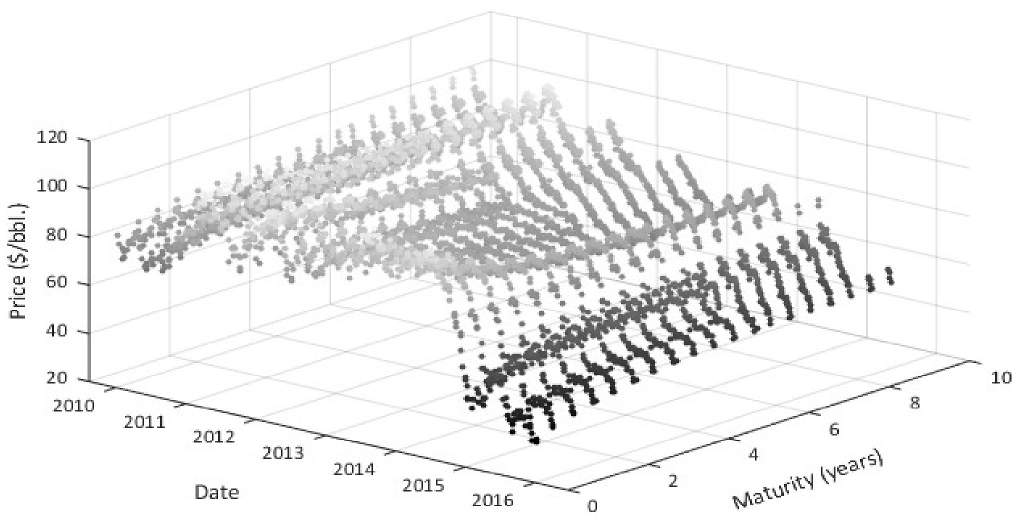

Fig. 4: Oil futures prices from 2010 to 2015 provided by NYMEX.

Table 2: Oil futures prices from 2010 to 2015 grouped by maturity bucket.

Maturity Bucket (years)	Mean Price ($/bbl.)	Price S.D.	Mean Maturity (years)	Min. Price ($/bbl.)	Max. Price ($/bbl.)	N° of Observations
0-1	85.4	17.7	0.4	36.6	113.7	786
1-2	85.0	14.5	1.5	45.4	110.7	621
2-3	84.0	12.7	2.5	48.5	107.9	625
3-4	83.5	11.6	3.5	50.9	106.2	627
4-5	83.4	11.0	4.5	52.5	105.6	631
5-6	83.5	10.8	5.5	53.5	105.6	622
6-7	83.8	10.9	6.5	54.2	105.9	625
7-8	84.1	11.1	7.5	54.6	106.3	626
8-9	84.6	11.6	8.4	54.9	107.0	461
Total	**84.2**	**12.8**	**4.2**	**36.6**	**113.7**	**5624**

4.3. Risk Premiums Implied from the Data

As explained in **Section 3.1**, empirical risk premiums can be derived directly form the data by comparing analysts' forecasts with futures prices of similar maturity[12]. Since oil futures contracts longest maturity does not exceed 9 years, it is not possible to calculate the data risk premiums exceeding this term. Then, if $\widehat{E_t}(S_T)$ is a price forecast at time t, for maturity T, and $F_{t,\hat{T}}$ is its closest futures (in maturity) for the same date, following **Equation 10** the data risk premium corresponding to that time is computed as:

$$\pi_{t,T} = \frac{\log\left(\frac{\widehat{E_t}(S_T)}{F_{t,\hat{T}}}\right)}{T} \tag{21}$$

The mean data risk premiums for each maturity bucket is presented in Table 3. Notice that the annual data risk premium is decreasing with maturity.

[12] Forecasts with more than one year of difference with the nearest future contract are not used to calculate data risk premiums.

Table 3: Mean Annual Data Risk Premium from 2010 to 2015 by maturity bucket.

Maturity Buckets (years)	Mean Data Risk Premium (%)
0.5 – 1.5	7.6%
1.5 – 2.5	6.7%
2.5 – 3.5	5.2%
3.5 – 4.5	3.3%
4.5 – 5.5	2.9%
5.5 – 6.5	3.2%
6.5 – 7.5	3.2%
7.5 – 8.5	3.1%
8.5 – 9.5	3.0%

5. Results

This section presents the results from calibrating the Cortazar and Naranjo (2006) N-factor model, described in Section 3, using a 3-factor[13] specification and different calibration data. In terms of the calibration data, two sets are available: futures prices (F) and analysts' forecasts (A). Results using jointly both data sets (FA-Model), only-analysts' data (A-Model), and the traditional only-futures data (F-Model), are presented. The behavior of the futures curve, the expected spot price curve and the risk premiums are analyzed.

