Oil Price Predictability and Risk Premia Based on Market Fundamentals

Daniel Cuzzi[†]

João Victor Issler[†]

Abstract

Understanding how oil prices relate to global fundamentals is crucial for both economic policy and investment strategy. Our findings reveal that oil prices exhibit a long-term relationship with global industrial production and oil output, distinguishing between OPEC and non-OPEC production. Our equilibrium measure provides unbiased forecasts, with cyclical deviations diminishing within two years. Forecasting exercises demonstrate a reduction in mean squared error by over 30% compared to a no-change model and by over 15% compared to forward prices with a two-year maturity. We also estimate a time-varying risk premium, which is negatively correlated with global stock indices, in contrast to the risk premiums from the no-change model, which estimate only half of the risk premium for holding oil.

JEL classifications: *C5,G12,Q41*

Keywords: *Oil fundamental value, oil price forecasting, oil price risk premia*

[†]FGV EPGE Brazilian School of Economics and Finance

1 Introduction

As oil became a primary input in the global economy, a substantial body of literature emerged exploring oil price economic drivers and the consequential impacts on macro variables. Predicting oil prices is relevant for policymakers, businesses, and asset managers primarily due to its impact on inflation and on economic activity through energy costs. We track oil prices through long-term equilibrium with market fundamentals, interpreted as the oil fundamental value. Our findings suggest that the error correction mechanism plays a crucial role in explaining cumulative oil returns over time. We show that the forecast generated by our model outperforms future prices and the no change model in terms of MSE reduction and our model provides an unbiased forecast for oil prices over a two-year horizon. Additionally, we recover the implied risk premia and we bring evidences of a sharp discrepancy of the risk measure comparing with the term premia, or the risk premium considering a no change model.

To model the long-term trend, we focus on the information in commodity economic fundamental levels, with the equilibrium value being the shared trend within the system. Industrial Production serves well as an economic activity tracker and a supply-side measure reflecting global production levels, acting as a demand source of information for the oil market. Along with the literature in structural analysis and forecasting, these models assume no equilibrium in levels. If this equilibrium exists, there is a price level such that deviations are transitory. We find evidence of this mechanism following He et al. (2010) and Lardic and Mignon (2006), and more importantly, we find that these deviations explain future oil price movements.

We contribute to the discussion on cointegrating relationships in the oil market by finding statistical equilibrium for oil prices with information on economic activity level, oil production, and oil prices. The most common approach is to use GDP; some studies use data for the US or extend to G7 countries such as Lardic and Mignon (2006) or the fright index following Kilian (2009) applied for cointegration in He et al. (2010). We utilize the Global Industrial Production data discussed in Hamilton (2021), which captures not only OECD data on industrial production but also data for Brazil, China, India, Indonesia, Russia, and South Africa. What also improves our estimates is that as GDP is measured quarterly while the industrial production index is monthly, we are allowed for higher frequency data in our analyses.

A positive shock to industrial production comes together with the derived demand for input as studied in Issler et al. (2014). Because supply is relatively restricted in the short-term, most of the effect goes to prices. Indeed, most short-term oil price variations come from demand shocks Kilian (2009); Alquist et al. (2013); Lippi and Nobili (2012); Duarte et al. (2021). Because we want to track persistent demand we focus on level relationships rather then modeling oil short-term fluctuations. In this sense cointergation is essential in our context. Industrial production is a slow moving demand information as it is actually a supply side fundamental of the global economy. In our dataset, cointegration is only robust when desegregating the supply side between non-OPEC and OPEC oil production, motivated by evidence of heterogeneous data-generating processes and also evidenced in our dataset. The long-term equilibrium implies that an error correction mechanism exists, where some variables respond to the lagged disequilibrium, adjusting for the long-term. We show that oil prices respond to the disequilibrium, and we interpret the central tendency of the oil prices as its fundamental value.

We use this model to generate forecasts for oil prices, showing that it is unbiased over a two-year horizon and reduces MSE compared to random walk and forward prices. As in Pagano and Pisani (2009) we find significant forecast error in crude oil futures. We estimate a time-varying risk premium and find that it is negatively correlated with global stock indices, in contrast to the risk premiums predicted by a random walk model, which also estimate only half of the risk premium for holding oil. We estimate a time-varying risk premium that is negatively correlated with global stock indices, contrasting with the risk premiums predicted by a random walk model that do not correlate with global stocks return, and also estimate only half of the actual risk premium associated with holding oil. Our analysis reveals that long-term contracts, with an average maturity of around 2 years, are priced approximately 7% below the long-term value, compared to an estimated 3.5% under a random walk process.

2 Data

We utilize monthly data spanning from January 1993 to May 2023. The oil-price data is sourced from the FRED database of the St. Louis Federal Reserve and specifically comprises the global price of West Texas Intermediate (WTI) crude oil. We obtain the real oil price after deflating the nominal price using the US Consumer Price Index (CPI), also retrieved from the FRED database. Additionally, we incorporate the Global Industrial Production index, following the approach by Hamilton Baumeister and Hamilton (2019). These series are seasonally adjusted and serve as a robust proxy for global economic activity and combine OECD Industrial production with Brazil, China, India, Indonesia, Russia, and South Africa. The level of industrial production reflects a substantial demand for oil, making it a reliable indicator of aggregate demand. For data on oil production, we accessed information from the US Energy Information Administration open data.

