BANK OF CANADA
BANQUE DU CANADA

Forecasting the Real Price of Oil in a Changing World: A Forecast Combination Approach

by Christiane Baumeister and Lutz Kilian

Forecasting the Real Price of Oil in a Changing World: A Forecast Combination Approach

by

Christiane Baumeister[1] and Lutz Kilian[2]

[1]International Economic Analysis Department
Bank of Canada
Ottawa, Ontario, Canada K1A 0G9
cbaumeister@bankofcanada.ca

[2]Department of Economics
University of Michigan
Ann Arbor, MI 48109-1220
Corresponding author: lkilian@umich.edu

Acknowledgements

Abstract

The U.S. Energy Information Administration regularly publishes short-term forecasts of the price of crude oil. Traditionally, such out-of-sample forecasts have been largely judgmental, making them difficult to replicate and justify, and not particularly successful when compared with naïve no-change forecasts, as documented in Alquist, Kilian and Vigfusson (2013). Recently, a number of alternative econometric oil price forecasting models have been introduced in the literature and shown to be more accurate than the no-change forecast of the real price of oil. We investigate the merits of constructing real-time forecast combinations of six such models with weights that reflect the recent forecasting success of each model. Forecast combinations are promising for four reasons. First, even the most accurate forecasting models do not work equally well at all times. Second, some forecasting models work better at short horizons and others at longer horizons. Third, even the forecasting model with the lowest mean-squared prediction error (MSPE) may potentially be improved by incorporating information from other models with higher MSPEs. Fourth, one can think of forecast combinations as providing insurance against possible model misspecification and smooth structural change. We demonstrate that over the past 20 years suitably constructed real-time forecast combinations would have been more accurate than the no-change forecast at every horizon up to two years. Relative to the no-change forecast, forecast combinations reduce the MSPE by up to 18 per cent. They also have statistically significant directional accuracy as high as 77 per cent. We conclude that suitably constructed forecast combinations should replace traditional judgmental forecasts of the price of oil.

JEL classification: Q43, C53, E32
Bank classification: Econometric and statistical methods; International topics

Résumé

L'Energy Information Administration des États-Unis publie périodiquement des prévisions à court terme du prix du pétrole brut. Les prévisions de ce type établies hors échantillon ont jusqu'à présent fait une large place au jugement, ce qui les rend difficiles à reproduire et à justifier. En outre, elles ne sont pas plus précises qu'une prévision naïve du prix de l'or noir basée sur une marche aléatoire, ainsi que le montrent Alquist, Kilian et Vigfusson (2013). Des modèles économétriques d'un nouveau genre, qui donnent des prévisions plus exactes du prix réel du pétrole que cette prévision naïve, ont récemment fait leur apparition dans la littérature. Les auteurs se penchent sur l'intérêt d'agréger les prévisions en temps réel de six de ces modèles en attribuant à ces derniers des poids conformes à leur efficacité récente en prévision. La combinaison de prévisions apparaît comme une voie prometteuse pour quatre raisons. D'abord, même les modèles offrant la meilleure qualité prédictive peuvent voir leur efficacité varier d'une période à l'autre. Deuxièmement, certains modèles produisent des prévisions plus justes aux horizons rapprochés, et d'autres, aux horizons éloignés. Troisièmement, même le modèle qui

présente la plus petite erreur quadratique moyenne de prévision (EQMP) est susceptible d'être amélioré par l'intégration d'information issue d'autres modèles ayant une EQMP plus élevée. Enfin, les combinaisons de prévisions peuvent être conçues comme un moyen de se prémunir contre de possibles erreurs de spécification et d'éventuels changements structurels progressifs. Les auteurs montrent que, sur les vingt dernières années, des combinaisons de prévisions en temps réel convenablement construites auraient donné de meilleurs résultats qu'un modèle de marche aléatoire aux horizons qui ne dépassent pas deux ans. Par rapport à celui-ci, les combinaisons de prévisions se distinguent par une baisse de l'EQMP qui peut atteindre 18 %. De plus, les taux de réussite observés dans la prévision du sens des variations s'élèvent jusqu'à 77 % et sont significativement supérieurs à 50 %. Les auteurs concluent que des combinaisons de prévisions convenablement construites devraient remplacer les prévisions du prix du pétrole qui font traditionnellement intervenir le jugement.

Classification JEL : Q43, C53, E32
Classification de la Banque : Méthodes économétriques et statistiques; Questions internationales

Non-Technical Summary

Recently, a number of alternative econometric oil price forecasting models have been introduced in the literature and shown to be more accurate than the no-change forecast of the real price of oil.