5.1. Joint Model Estimation (FA-Model)

The Joint Model estimation, FA-Model, uses both the analysts' price forecasts and futures data to calibrate the 3-Factor Model. To motivate the discussion, **Figures 5** and **6** illustrate the results for the futures and expected spot curves, under different calibrations, for two specific dates, when futures are in contango (04-14-2010) and when futures are in backwardation (07-09-2014). Notice that in all cases the curves fit reasonably well the futures prices and analysts' forecasts observations when using the FA-Model. On the contrary, when using the traditional F-Model, the expected price curves are well below the analysts' forecasts.

[13] Results for a 2-factor specification are qualitatively similar to those for the 3-factor model, but are not reported.

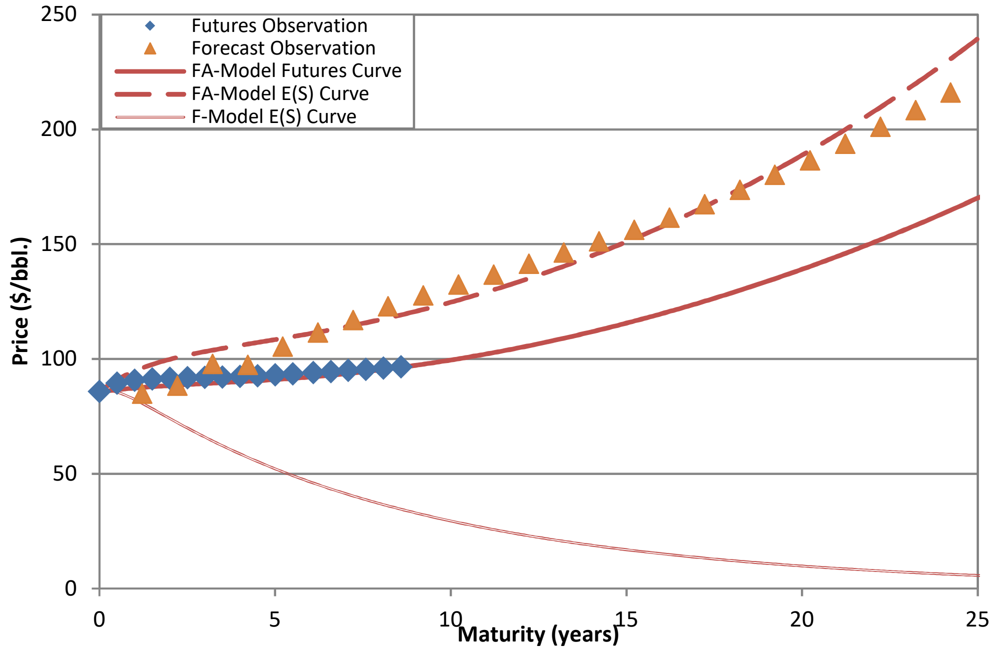

Fig. 5: Futures, expected spot curves and observations for 04-14-2010. Curves include FA- and F- Models. Parameter estimation from 2010 to 2014.