Figure 1: The Industrial Production Index by Baumeister and Hamilton (2019) and the real oil price log levels in the top and 9month moving average of their return in the bottom.

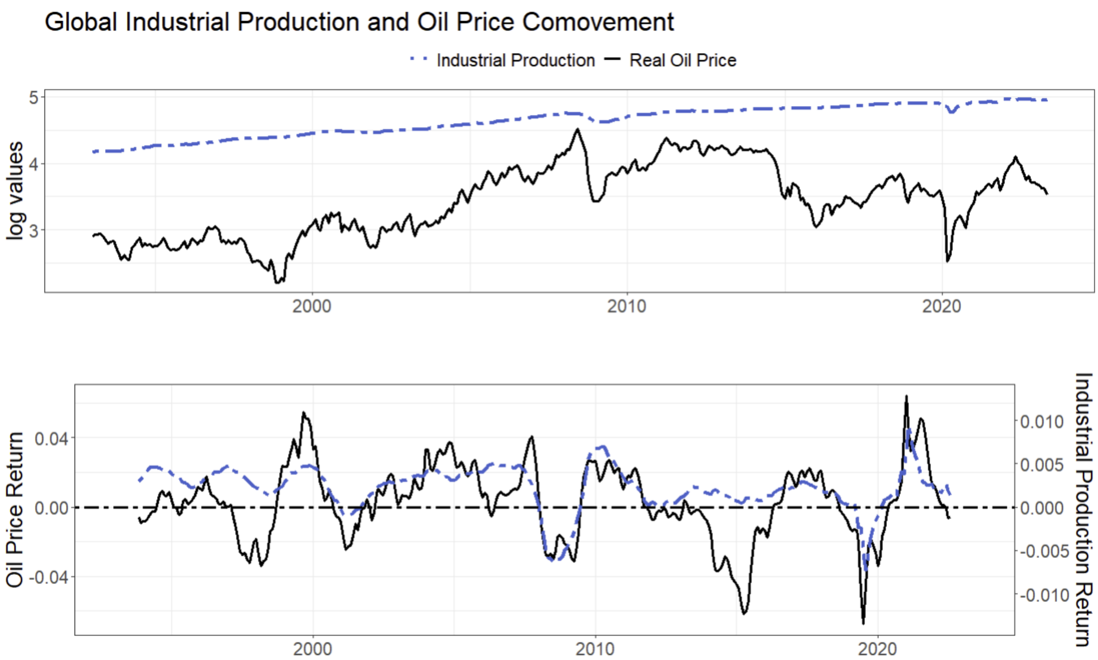
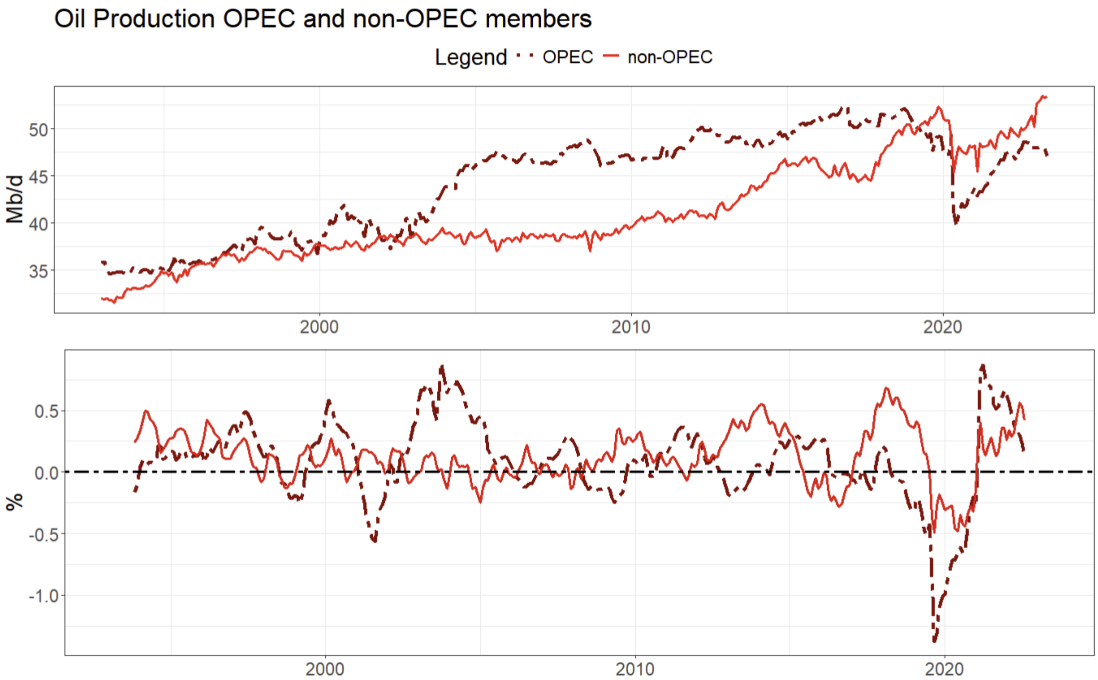


Figure 2: Top: Level of OPEC and non-OPEC production in millions of barrels per day. Bottom: the 9-month moving average of the return of the series.



