Are Product Spreads Useful for Forecasting Oil Prices? An Empirical Evaluation of the Verleger Hypothesis

Christiane Baumeister — *Bank of Canada*

Lutz Kilian — *University of Michigan, CEPR*

Xiaoqing Zhou — *University of Michigan*

EIA 2014 Workshop on
Financial and Physical Oil Market Linkages
October 6, 2014

Background

- Oil price forecasts affect the economic outlook of oil-importing and oil-exporting countries.

- Users of real oil price forecasts include governments at the state and federal level, international organizations, central banks, and many industries (e.g., airline companies, car manufacturers, utilities)

- Old consensus in the literature:
    - ➢ The real price of oil is inherently unpredictable.
    - ➢ Nothing beats the no-change forecast (random walk) at short and long horizons.

- Several alternatives have been developed in recent years.

Oil Price Predictors Used in the Recent Literature

- Forecasts based on the oil futures prices

  e.g., Chernenko, Schwarz and Wright 2004; Knetsch 2007; Alquist and Kilian 2010; Reeve and Vigfusson 2011; Baumeister, Guérin and Kilian 2014

- Professional and survey forecasts

  e.g., Sanders, Manfredo and Boris 2008; Alquist, Kilian and Vigfusson 2013; Kilian and Hicks 2013; Baumeister and Kilian 2014; Baumeister, Kilian and Lee 2014; Bernard, Khalaf, Kichian and Yelou 2014

- Forecasts based on changes in non-oil industrial commodity prices, exchange rates and oil company stock prices

  e.g., Chen, Rogoff and Rossi 2010; Baumeister and Kilian 2012; Alquist, Kilian and Vigfusson 2013; Chen 2013; Baumeister, Guérin and Kilian 2014

- Forecasts based on changes in oil inventories, oil production, global real economic activity

  e.g., Alquist, Kilian and Vigfusson 2013; Baumeister and Kilian 2012, 2014; Baumeister, Guérin and Kilian 2014

New forecasting approach

- Demand for crude oil derives from the demand for refined products such as gasoline and heating oil.

- Spot prices for petroleum products are primary determinants of crude oil prices.

- Difference between refined product market prices and the purchase price of crude oil should have predictive power for the price of crude oil (*Verleger hypothesis*).

How to arrive at a forecasting model

- Basic idea:

$$P_t^{oil} = \sum_{i=1}^{n} w_i P_t^i$$

How to arrive at a forecasting model

- Basic idea:

$$P_{t+h}^{oil} = \sum_{i=1}^{n} w_i P_{t+h}^{i}$$

- Given date t information, the conditional expectation is:

$$E(P_{t+h|t}^{oil}) = \sum_{i=1}^{n} w_i E(P_{t+h|t}^{i})$$

- Expectation of the spot market price for product i can either be the futures price or the spot price of product i

  $\Rightarrow$ Let's focus on futures prices.

How to arrive at a forecasting model

- Transform into a regression model that we can estimate:

$$E(P_{t+h|t}^{oil}) = \sum_{i=1}^{n} w_i F_{t+h,t}^{i}$$

$$\log\left(E\left(\frac{P_{t+h|t}^{oil}}{P_t^{oil}}\right)\right) = \log\left(\sum_{i=1}^{n} w_i F_{t+h,t}^{i}\right) - \log P_t^{oil}$$

$$\Delta p_{t+h|t}^{oil} = \alpha + \beta\left[\log\left(\sum_{i=1}^{n} w_i F_{t+h,t}^{i}\right) - p_t^{oil}\right] + \varepsilon_{t+h}$$

- Construct the forecast of the real price of oil:

$$\widehat{P_{t+h|t}}^{oil} = P_t^{oil} exp\left\{\hat{\alpha} + \hat{\beta}\left[\log\left(\sum_{i=1}^{n} w_i F_{t+h,t}^{i}\right) - p_t^{oil}\right]\right\}$$

$$\widehat{R_{t+h|t}}^{oil} = R_t^{oil} exp\left\{\hat{\alpha} + \hat{\beta}\left[\log\left(\sum_{i=1}^{n} w_i F_{t+h,t}^{i}\right) - p_t^{oil}\right] - E\left(\pi_{t+h}^{h}\right)\right\}$$

$\Rightarrow$ Let's consider several special cases.