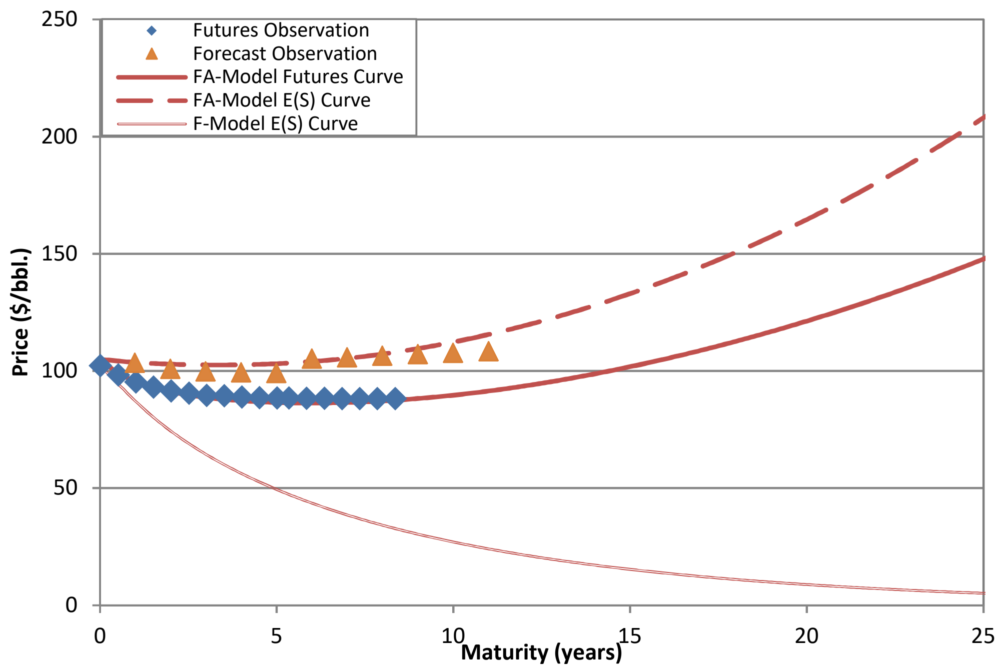

Fig. 6: Futures, expected spot curves and observations for 07-09-2014. Curves include FA- and F- Models. Parameter estimation from 2010 to 2014.

As discussed previously, the F- and FA-Models estimate both the true and the risk-adjusted distributions, from which futures prices and expected spot prices can be obtained. Analysts' forecasts and Futures pricing errors for both models are computed and presented in **Tables 4** and **5.** Parameter values obtained using the Kalman filter and using weekly data from 2010 to 2014 are reported in the Appendix. It is worth reporting that by using this new FA approach most risk premium parameters become now statistically significant. Notice that since the parameters of the model are estimated with data from 2010 to 2014, the results for 2015 are out of sample.

Table 4: Price forecasts Mean Absolute Errors for the F- and FA-Models for each maturity bucket and time window, between 2010 and 2015. Errors are calculated as percentage of price forecasts. Parameter estimation from 2010 to 2014.

Buckets	N° of	In Sample (2010-2014)		Out of Sample (2015)		Total (2010-2015)	
(years)	Observations	F-Model	FA-Model	F-Model	FA-Model	F-Model	FA-Model
0-1	149	11,7%	3,8%	15,6%	4,3%	12,3%	3,9%
1-2	284	22,0%	4,8%	22,3%	5,0%	22,1%	4,9%
2-3	236	33,9%	7,1%	31,0%	6,7%	33,3%	7,1%
3-4	190	41,1%	9,4%	39,3%	7,5%	40,7%	8,9%
4-5	141	47,5%	10,8%	46,4%	7,9%	47,2%	10,0%
5-10	122	61,6%	7,7%	59,6%	4,5%	61,1%	6,9%
10-28	110	90,2%	5,4%	65,7%	4,4%	86,2%	5,2%
Total	**1232**	**38,6%**	**6,8%**	**37,6%**	**6,0%**	**38,4%**	**6,6%**

Table 4 shows the mean absolute errors between analysts' forecasts and model expected spot prices generated by the FA-Model versus the F-Model. It is clear that the FA-Model has a significantly better fit for all time windows and buckets.

Furthermore, **Table 5** shows the mean absolute errors between observed futures prices and model futures prices. As expected, the benefit of obtaining a better fit in the expected spot prices, by including analysts' forecasts, comes at the expense of increasing the mean absolute error on the futures prices. Nevertheless, the error increase is only 1%.

Table 5: Futures Mean Absolute Errors for the F- and FA-Models for each maturity bucket and time window, between 2010 and 2015. Errors are calculated as percentage of futures prices. Parameter estimation from 2010 to 2014.

Buckets (years)	N° of Observations	In Sample (2010-2014)		Out of Sample (2015)		Total (2010-2015)	
		F-Model	FA-Model	F-Model	FA-Model	F-Model	FA-Model
0-1	786	0,5%	1,8%	0,8%	2,4%	0.6%	1.9%
1-2	621	0,4%	1,7%	1,1%	1,8%	0.5%	1.7%
2-3	625	0,2%	1,6%	0,4%	1,4%	0.3%	1.6%
3-4	627	0,3%	1,5%	0,7%	1,5%	0.4%	1.5%
4-5	631	0,4%	1,3%	1,0%	1,8%	0.5%	1.3%
5-6	622	0,3%	1,1%	1,0%	1,9%	0.4%	1.2%
6-7	625	0,2%	1,0%	0,4%	1,3%	0.2%	1.1%
7-8	626	0,2%	1,0%	0,7%	1,3%	0.3%	1.1%
8-9	461	0,4%	1,2%	1,7%	2,6%	0.6%	1.4%
Total	**5624**	**0,3%**	**1,4%**	**0,8%**	**1,8%**	**0.4%**	**1.4%**

In summary, the FA-Model has the advantage of generating a more reliable expected spot curve, with only a moderate effect for the goodness of fit for the futures.