3 The Long Term Equilibrium

3.1 Cointegration Estimates

Let $y_t = (p_t, ip_t, q_t^o, q_t^r)$ represent the natural logarithms of the real oil price, OPEC and non-OPEC oil production, and global industrial production. All variables exhibit a unit root, as shown by Phillips-Perron tests in the Appendix. However, this is not true for the cointegrating error, as presented in Table 1. For different specifications, the Johansen procedure confirms with 99% confidence that there is a single cointegrating vector, which allows us to estimate the equilibrium using the following OLS regression:

$$p_t = \hat{\beta}_1 ip_t + \hat{\beta}_2 q_t^o + \hat{\beta}_3 q_t^r + e_t^c.$$

Cointegration tests, assuming that oil production can be treated as a single input in the data generating process where $q_t = q_t^o + q_t^n$, fail to show cointegration, as indicated in Table 2. There is heterogeneity in the supply side and the relationship between oil production for each group within the data generating process. For any specification, the long-term relationship between OPEC production and prices is closer to zero and sometimes slightly positive, depending on the specification. This occurs because the relationship between OPEC production and prices is relatively more influenced by demand rather than the competitive side of the market, where production growth is associated with productivity gains and discoveries, characterizing supply shocks and a stronger negative relationship with prices. As these are reduced-form estimates, they can only be interpreted statistically within the equilibrium condition and for dynamic estimates, not economically.

The fundamental value corresponds to the expected real oil price given market conditions:

$$p_t^* = \hat{\beta}_1 ip_t + \hat{\beta}_2 q_t^o + \hat{\beta}_3 q_t^r,$$

and the equilibrium condition is related to $p_t - p_t^* = e_t^c$, where p_t and p_t^* are integrated of order 1 (I(1)) and e_t^c is integrated of order 0 (I(0)). Results, presented in Table 2, indicate that a one percent increase in non-OPEC oil production is associated with a 3.6 percent decrease in real oil prices. Conversely, an increase in OPEC oil production is associated with a 2.7% increase in real oil prices in equilibrium. We show in the structural analysis that this is due to the primary source of comovement between the variables: non-OPEC production varies more with supply shocks, while OPEC production is more affected by demand shocks in the medium to long term. The relationship with global industrial production suggests an approximate 3% increase in the commodity price for a 1% increase in global production.

3.2 Return to Fundamental Levels

Central banks and market players recognize the importance of the oil price as a crucial variable, relevant for evaluating the risks associated with macroeconomic developments. We assess the explanatory capability of the fundamental value and compare results with benchmarks. Consistent with the arguments by Cogley (2002) in studying inflation convergence, later applied by Burger et al. (2022) in capital flows, if $\mathbb{E}_t[p_{t+h^*}] = p_t^*$, where h is a long-term horizon over which we anticipate the real oil price will converge to its

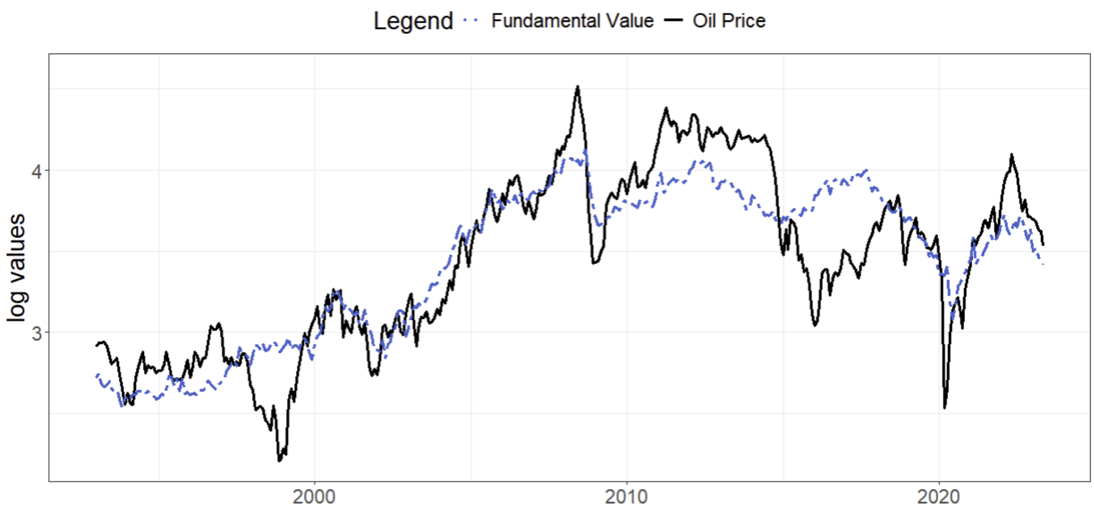


Figure 3: In dashed blue the estimated oil fundamental value and in the real oil price in black. Sample from 1993.1 to 2023.9