The Single Futures Spread Model

$$\widehat{R_{t+h|t}}^{oil} = R_t^{oil} exp\left\{\hat{\alpha} + \hat{\beta}\left[\boldsymbol{f^{i}_{t+h,t}} - p_t^{oil}\right] - E\left(\pi^h_{t+h}\right)\right\}$$

where $f^i_{t+h,t}$ is the log of the futures price of product i with $i \in \{gasoline, heating\ oil\}$ at time t with maturity h periods

The Crack Spread Futures Model

- Refined products are produced from a barrel of crude oil in approximately fixed proportions.
- Typically: 3 barrels of crude oil converted to

  2 barrels of gasoline and 1 barrel of heating oil

$\Rightarrow$ 3:2:1 crack spread can be expressed as follows:

$$F_{t+h,t}^{CS} \equiv \frac{2}{3} F_{t+h,t}^{gas} + \frac{1}{3} F_{t+h,t}^{heat} - F_{t+h,t}^{oil}$$

$$\widehat{R_{t+h|t}}^{oil} = R_t^{oil} exp\{\hat{\alpha} + \hat{\beta}\left[\boldsymbol{f_{t+h,t}^{CS}}\right] - E\left(\pi_{t+h}^{h}\right)\}$$

The Weighted Product Futures Spread Model

- Apply same weights for gasoline and heating oil futures prices as used in the crack spread but together with the current spot price for crude oil

$$\widehat{R_{t+h|t}}^{oil} = R_t^{oil} exp\left\{\hat{\alpha} + \hat{\beta}\left[\log\left(\boldsymbol{\frac{2}{3}F_{t+h,t}^{gas}} + \boldsymbol{\frac{1}{3}F_{t+h,t}^{heat}}\right) - p_t^{oil}\right] - E\left(\pi_{t+h}^{h}\right)\right\}$$

Equal-Weighted Forecast Combination of Single Product Futures Spread Models

- A more flexible approach to combining information from single product spreads is to assign equal weight to the gasoline and heating oil futures spread forecasts

$$\widehat{R_{t+h|t}}^{oil} = \frac{1}{2}\sum_{k=1}^{2}\widehat{R_{t+h|t}}^{oil,k}$$

Using Spot Prices in Product Spread Models

- Product prices and crude oil prices move together in the long run (e.g., Lanza et al. 2005; Kilian 2010)

- Postulate cointegration between product spot prices and the spot price of crude oil such that

$$\log\left(\sum_{i=1}^{n} w_i P_t^i\right) - p_t^{oil} \sim I(0)$$

- Forecast of the real price of oil:

$$\widehat{R_{t+h|t}}^{oil} = R_t^{oil} exp\left\{\hat{\alpha} + \hat{\beta}\left[\log\left(\sum_{i=1}^{n} w_i \boldsymbol{P_t^i}\right) - p_t^{oil}\right] - E\left(\pi_{t+h}^{h}\right)\right\}$$