In this paper, we investigate the merits of constructing real-time forecast combinations of six such models with weights that reflect the recent forecasting success of each model. The forecasting models are: forecasts from vector autoregressive models of the global oil market that include the key determinants of the real price of oil according to economic theory, forecasts based on recent changes in the price index of non-oil industrial raw materials, forecasts based on West Texas Intermediate (WTI) oil futures prices, the no-change forecast, forecasts based on the spread of the U.S. spot price of gasoline relative to the WTI spot price of crude oil, and a time-varying parameter forecasting model allowing the U.S. gasoline spread and the U.S. heating oil spread to contribute to the oil price forecast with smoothly varying weights.

Forecast combinations are promising for four reasons. First, even the most accurate forecasting models do not work equally well at all times. Second, some forecasting models work better at short horizons and others at longer horizons. Third, even the forecasting model with the lowest mean-squared prediction error (MSPE) may potentially be improved by incorporating information from other models with higher MSPE. Fourth, one can think of forecast combinations as providing insurance against possible model misspecification and smooth structural change.

Our results show that real-time forecast combinations lead to a considerable reduction in MSPE relative to the no-change forecast. We also demonstrate that they have statistically significant directional accuracy. Over the past 20 years, suitably constructed real-time forecast combinations would have produced more accurate out-of-sample forecasts of the real price of oil than the no-change forecast at every horizon up to two years.

1. Introduction

Since long-term oil contracts were abandoned around 1980, one of the most challenging forecasting problems has been how to forecast the price of crude oil. One of the few regular producers of oil price forecasts has been the U.S. Energy Information Administration (EIA), which constructs monthly and quarterly forecasts of the price of crude oil at horizons up to two years. EIA oil price forecasts help guide natural resource development and investments in infrastructure. They also play an important role in preparing budget and investment plans. Users of oil price forecasts include international organizations, central banks, governments at the state and federal level, and a range of industries including utilities and automobile manufacturers.

Traditionally, the EIA's short-term oil price forecasts have been largely judgmental, making them difficult to replicate and justify. Nor have these forecasts been particularly successful when compared with naïve no-change forecasts, as documented in Alquist, Kilian and Vigfusson (2013). Indeed, many pundits have suggested that changes in the price of oil are inherently unforecastable and that attempts to forecast the price of crude oil are pointless. These agnostics view the current price of oil as the best forecast of the future price (see, e.g., Davies 2007; Hamilton 2009). In recent years, however, a number of new econometric forecasting models have been introduced in the literature and shown to be more accurate out of sample than the no-change forecast of the real price of oil, even after taking account of real-time data constraints (see, e.g., Baumeister and Kilian 2012, 2013a; Baumeister, Kilian and Zhou 2013).[1]

In this paper, we investigate the merits of constructing combinations of six such forecasting models with weights that reflect each model's recent forecasting success. The objective is to generate forecasts that do not require judgment and are available in real time. We restrict attention to forecast horizons between 1 month and 24 months, consistent with the objective of the EIA. The forecasting

---

[1] In recent years, there has been a resurgence in research on how to forecast the price of commodities in general and the price of oil in particular, at least at horizons up to a year. This literature has examined in depth the predictive power of oil futures prices, the predictive content of changes in oil inventories, oil production, macroeconomic fundamentals, product spreads, and exchange rates, as well as the forecasting ability of professional and survey forecasts. Other contributors to this literature include Chernenko, Schwarz and Wright 2004; Knetsch 2007; Sanders, Manfredo and Boris 2008; Alquist and Kilian 2010; Chen, Rogoff and Rossi 2010; Reeve and Vigfusson 2011; Baumeister and Kilian 2013b; Chinn and Coibion 2013.

models considered include forecasts from vector autoregressive (VAR) models of the global oil market, forecasts based on recent changes in the price index of non-oil industrial raw materials, forecasts based on West Texas Intermediate (WTI) oil futures prices, the no-change forecast, forecasts based on the spread of the U.S. spot price of gasoline relative to the WTI spot price of crude oil, and a time-varying parameter (TVP) forecasting model allowing the U.S. gasoline spread and the U.S. heating oil spread to contribute to the oil price forecast with smoothly varying weights. All models are estimated recursively, as is standard in this literature, and subject to real-time data constraints.[2] The weights attached to each forecast are constructed in real time as well. We consider inverse mean-squared prediction error (MSPE) weights based on recursive and rolling windows.

Forecast combinations are promising for four reasons. First, even the most accurate forecasting models do not work equally well at all times. Baumeister and Kilian's (2012) oil price forecasting model, for example, works well during times when economic fundamentals show persistent variation, as was the case between 2002 and 2011, but less at other times. Likewise, there is considerable variation over time in the ability of oil futures prices to forecast the price of oil.

Second, previous research has shown that some forecasting models work better at short horizons and others at longer horizons. For example, forecasting models based on economic fundamentals tend to enjoy superior accuracy at horizons up to 3 months, whereas models based on the spread of refined product prices relative to the price of crude oil tend to be most accurate at horizons between 12 and 24 months.