fundamental level, subtracting both sides by p_t yields $\mathbb{E}_t[p_{t+h}] - p_t = -p_t + p^*_t$. Rewriting and using the definition $\mathbb{E}_t[p_{t+h}] = p_{t+h} + u_{t+h}$ we have that $p_{t+h} - p_t = e^c_t + u_{t+h}$. Considering $\alpha_{1,h^*} = -1$, we obtain the following specification:

$$p_{t+h} - p_t = \alpha_{1,h} e^c_t + u_{t+h} \tag{1}$$

The cointegration error appears on the right-hand side, explaining the cumulative returns. We can estimate this in an OLS estimation by replicating the local projection h periods ahead. Equation (2) parallels the analysis conducted by Cogley (2002) on inflation and Burger et al. (2022) on capital flows. It suggests that if the relationship holds, the gap between expected real oil prices h periods ahead and current real oil prices is the negative of today's difference between p_t and p^*_t.

$$\mathbb{E}[p_{t+h^*} - p_t] = -e^c_t \tag{2}$$

We test whether deviations of current real oil prices from the natural level are inversely related to subsequent changes in real oil prices. Cogley (2002) emphasized that $\alpha_{0,h}$ should equal zero; otherwise, p^*_t would be biased. However, the focus is primarily on $\alpha_{1,h}$ following Burger et al. (2022). If p^* reflects real oil prices' long-term trend, we obtain

$\alpha_{1,h} = -1$ for medium-run horizons. A $\alpha_{1,h} = -1$ estimate implies that the gap between real oil prices and p^* represents its transitory component, and real oil prices are expected to converge to p^* in h periods.

The fundamental level represents the expected oil price or the value in the absence of volatility. Utilizing an equilibrium measure, our objective is to enhance predictions as stationary shocks fade. Conducting the Cogley (2002) test in line with Burger et al. (2022) for horizons ranging from 1 to 30 months ($h = 1, \ldots, 30$), We estimate Equation (2)[1]. Note that this analysis is out of sample, utilizing the period t gap between the actual real oil price and the predetermined p^* to predict the h period-ahead change in oil returns. The proximity of $\alpha_{1,h}$ to -1 serves as a summary measure of the model's performance.

Our findings indicate that the mean return occurs over an average period of two years, which intriguingly aligns with results obtained for other economic variables, such as real exchange rates (Rossi (2013)) and capital flows (Burger et al. (2022)). Results suggest that convergence is fully archived, and the disequilibrium of the oil price to our fundamental measure vanishes in around 2 years.

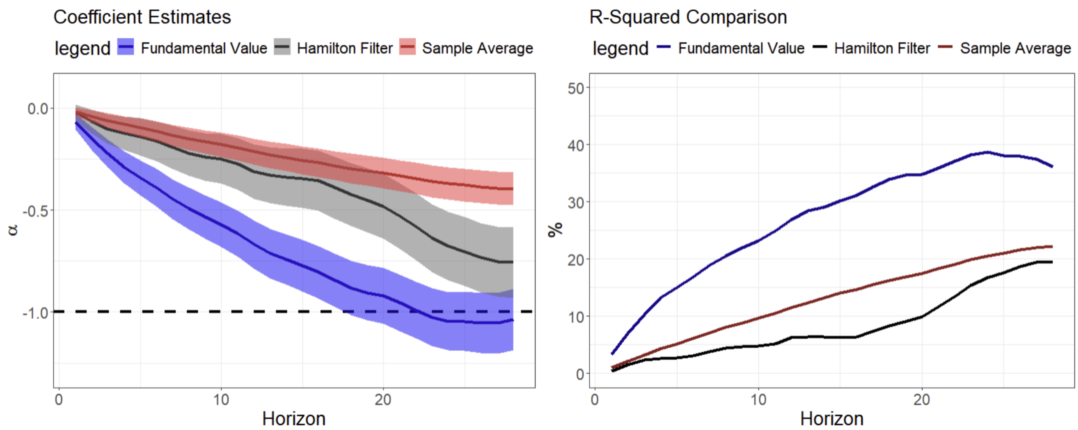


Figure 4: Estimates of equation 2. Forecast horizon going from 1 to 30. In the left-hand side, we have the slope of the model with 95% bounds. We can interpret it as a share of mean return which is complete in the case of cointegration in which We cannot reject -1 for α_h for the twentieth month on. R-squared of the models on the right-hand side

We perform medium-term forecasting, the same exercise for the fundamental value

[1]
$$\sum_{1 \leq i \leq h} \Delta p_{t+i} = \alpha_{0,h} + \alpha_{1,h} e_t^c + e_{t+h}$$

but using the Hamilton Filter, following the author's suggestion by filtering two years of cycles, a practice consistent with our findings. This filter is known to perform well as a trend tracker[2] and this historical average. The Hamilton Filter proposed in Hamilton (2018) is a long-term tracker of the variable of interest; it eliminates the cycle which is based on the trend and cycle decomposition idea put forth by Beveridge and Nelson (1981). Burger et al. (2022) show that the Hamilton Filter has the same performance for capital flow forecast compared to their equilibrium measure. Here we had a similar result where Hamilton filter performed well with coefficients around 0.9 in two years. The explanation power is a bit higher for the Hamilton Filter and fluctuations of the long-term measure produce similar variance compared to the fundamental value which is 50% of the 1% monthly standard deviation of oil return.