Key Parameters for Real-Time Forecasting Horserace

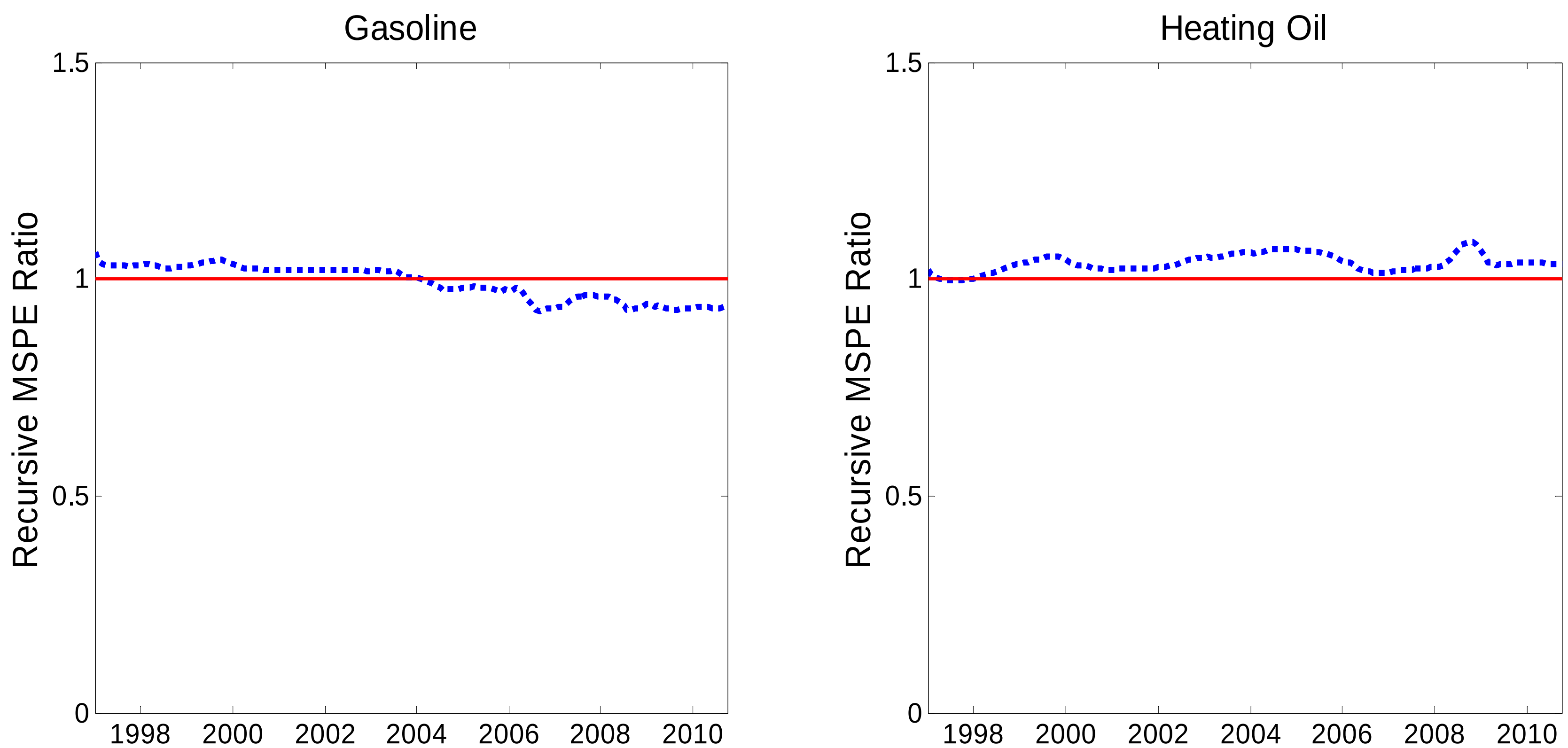
- Real-time out-of-sample forecasts for the monthly real WTI spot price and real refiners' acquisition cost of oil imports
- Evaluation period: 1992.1 to 2012.9
- Data for 1992.1-2012.9 in the 2013.3 vintage are treated as ex-post revised data when evaluating the forecast accuracy
- Evaluation criteria:
    - Recursive mean-squared prediction error (MSPE)
    - Success ratio
- Futures prices and spot prices for gasoline and heating oil
- Extend forecast horizon to 2 years: $h \in \{1,3,6,9,12,15,18,21,24\}$