5.2. Analysts' Consensus Curve using only Analysts' Forecasts (A-Model)

In the previous section, futures and expected spot curves for the FA-Model, calibrated using both futures and analysts' forecasts, were presented. In that setting each curve is affected by both sets of data. In this section we calibrate the model using only analysts' forecasts, modeling only the dynamics of the spot price. Thus, the expected spot curve represents an analysts' consensus curve that optimally considers all previous forecasts. The A-Model parameter values are presented in the Appendix. Given that futures data is not used, no futures curve or risk premium parameters are obtained.

Table 6 compares the mean absolute errors of the analysts' consensus curve in both models. As expected, the A-Model that only uses analysts' forecast data fits better this data than the FA-Model that includes also futures prices. This holds for every time window and maturity bucket.

Table 6: Price forecasts Mean Absolute Errors for the A- and FA-Models for each maturity bucket and time window, between 2010 and 2015. Errors are calculated as percentage of price forecasts. Parameter estimation from 2010 to 2014.

Buckets (years)	N° of Observations	In Sample (2010-2014)		Out of Sample (2015)		Total (2010-2015)	
		FA-Model	A-Model	FA-Model	A-Model	FA-Model	A-Model
0-1	149	3,8%	2,1%	4,3%	2,7%	3,9%	2,2%
1-2	284	4,8%	2,3%	5,0%	3,2%	4,9%	2,5%
2-3	236	7,1%	2,6%	6,7%	2,5%	7,1%	2,5%
3-4	190	9,4%	2,9%	7,5%	3,0%	8,9%	2,9%
4-5	141	10,8%	2,5%	7,9%	2,5%	10,0%	2,5%
5-10	122	7,7%	1,5%	4,5%	2,9%	6,9%	1,9%
10-28	110	5,4%	1,3%	4,4%	3,6%	5,2%	1,7%
Total	**1232**	**6,8%**	**2,3%**	**6,0%**	**2,9%**	**6,6%**	**2,4%**

Table 7: Expected Spot Mean Price and Annual Volatility of the FA- and A-Models, for each equal size maturity bucket between 2010 and 2015. Volatility of the curve at maturities in the middle of each bucket are presented. Parameter estimation from 2010 to 2014.

Maturity Buckets (years)	Mean Price ($/bbl.)		Annual Volatility (%)	
	FA-Model	A-Model	FA-Model	A-Model
0-5	95.1	95.4	18.3%	95.3%
5-10	103.2	104.8	22.1%	170.9%
10-15	118.9	115.1	26.8%	193.5%
15-20	144.9	133.0	29.2%	200.9%
20-25	182.2	158.6	30.3%	203.2%
Total	**129.0**	**121.5**	**25.4%**	**170.8%**

Table 7 reports the expected spot mean price and annual volatility for the FA- and A-Models, for each maturity bucket between 2010 and 2015. The first two columns show that the mean expected spot prices for the FA- and A- models are similar for short-term maturity buckets[14]. The last two columns report the volatility of expected prices obtained for the two models. Since the analysts´ forecasts are very noisy, the A-Model generates an analysts' consensus curve that is between 5 and 8 times more volatile than the one from the FA-Model.

In summary, the analysts' consensus curve can be obtained from the FA or the A-Models. The former has the advantage of generating a less volatile curve, while the latter generates a better fit. The difference between the means of both curves increases with maturity.

5.3. Long-Term Futures Price Estimation using also Analysts' Price Forecasts (FA-Model)

As has been argued earlier, estimation of long-term futures prices by extrapolation is subject to estimation errors. In addition, oil futures' longest maturity is around 9 years, while there are oil price forecasts for maturities of over 25 years. In this section, the impact on long-term futures prices of using analysts' price forecasts, in addition to futures, is explored.