4 Forecasting the Oil Price

Short-term predictability and outperformance against random walk can be found in Alquist et al. (2013) for the 3 month horizon using oil futures Baumeister and Kilian (2012) presents nowcasting models using macroeconomic aggregates that performs well in the short-run yet the enhancement of predictability against the random walk diminishes as the horizon extends beyond 1 year. We present results focusing on long-term predictability aligned with the scope of our model. We present the RMSE ratio comparing model and forward prices with the random walk. Is interesting to note that the forward curve outperforms random walk to forecast oil prices reducing forecast error in more than 10%. But our model performs even better reducing RMSE in more than 30% for the two years horizon. In the next figure we can visualize the spot price together with the future prices in grey and model projection in red. For each period we observe 30 months ahead of the forecasts given by the forward curve and model forecast. The commodities super cycle is an important example to understand how the model works. In this period oil went from around 60 and more than doubled in a year with no significant changes in the long-term fundamentals. We can visualize the parallel forward curves in

[2]The specification is given by: Hamilton (2018)

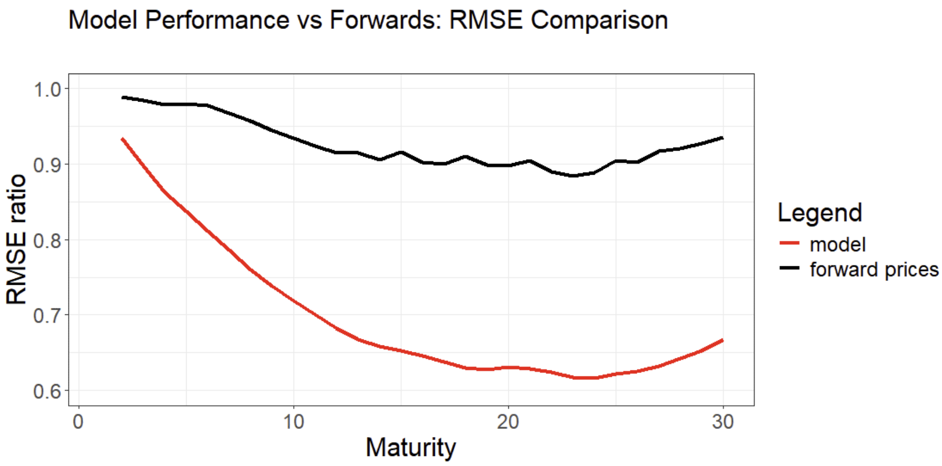


Figure 5: Comparison of Root Mean Square Errors of each forecaster versus random walk. In red, results generated by our model with a sharp reduction of the forecasting error and in grey forward prices also generating predictability but with a weaker performance.

this period where market priced the shock to whole term structure as if it was a permanent shock but based in our results not much change in the equilibrium value despite the surprisingly high oil price increase in the period.

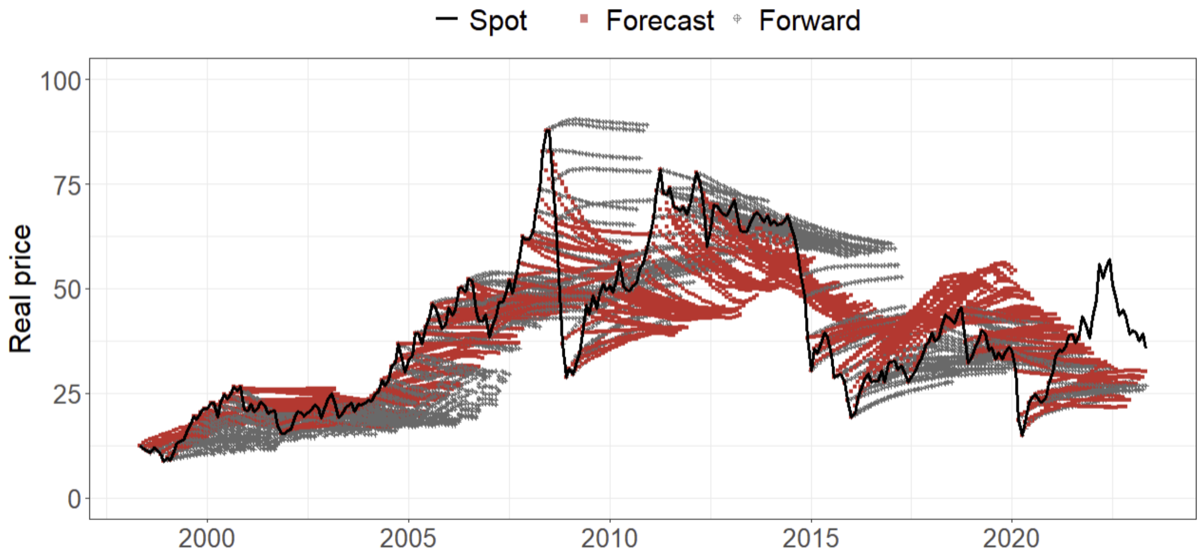


Figure 6: In solid black line is the oil spot price, in red forecasts generated by our model and in grey future prices from 1998 to 2923.5