Forecast Accuracy of Spot Spread Models for Real WTI Price

	$\hat{\alpha}, \hat{\beta}$				$\alpha = 0, \hat{\beta}$			
Horizon	(1) Gasoline spot spread	(2) Heating oil spot spread	(3) Equal-weighted combination of (1) and (2)	(4) Weighted spot product spread	(1) Gasoline spot spread	(2) Heating oil spot spread	(3) Equal-weighted combination of (1) and (2)	(4) Weighted spot product spread
				MSPE Ratios Relative to No-Change Forecast				
1	1.015	1.010	1.009	1.017	**0.999**	1.008	1.003	1.002
3	1.032	1.028	1.018	1.040	**0.998**	1.023	1.008	1.007
6	1.015	1.043	1.024	1.032	**0.978****	1.037	1.006	**0.998**
9	1.067	1.056	1.060	1.067	**0.965***	1.052	1.006	**0.989**
12	1.016	1.051	1.028	1.057	**0.940***	1.040	**0.987**	**0.970****
15	**0.993**	1.035	1.004	1.053	**0.936***	1.031	**0.980**	**0.966**
18	1.026	1.006	1.011	1.062	**0.969****	1.041	1.001	**0.990**
21	1.025	**0.995**	1.006	1.048	**0.987**	1.058	1.017	1.004
24	**0.979**	1.006	**0.981**	1.018	**0.940***	1.054	**0.990**	**0.968****
				Success Ratios				
1	0.462	**0.546**	**0.506**	0.454	**0.554**	**0.534**	**0.566**	**0.562**
3	0.445	**0.518**	0.494	0.453	**0.575**	0.482	**0.555**	**0.563**
6	0.443	**0.557**	**0.512**	0.455	**0.541**	0.459	**0.508**	**0.541**
9	0.461	**0.585**	**0.573**	0.494	0.419	0.419	0.465	0.469
12	0.445	**0.576**	**0.525**	0.483	**0.504**	0.370	0.416	0.437
15	0.477	**0.592****	**0.506**	0.443	0.494	0.434	0.396	0.438
18	0.474	**0.603***	**0.530**	**0.535**	0.440	0.397	0.440	0.435
21	0.485	**0.555**	**0.520**	**0.507**	0.437	0.349	0.384	0.415
24	0.451	**0.531**	0.465	**0.504**	0.491	0.367	0.416	0.474

Unusual or systematic?

Real-Time Recursive MSPE Ratio of α=0 Spot Spread Models Relative to No-Change Forecast of Real RAC at h=24 Months



Toward a generalization of product spread models

- Gasoline spread models tend to be more accurate than heating oil spread models.

- Global price of crude oil is determined by the refined product in highest demand.

- Predictive relationship might not be stable: allow weights assigned to gasoline and heating oil spreads to evolve smoothly over time.

Two forecasting approaches

1. Forecast combinations based on inverse MSPE weights:

$$\widehat{R_{t+h|t}}^{oil} = \sum_{i=1}^{2} v_{k,t}\, \widehat{R_{t+h|t}}^{oil,k}$$

with $\quad v_{k,t} = \frac{m_{k,t}^{-1}}{\sum_{j=1}^{2} m_{j,t}^{-1}}$

where $m_{k,t}$ is the rolling or recursive MSPE of model k in period t

2. Time-varying parameter model:

$$\Delta s_{t+h}^{oil} = \alpha_t + \beta_{1t}\left(s_t^{gas} - s_t^{oil}\right) + \beta_{2t}\left(s_t^{heat} - s_t^{oil}\right) + \varepsilon_{t+h}$$

where s_t^{gas} is the log of the nominal US spot price of gasoline
s_t^{heat} is the log of the nominal US spot price of No.2 heating oil
s_t^{oil} is the log nominal WTI spot price of crude oil.

and $\theta_t = \theta_{t-1} + \xi_t$ with $\theta_t = [\alpha_t \;\; \beta_{1t} \;\; \beta_{2t}]'$