To motivate this section **Figure 7** shows futures curves from for the FA- and F-Models on 04-14-2010 and compares them to the analysts' forecasts for the same date. It can be seen that both futures curves for long maturities are very different. On the other hand, both curves are very similar for short and medium term maturities, for which there is futures data. Given that there are no long-term futures data to validate any of the curves, we present the FA-Model futures curve as a valuable alternative that considers analysts' opinions.

Table 8 shows the mean price and annual volatility of the futures curves (FA- and F-Models) for every maturity. As can be seen, the inclusion of expectations data, when using the FA-Model, significantly affects the mean futures curve in the long-term, without considerably changing it in the short-term. Again, as was the case for the expected spot curves in the previous section, the longer the maturity the greater the difference between both curves[15]. Given the fact that analysts' forecasts are very volatile, the effect of using them almost doubles the volatility of the futures curves when using the FA-Model.

---

[14] In fact, differences in mean prices are significant at the 99% level for maturity buckets over 10 years.
[15] Differences in mean curves are significant at the 99% level for maturity buckets from 10 to 25 years.

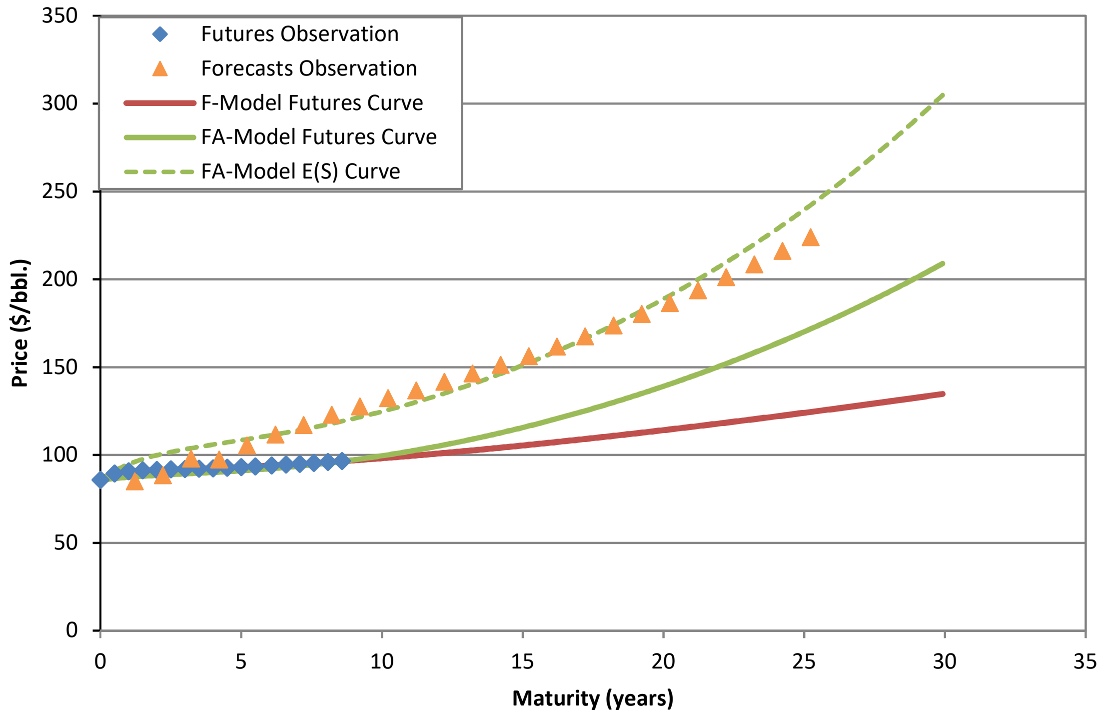

Fig. 7: Futures under the FA-, and F-Models, Expected spot curve under the FA-Model, forecasts and futures observations, for 04-14-2010. Parameter estimation from 2010 to 2014.

Table 8: Futures Mean Price and Annual Volatility of the FA- and F-Models, for each maturity bucket between 2010 and 2015. Volatility of the curve at maturities in the middle of each bucket are presented. Parameter estimation from 2010 to 2014.