5 Forward Prices and the Fundamental Value

One approach to price futures involves incorporating a *risk premium* into the expected value of the spot price. In this section, we study the risk premia generated by utilizing our model as forecaster and compare it to the term premia. We define *term premia* as difference between the forward and the spot price. We model future contracts based on the assumption that p_t tends towards p_t^* as previously evidenced. Let $f_{t,h}$ represent the future contract for the underlying oil price p_t. Pagano and Pisani (2009) documents the significant forecast error on crude oil futures. A long position in oil futures has a random payoff of: $f_t^h - p_{t+h}$. No-arbitrage conditions require:

$$f_t^{(h)} = E_t[p_{t+h}] - rp_t^{(h)}$$

The expected value of the h-step ahead oil price is the price today adjusted by the share of the disequilibrium we expect to vanish.

$$\mathbb{E}_t[p_{t+h} - p_t] = \alpha^h e_t^c$$

Pricing the h-month forward contract yields:

$$f_t^{(h)} = (1 + \alpha^h)p_t - \alpha^h p_t^* - rp_t^h$$

where rp_t denotes the risk premium. When the expected value is equal to price today as it is in a random walk model risk premium will match the inclination of the forward structure:

$$rp_{t,h} = f_{t,h} - E_t[p_{t+h}]$$
$$tp_{t,h} = f_{t,h} - p_t$$

Recall that p_t^* is a value based on fundamentals, deviations from which are transitory. The disequilibrium $e_t^c = p_t - p^*$ vanishes over time and we estimate that the gap closes within 18 to 24 months. So α_h tends towards -1 in two years as the real oil price converges to the fundamental value. If the the price of the 24th maturity in the oil price term structure matches our model forecasts the risk premium will be zero but the term premia in general different from zero. We can discuss non-arbitrage based in our results for the horizons for which our model fits well as a expected value as happens when cycle vanishes.

5.1 Long-Term Risk Premium

We estimate the time varying risk premium that can be visualized in Figure 8 together with the term premia. Is interesting to note the divergence we find contrasting with our estimates and on the right hand side we present the average premium per maturity where we can note that the according to our model that reduces significantly RMSE against the no change model the premium faced by agents that are hedging producers on the real side of the economy is doubled when compared to the benchmark.

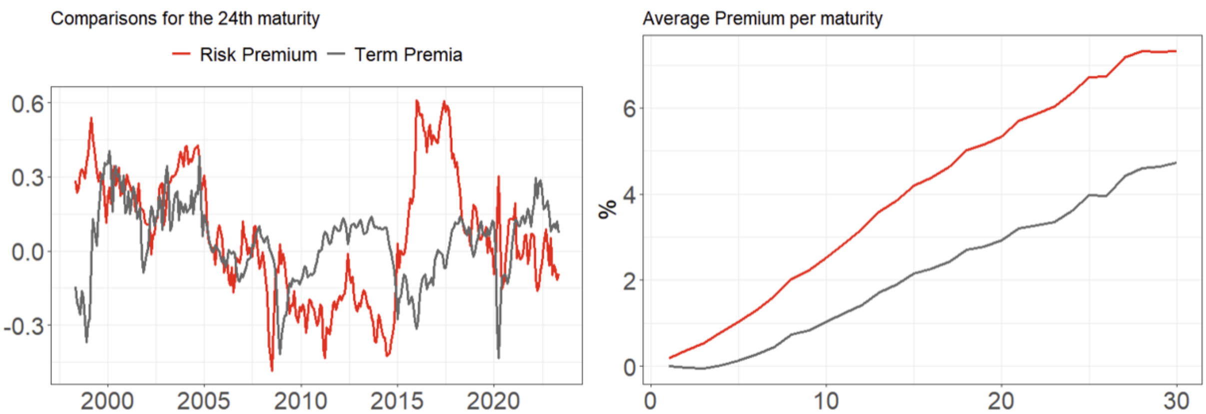


Figure 7: Estimates of the risk premia given by the model and the term premia recovered from the forward prices from 1997 to 2023.5 left side and on the right hand side their average estimates for per maturity from 1 to 30 months

In Figure 9 we present the decomposition of the risk premium in the two terms. In red the disequilibrium that is an estimate of the disequilibrium according to our measure of fundamental value which we estimate to vanish in two years and in grey we have what market is pricing to occur in two years. First thing to note is that the volatility implied from the term premia is 15% what is 60% of the 25% generate by our model.