Construct forecast by Monte Carlo integration as the mean of simulated forecasts conditional on the most recent data:

$$R_{t+h|t}^{oil} = R_t^{oil} \exp\left(\hat{\alpha}_t + \hat{\beta}_{1t}\left[s_t^{gas} - s_t\right] + \hat{\beta}_{2t}\left[s_t^{heat} - s_t\right] - E_t(\pi_{t+h}^{h})\right)$$

Forecast Accuracy of TVP Product Spread Models for the Real U.S. Refiners' Acquisition Cost for Oil Imports

Horizon	Gasoline and Heating Oil Spreads		Gasoline Spread Only	
	$\hat{\alpha}_t, \hat{\beta}_{1t}, \hat{\beta}_{2t}$	$\hat{\alpha}_t = 0,$ $\hat{\beta}_{1t}, \hat{\beta}_{2t}$	$\hat{\alpha}_t, \hat{\beta}_{1t}$	$\hat{\alpha}_t = 0, \hat{\beta}_{1t}$
	Spot spread	Spot spread	Spot spread	Spot spread
	MSPE Ratios Relative to No-Change Forecast			
1	**0.98**	**1.00**	**0.97**	**0.99**
3	**0.99**	**1.00**	**0.99**	1.01
6	**0.96**	**0.97**	**0.96**	1.01
9	**0.91**	**0.89**	**0.90**	**0.96**
12	**0.89**	**0.87**	**0.85**	**0.91**
15	**0.95**	**0.88**	**0.86**	**0.90**
18	1.01	**0.91**	**0.95**	**0.92**
21	1.11	**0.91**	**0.93**	**0.91**
24	**0.98**	**0.80**	**0.87**	**0.81**

Forecast Accuracy of TVP Product Spread Models for the Real U.S. Refiners' Acquisition Cost for Oil Imports

Horizon	Gasoline and Heating Oil Spreads		Gasoline Spread Only	
	$\hat{\alpha}_t, \hat{\beta}_{1t}, \hat{\beta}_{2t}$	$\hat{\alpha}_t = 0,$ $\hat{\beta}_{1t}, \hat{\beta}_{2t}$	$\hat{\alpha}_t, \hat{\beta}_{1t}$	$\hat{\alpha}_t = 0, \hat{\beta}_{1t}$
	Spot spread	Spot spread	Spot spread	Spot spread
		Success Ratios		
1	**0.50**	**0.59**	**0.57**	**0.59**
3	**0.52**	**0.58**	**0.56**	**0.59**
6	**0.50**	**0.54**	**0.53**	**0.58**
9	0.49	**0.56**	**0.51**	**0.59**
12	0.48	**0.61**	**0.52**	**0.62**
15	0.44	**0.65**	**0.53**	**0.64**
18	0.50	**0.66****	**0.55**	**0.65**
21	**0.52**	**0.66**	**0.65**	**0.66**
24	**0.50**	**0.63**	**0.58**	**0.63**

Punchline

- Product spot price spreads contain useful predictive information for horizons between one and two years.

- Model that allows for smooth structural change is the most accurate product spread forecasting approach.

- Proliferation of forecasting models for the real price of crude oil with different strength and weaknesses: Which one to use in practice? (Baumeister and Kilian, 2014; Baumeister, Kilian and Lee, 2014)

Three Reasons for Considering Forecast Combinations

1. Some forecasting models are more accurate at short horizons and others at longer horizons.

2. Even the forecasting model with the lowest MSPE may potentially be improved by incorporating information from other models with higher MSPE.

3. Even the most accurate forecasting models do not work equally well at all times.

One can think of forecast combinations as providing insurance against possible model misspecification and smooth structural change.