Maturity Buckets (years)	Mean Price ($/bbl.)		Annual Volatility (%)	
	F-Model	FA-Model	F-Model	FA-Model
0-5	84.3	84.3	17.2%	18.3%
5-10	84.2	84.4	15.4%	22.1%
10-15	88.5	92.8	16.5%	26.8%
15-20	95.0	108.8	17.1%	29.2%
20-25	103.0	131.7	17.3%	30.3%
Total	**91.0**	**100.5**	**17.9%**	**25.4%**

5.4. Data Risk Premium Curves

Having reliable expected spot and the futures curves allows for the estimation of the term structure of risk premiums implied by their difference. As stated earlier the calibration of the F-Model provides most of the time statistically insignificant risk premium parameters, thus expected spot curves are unreliable. On the contrary, adding analysts´ forecast data addresses this issue.

Figure 8 shows the model term structure of risk premiums implicit in the difference of the expected spot and futures curves for the FA- and F- Models. In our model, changes in futures prices and analysts forecasts (i.e. expected spot prices) are driven by the three stochastic factors. More appropriately, the time series of the three factors and all the parameters of the model are jointly estimated using the Kalman Filter to fit the futures prices and analyst forecast data. Thus, futures prices (and also their slope) are a function of these three factors. One implication of this model is that the risk premium depends only on maturity and not on the state variables, so there is a constant risk premium curve[16] for each model over the whole sample period. The figure also shows the data risk premiums, obtained directly from the difference between price forecasts and their closest future price observation, averaged for each maturity over the whole sample period 2010 and 2015.

Several insights can be gained from **Figure 8**. First, the FA-model risk premiums are very close to the mean data risk premiums. Second, the term structure seems to be downward sloping, with annual risk premiums in the range of 2% to 10%. Finally, as expected, the F-Model is not able to obtain a credible estimation of risk premiums.

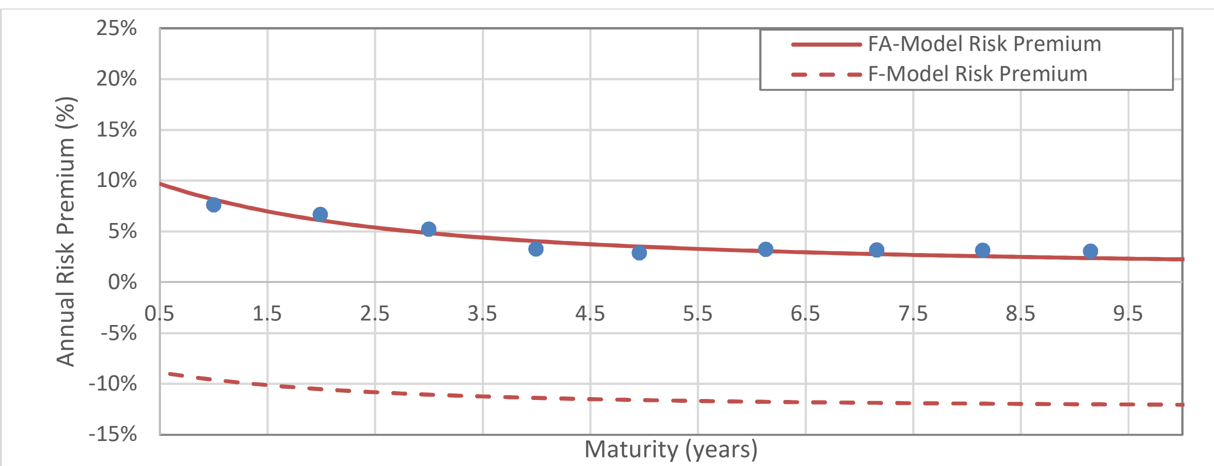

Fig. 8: Annual model risk premium term-structure for the FA- and F-Models, and annual mean data risk premiums. The data risk premiums are implicit from the difference between price forecasts and their closest future price observation, for every date between 2010 and 2015. Parameter estimation from 2010 to 2014.

[16] In a separate work some of the authors of this paper are considering a model with stochastic risk premium to better explain volatility

6. Conclusion

Even though commodity-pricing models have been successful in fitting futures prices, they do not generate accurate true distributions of spot prices. This paper proposes to calibrate these models using not only observations of futures prices, but also analysts´ forecasts of spot prices.

The Cortazar and Naranjo (2006) N-factor model is implemented for three factors, and estimated using the Kalman Filter. The model is calibrated using the traditional only-futures data (F-Model), an alternative only-analysts' data (A-Model), and a joint calibration using both sets of data (FA-Model). Futures data is from NYMEX contracts, and analysts´ forecasts from Bloomberg, IMF, World Bank, and EIA. Weekly oil data from 2010 to 2015 is used.