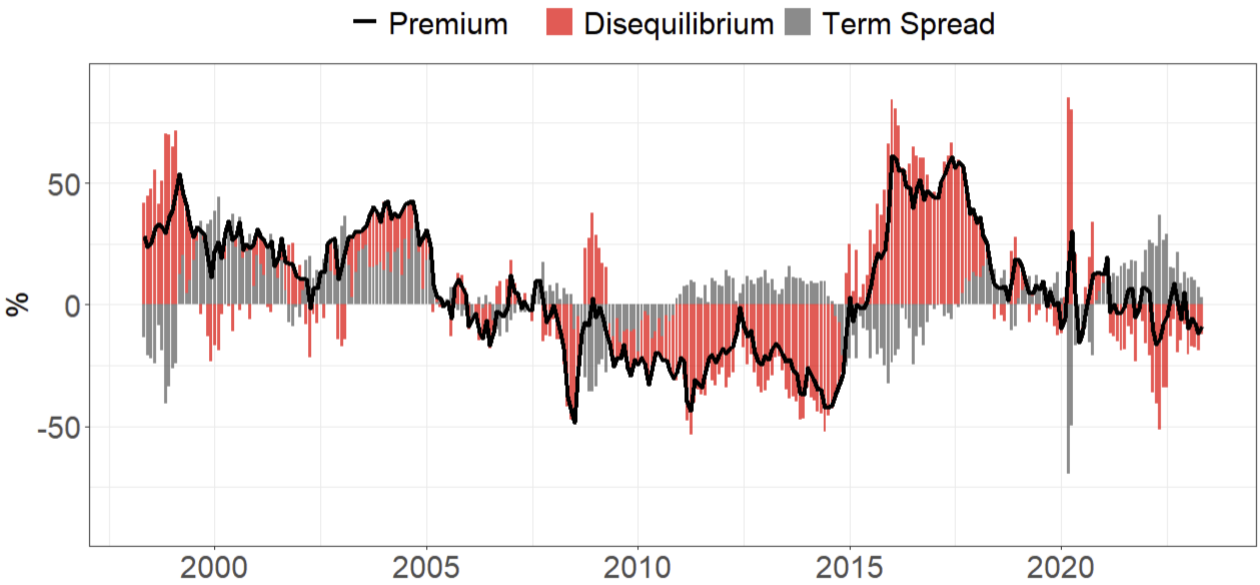


Figure 8: The solid black line is our estimates of the risk premium. In red is the short run deviation predicted by our model and in grey we have the term premia for the 24th maturity.

$$rp_t^{(24)} = -(f_t^{24} - p_t^*) = -spread_t^{24} - e_t^c$$

5.2 Global Financial Cycles

The peak of the forecasting performance occurs in 24 months as volatility is not being modeled because we are track oil price trend. We compute simple OLS to study the correlation between the risk premia, the disequilibrium and the spread for the 2-year contract with the SP-500 and the Global MSCI return. The relationship of with the cointegration error is positive reflecting that a positive return on SP-500 is related to increase in oil prices vs the long-term disequilibrium. The risk premium is negatively correlated with global stocks return. If we want to assume a no-change model and assume that p_t is a forecaster for the future price we have that the risk premium is the negative of the spread. For this model the risk premium is not correlated with global stocks return.

	$rp_{t,24}$	$spread_{t,24}$	e_t^c	$rp_{t,24}$	$spread_{t,24}$	e_t^c
SP-500	−0.1713**	0.0647	0.3007***			
	(0.0571)	(0.0578)	(0.0553)			
Global MSCI				−0.2200***	0.0954	0.3578***
				(0.0565)	(0.0577)	(0.0541)
Observations	300	300	300	300	300	300
F Statistic	9.010**	1.252	29.620***	15.150***	2.735	43.740***
R-squared	0.0294	0.0042	0.0904	0.0484	0.0091	0.1280

Table 1: SImple OLS between global stocks and SP-500 with the risk premium, spread and the cointegratio error from 1998.1-2023.5

6 Conclusion

The interplay between oil prices and macroeconomic variables has long been a central focus of economic research, reflecting oil's role as a key input in the global economy. Understanding the estimated value of oil in the absence of transitory shocks can help policymakers and market participants better comprehend how fundamentals influence price dynamics.

To model the long-term trend, we focus on commodity economic fundamentals, identifying a long-term equilibrium relationship between oil prices, global industrial production, and oil supply, with distinctions between OPEC and non-OPEC production. Our measure provides unbiased forecasts over a 24-month horizon and explains approximately 40% of cumulative returns in oil prices over the same period. This equilibrium-based forecasting model outperforms both futures and random walk models, reducing forecasting errors by about 35% compared to a random walk and 15% compared to forward prices for long-term maturity contracts.

We estimate a time-varying risk premium that is negatively correlated with global stock indices, contrasting with the risk premiums predicted by a random walk model that do not correlate with global stocks return, and also estimate only half of the actual risk premium associated with holding oil. Our analysis reveals that long-term contracts, with an average maturity of around 2 years, are priced approximately 7% below the long-term value, compared to an estimated 3.5% under a random walk process.

References

Alquist, R., L. Kilian, and R. J. Vigfusson (2013): "Forecasting the price of oil," in *Handbook of economic forecasting*, Elsevier, vol. 2, 427–507.

Baumeister, C. and J. D. Hamilton (2019): "Structural interpretation of vector autoregressions with incomplete identification: Revisiting the role of oil supply and demand shocks," *American Economic Review*, 109, 1873–1910.

Baumeister, C. and L. Kilian (2012): "Real-time forecasts of the real price of oil," *Journal of business & economic statistics*, 30, 326–336.

Beveridge, S. and C. R. Nelson (1981): "A new approach to decomposition of economic time series into permanent and transitory components with particular attention to measurement of the 'business cycle'," *Journal of Monetary economics*, 7, 151–174.