Most Promising Candidate Models for Forecast Combination

1. No-change forecast (Hamilton 2009)
2. Iterated forecasts from recursively estimated VAR(12) model motivated by global oil market model of Kilian and Murphy (2014)
3. WTI oil futures spread model (Alquist and Kilian 2010)
4. Model based on changes in non-oil industrial raw materials (Baumeister and Kilian 2012)
5. Gasoline spot spread model with $\alpha = 0$
6. TVP model of gasoline and heating oil spot price spreads

Real-Time Forecast Accuracy of Forecast Combinations

	Real U.S. RAC for Oil Imports 6 MODELS		Real WTI Price 6 MODELS	
	Equal Weights	Recursive Weights	Equal Weights	Recursive Weights
Horizon	**Recursive MSPE Ratios**			
1	**0.922**	**0.927**	**0.911**	**0.917**
3	**0.906**	**0.912**	**0.906**	**0.914**
6	**0.957**	**0.964**	**0.962**	**0.968**
9	**0.948**	**0.955**	**0.952**	**0.958**
12	**0.912**	**0.918**	**0.920**	**0.925**
15	**0.913**	**0.922**	**0.922**	**0.934**
18	**0.962**	**0.979**	**0.963**	**0.986**
21	1.025	1.030	1.023	1.032
24	**0.992**	**0.987**	**0.984**	**0.987**
Horizon	**Success Ratios**			
1	**0.570***	**0.588****	**0.517**	**0.512**
3	**0.588***	**0.592***	**0.567****	**0.576***
6	**0.556**	**0.535**	**0.543**	**0.517**
9	**0.575**	**0.575**	**0.562**	**0.562**
12	**0.614***	**0.614***	**0.605***	**0.596***
15	**0.645***	**0.626***	**0.612***	**0.608***
18	**0.611***	**0.553***	**0.572****	**0.548***
21	**0.550**	**0.564**	**0.550**	**0.574**
24	**0.566**	**0.531**	**0.556**	**0.520**

Real-Time Forecast Accuracy of Forecast Combinations

	Real U.S. RAC for Oil Imports			Real WTI Price		
	6 MODELS		Drop NC and GAS SPREAD	6 MODELS		Drop NC and GAS SPREAD
	Equal Weights	Recursive Weights	Equal Weights	Equal Weights	Recursive Weights	Equal Weights
Horizon	**Recursive MSPE Ratios**					
1	**0.922**	**0.927**	**0.897**	**0.911**	**0.917**	**0.880**
3	**0.906**	**0.912**	**0.874**	**0.906**	**0.914**	**0.873**
6	**0.957**	**0.964**	**0.949**	**0.962**	**0.968**	**0.956**
9	**0.948**	**0.955**	**0.939**	**0.952**	**0.958**	**0.943**
12	**0.912**	**0.918**	**0.892**	**0.920**	**0.925**	**0.902**
15	**0.913**	**0.922**	**0.893**	**0.922**	**0.934**	**0.906**
18	**0.962**	**0.979**	**0.957**	**0.963**	**0.986**	**0.959**
21	1.025	1.030	1.065	1.023	1.032	1.064
24	**0.992**	**0.987**	1.029	**0.984**	**0.987**	1.017
Horizon	**Success Ratios**					
1	**0.570***	**0.588****	**0.554***	**0.517**	**0.512**	**0.517**
3	**0.588***	**0.592***	**0.609***	**0.567****	**0.576***	**0.592***
6	**0.556**	**0.535**	**0.556**	**0.543**	**0.517**	**0.543**
9	**0.575**	**0.575**	**0.580****	**0.562**	**0.562**	**0.562**
12	**0.614***	**0.614***	**0.609***	**0.605***	**0.596***	**0.605***
15	**0.645***	**0.626***	**0.650***	**0.612***	**0.608***	**0.617***
18	**0.611***	**0.553***	**0.601***	**0.572****	**0.548***	**0.577****
21	**0.550**	**0.564**	**0.550**	**0.550**	**0.574**	**0.550**
24	**0.566**	**0.531**	**0.561**	**0.556**	**0.520**	**0.551**

Conclusion

- Combining the best performing oil price forecasting models with equal weights dominates selecting one model and using it for all horizons.

- Even more accurate forecasts are obtained when allowing the forecast combination to change across horizons.

- This approach is not always more accurate than the single most accurate model by horizon, but its accuracy is much more stable.