There are several interesting conclusions that can be derived from the results presented. The first is that in order to obtain reasonable expected spot curves, analysts´ forecasts should be used, either alone (A-Model), or jointly with futures data (FA-Model). Second, using both futures and forecasts (FA-Model), instead of using only forecasts (A-Model), generates expected spot curves that do not differ considerably in the short/medium term, but long term estimations are significantly different and the volatility of the curve is substantially reduced. Third, the inclusion of analysts´ forecasts, in addition to futures, in the FA-Model, instead of only futures prices (F-Model) does not alter significantly the short/medium part of the futures curve, but does have a significant effect on long-term futures estimations, and increases the volatility of the curve. Finally, that in order to obtain a statistically significant risk premium term structure, both data sets must be used jointly.

The information provided by experts in commodity markets, reflected in analysts' and institutional forecasts, is a valuable source that should be taken into account in the estimation of commodity pricing models. This paper is a first attempt in this direction.

7. References

Alquist, R., Kilian, L. & Vigfusson, R. (2013). Forecasting the Price of Oil, *Handbook of Economic Forecasting* 2, 1-46.

Altavilla, C., R. Costantini, & Giacomini R. (2014).Bond Returns and Market Expectations. *Journal of Financial Econometrics*, 12, 708-729.

Altavilla, C., Giacomini, R. & Ragusa, G. (2016). Anchoring the Yield Curve Using Survey Expectations, *Working Paper*, UCL.

Auffhammer. (2007). The rationality of EIA forecasts under symmetric and asymmetric loss. *Resource and Energy Economics, 29*(1), 102-121.

Baker, S. D., & Routledge, B.R. (2011). The Price of Oil Risk. *Working Paper*

Baumeister, C., & Kilian, L. (2015). Forecasting the real price of oil in a changing world: a forecast combination approach. *Forthcoming: Journal of Business and Economic Statistics, 33*(3), 338-351.

Baumeister, C., & Kilian, L. (2016). A General Approach to Re- covering Market Expectations from Futures Prices With an Application to Crude Oil, *Working Paper* .

Baumeister, C., & Kilian, L., & Lee, T. (2014). Are there Gains from Pooling Real-Time Oil Price Forecasts?, *Energy Economics* 46, S33-S43.

Berber, A., & Piana, J. (2016). Expectations, fundamentals and asset returns: evidence from commodity markets. *Working Paper*.

Bernard, J.T., Khalaf, L., Kichian, M. & Yelou, C. (2015). Oil Price Forecasts for the Long-Term: Expert Outlooks, Models or Both, *Macroeconomic Dynamics*, forthcoming.

Bianchi, D., & Piana, J. (2016). Expected spot prices and the dynamics of commodity risk premia. *Working paper*.

Bolinger, M., Wiser, R., & Golvone, W. (2006). Accounting for fuel price risk when comparing renewable to gas-fired generation: the role of forward natural gas prices. *Energy Policy, 34*(1), 706-720.

Brennan, M., & Schwartz, E. (1985). Evaluating natural resources investments. *Journal of Business, 58*(2), 135-157.

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

Chernov, M., & Mueller, P. (2012). The term structure of inflation expectations. *Journal of financial economics*

Chiang, I., Ethan, H., Hughen, W. K., & Sagi, J. S. (2015). Estimating oil risk factors using information from equity and derivatives markets. The Journal of Finance, 70(2), 769-804

Chun, A.L., (2011). Expectations, Bond Yields, and Monetary Policy. *Review of Financial Studies* 24, 208-247.

Cortazar, G., & Eterovic, F. (2010). Can Oil Prices Help Estimate Commodity Futures Prices? The Cases of Copper and Silver. *Resources Policy*, 35(4) 283–291

Cortazar, G., & Naranjo, L. (2006). An N-factor gaussian model of oil futures. *Journal of Futures Markets*, 26(3), 243-268.

Cortazar, G., & Schwartz, E. (2003). Implementing a stochastic model for oil futures prices. *Energy Economics, 25*(3), 215-238.

Cortazar, G., Kovacevic, I., & Schwartz, E. (2015). Expected commodity returns and pricing models. *Energy Economics*, 60-71.

Cortazar, G., Milla, C., & Severino, F. (2008). A multicommodity model of futures prices: using futures prices of one commodity to estimate the stochastic process of another. *Journal of Futures Markets, 28*(6), 537-560.