Burger, J. D., F. E. Warnock, and V. C. Warnock (2022): "A natural level of capital flows," *Journal of Monetary Economics*, 130, 1–16.

Cogley, T. (2002): "A simple adaptive measure of core inflation," *Journal of money, credit and banking*, 94–113.

Duarte, A. M., W. P. Gaglianone, O. T. de Carvalho Guillén, and J. V. Issler (2021): "Commodity prices and global economic activity: a derived-demand approach," *Energy Economics*, 96, 105120.

Hamilton, J. D. (2018): "Why you should never use the Hodrick-Prescott filter," *Review of Economics and Statistics*, 100, 831–843.

——— (2021): "Measuring global economic activity," *Journal of Applied Econometrics*, 36, 293–303.

He, Y., S. Wang, and K. K. Lai (2010): "Global economic activity and crude oil prices: A cointegration analysis," *Energy Economics*, 32, 868–876.

Issler, J. V., C. Rodrigues, and R. Burjack (2014): "Using common features to understand the behavior of metal-commodity prices and forecast them at different horizons," *Journal of International Money and Finance*, 42, 310–335.

Kilian, L. (2009): "Not all oil price shocks are alike: Disentangling demand and supply shocks in the crude oil market," *American Economic Review*, 99, 1053–1069.

Lardic, S. and V. Mignon (2006): "The impact of oil prices on GDP in European countries: An empirical investigation based on asymmetric cointegration," *Energy policy*, 34, 3910–3915.

Lippi, F. and A. Nobili (2012): "Oil and the macroeconomy: a quantitative structural analysis," *Journal of the European Economic Association*, 10, 1059–1083.

Pagano, P. and M. Pisani (2009): "Risk-adjusted forecasts of oil prices," *The BE Journal of Macroeconomics*, 9.

Rossi, B. (2013): "Exchange rate predictability," *Journal of economic literature*, 51, 1063–1119.

7 Appendix

Table 2: Estimates of equation 1. Cointegration Results with and without constant for the sample period. Includes estimation prior to COVID-19 Crisis.

	Real oil price			
	1993.1 - 2019.12		1993.1 - 2023.5	
	(1)	(2)	(3)	(4)
non-OPEC oil production	−3.147***	−4.904***	−3.624***	−5.397***
	(0.160)	(0.306)	(0.218)	(0.332)
OPEC oil production	2.137***	0.486*	2.702***	−0.452
	(0.158)	(0.291)	(0.233)	(0.513)
Global Industrial Production	3.097***	4.786***	2.880***	5.533***
	(0.079)	(0.266)	(0.096)	(0.401)
Constant		28.477***		40.280***
		(4.311)		(5.933)
Observations	365	365	324	324
R^2	0.995	0.865	0.995	0.876
Residual Std. Error	0.268 (df = 362)	0.253 (df = 361)	0.264 (df = 321)	0.247 (df = 320)
Note:				*p<0.1; **p<0.05; ***p<0.01

Table 3: Phillips Perron Unit Root Tests: Null Hypothesis $\sim \mathcal{I}(1)$. Unit root rejected only for the cointegrating error.

Variable	Test Type	Dickey-Fuller Statistic	Truncation Lag Parameter	p-value
Real oil price	trend + Constant	-2.2446	5	0.4737
	Constant	-10.511	5	0.5218
Global Industrial Production	trend + Constant	-2.3551	5	0.4271
	Constant	-10.353	5	0.5306
non-OPEC oil production	trend + Constant	-2.7987	5	0.2398
	Constant	-16.476	5	0.1879
OPEC oil production	trend + Constant	-1.6545	5	0.7228
	Constant	-6.6099	5	0.7402
cointegrating error	trend + Constant	-3.6019	5	0.03306
	Constant	-25.027	5	0.02331

Table 4: Phillips-Perron Unit Root Tests: Null Hypothesis $\sim \mathcal{I}(1)$. Unit root rejected only for the cointegrating error.

Variable	Test Type	Dickey-Fuller Statistic	Truncation Lag Parameter	p-value
Real oil price	trend + Constant	-2.2596	5	0.4674
	Constant	-10.799	5	0.5057
Global Industrial Production	trend + Constant	-2.3551	5	0.4271
	Constant	-10.353	5	0.5306
non-OPEC oil production	trend + Constant	-3.1671	5	0.09373
	Constant	-19.506	5	0.07818
OPEC oil production	trend + Constant	-1.6545	5	0.7228
	Constant	-6.6099	5	0.7402
cointegrating error	trend + Constant	-2.4595	5	0.383
	Constant	-12.367	5	0.4179

Table 5: Johansen Cointegration Test: Maximal Eigenvalue Statistic (lambda max) Without Linear Trend and Constant in Cointegration

Test Type	**Test Statistic**	**10% Critical Value**	**5% Critical Value**	**1% Critical Value**
$r \leq 3$	2.22	7.52	9.24	12.97
$r \leq 2$	10.27	13.75	15.67	20.20
$r \leq 1$	17.92	19.77	22.00	26.81
$r = 0$	33.67	25.56	28.14	33.24