Cortazar, G., Schwartz, E., & Naranjo, L. (2007). Term-structure estimation in markets with infrequent trading. *International Journal of Financial Economics, 12*(4), 353-369.

Cutler, D.M., Poterba, J.M & Summers, L.H. (1990). Speculative Dynamics and the Role of Feedback Traders, *American Economic Review* 80, 63-68.

Dijk, D.V., Koopman, S. J. & Van Der Wel, M. (2012). Forecasting Interest Rates with Shifting Endpoints, *Tinbergen Institute Discussion Paper* 12-076/4.

Duffie, D., J. Pan & K. Singleton (2000). Transform analysis and asset pricing for affine jump-diffusions. *Econometrica* 68 (6), 1343–1376.

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

Greenwood, R., & Shleifer, A. (2014). Expectations of Returns and Expected Returns, *Review of Financial Studies* 27, 714-746.

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

Haugom, E., Mydland, O., & Pichler, A. (2016). Long term oil prices. *Energy Economics, 58*(1), 84-94.

Heston, S. L. (1993). A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options. *Review of Financial Studies* 6 (2), 327–343.

Kim, D. H., & Orphanides, A. (2012). Term structure estimation with survey data on interest rate forecasts. *Journal of Financial and Quantitative Analysis* 47, 241-272.

Koijen, R. S.J., Schmeling, M., & Vrugt, E.V.(2015). Survey Expectations of Returns and Asset Pricing Puzzles, *Working Paper.*

Lee, C. Y., & Huh, S. Y. (2017). Forecasting Long-Term Crude Oil Prices Using a Bayesian Model with Informative Priors. *Sustainability*, 9(2), 190.

Ready, R. C. (2016). Oil Consumption, Economic Growth, and Oil Futures: The Impact of Long-Run Oil Supply Uncertainty on Asset Prices. *Simon Business School Working Paper*, No. FR 14-14.

Sanders, D.R., Manfredo, M.R. & Boris, K. (2009). Evaluating Information in Multiple-Horizon forecasts: The DOE's Energy Price Forecasts, *Energy Economics* 31, 189-196.

Schwartz, E. (1997). The stochastic behavior of commodity prices: implications for valuation and hedging. *Journal of Finance, 52*(3), 927-973.

Schwartz, E., & Smith, J. (2000). Short-term variation and long-term dynamics in commodity prices. *Management Science, 46*(7), 893-911.

Stark, T. (2010). Realistic evaluation of real-time forecasts in the Survey of Professional Forecasters, *Tinbergen Institute Discussion Paper* 12-076/4.

Trolle, A. B. & Schwartz, E. S. (2009). Unspanned Stochastic Volatility and the Pricing of Commodity Derivatives. *Review of Financial Studies*, 22(11): 4423–4461.

Appendix

Three-factor F-Model, FA- and A-Model parameters, standard deviation (S.D.) and t-Test estimated from oil futures prices and price forecasts. Parameter estimation from 2010 to 2014.

Parameter	F-Model			FA-Model			A-Model		
	Estimate	S.D.	t-Test	Estimate	S.D.	t-Test	Estimate	S.D.	t-Test
κ2	1.015	0.011	92.490	0.940	0.023	40.877	0.316	0.030	10.494
κ3	0.200	0.003	74.208	0.170	0.004	47.314	0.259	0.021	12.508
σ1	0.175	0.003	52.173	0.311	0.003	102.803	2.044	0.126	16.278
σ2	0.531	0.006	91.077	0.241	0.004	56.060	9.566	4.759	2.010
σ3	0.251	0.004	58.302	0.455	0.008	58.918	9.989	4.649	2.149
ρ12	-0.162	0.003	-59.458	0.492	0.010	48.032	0.128	0.030	4.245
ρ13	-0.497	0.007	-66.317	-0.809	0.015	-52.635	-0.355	0.097	-3.661
ρ23	0.254	0.004	58.151	-0.693	0.012	-55.800	-0.972	0.029	-33.917
μ	-0.123	0.068	-1.818	0.002	0.000	44.564	-2.052	0.257	-7.990
λ1	-0.125	0.068	-1.844	0.007	0.003	2.605			
λ2	0.046	0.189	0.246	0.101	0.009	11.151			
λ3	0.000	0.001	0.029	0.010	0.007	1.429			
ξ	0.005	0.000	102.346	0.044	0.000	108.762	0.042	0.001	29.975