Third, even the forecasting model with the lowest MSPE may potentially be improved upon by incorporating information from other models with higher MSPEs. For example, Baumeister and Kilian (2013a) show that simple equal-weighted averages of forecasts based on oil futures prices and forecasts based on VAR models of the global oil market are more accurate than either model alone. This evidence suggests that there is little point in ranking models to determine the one model that is most accurate.

---

[2] See Baumeister and Kilian (2013a) for a comparison of oil price forecasts based on rolling and recursive regression estimates.

Whether forecast combinations will improve on the single most accurate model is by no means a foregone conclusion, however. Baumeister and Kilian (2013a), for example, find that one and only one forecast combination among the many forecasting models they consider improves on its individual components.

Fourth, one can think of forecast combinations as providing insurance not only against possible model misspecification, but also against smooth structural change. Such structural change may arise, for example, from changes in market structure, in the structure of the global economy or in the accessibility of crude oil.

The focus of our analysis is both the U.S. refiners' acquisition cost for crude oil imports, which is commonly viewed as a proxy for the global price of oil, and the spot price of WTI crude oil commonly cited in the press. There are two problems with modelling WTI prices. One is that the WTI spot price was subject to government regulation until the early 1980s and hence is not representative of the market price of oil. The other is that the WTI price has suffered from structural instability since 2011, when restrictions on U.S. crude oil exports prevented arbitrage between the WTI price in the United States and the price of Brent crude oil in the United Kingdom. As a consequence, generating WTI forecasts in some cases requires suitable modifications of the baseline forecasting model (see Baumeister and Kilian 2013a).[3]

For each oil price series, we investigate a number of different approaches to combining forecasts. Although the accuracy of the forecast combinations remains remarkably robust across alternative specifications, our results indicate that inverse MSPE weights based on the most recent 12 months of data generate more accurate forecasts than weights based on longer windows or weights obtained from recursive estimates. We also demonstrate that, in practice, the inclusion of only four forecasting methods – the oil market VAR model, a model based on non-oil commodity prices, a model based on oil futures

---

[3] We do not report results for the Brent price of crude oil. The reason is that there are no suitable data available for applying some of the forecasting models considered in this paper. For example, there do not exist long-enough time series for Brent futures prices at longer maturities, and no suitable spot price data are available for the Rotterdam gasoline and heating oil markets. Without access to these data, there is no reason to expect forecast combinations to replicate the successes reported in this paper for other oil price measures. We note, however, that the Brent price has remained stable in relation to the U.S. refiners' acquisition cost for oil imports even in recent years, so to some extent our results for the refiners' acquisition cost are expected to be representative of the Brent price and could be mapped into forecasts of that price, as discussed in Baumeister and Kilian (2013a).

spreads and a time-varying product spread model – suffices to generate the most accurate forecast combination. We show, in particular, that allowing the combination to assign positive weight to the no-change forecast increases the MSPE at all horizons, providing a powerful argument against the agnostic position that the real price of oil is unforecastable.

Our results demonstrate that over the past 20 years suitably constructed real-time forecast combinations would have been more accurate than the no-change forecast at every horizon up to two years. Relative to the no-change forecast, forecast combinations may reduce the MSPE by up to 18 per cent. They also have statistically significant directional accuracy as high as 77 per cent. We conclude that suitably constructed forecast combinations should replace traditional judgmental forecasts of the real price of oil.

The remainder of the paper is organized as follows. In section 2, we briefly review the six forecasting models to be combined in section 3 and the data used in constructing these forecasts. In section 3, we investigate how to choose the weights used in the forecast combination. Section 4 formally evaluates whether the forecast accuracy may be improved further by dropping one or more of the models under consideration. We show that a more parsimonious forecast combination involving only four forecasting models performs best overall. We also examine how the gains in accuracy of this combination forecast accumulate over time. In section 5, we discuss a number of visual diagnostics designed to inform users of the forecast combination about how the loadings on each model change in real time. Section 6 extends the analysis to forecasts of the real price of oil at quarterly horizons. We conclude in section 7.

2. Six Approaches to Forecasting the Real Price of Oil

In this section, we briefly review the forecasting models to be combined in section 3. Our focus is on forecasts of the real price of oil at horizons between 1 and 24 months. The maximum forecast horizon is dictated by the needs of the EIA. We also follow the EIA in focusing on monthly averages of the price of oil. Each of the models below has been shown to generate more accurate real-time out-of-sample forecasts than the no-change forecast at least at some forecast horizons, although not all models have been

examined at horizons beyond 12 months. All models are estimated subject to real-time data constraints using a suitably extended and updated version of the real-time database developed for Baumeister and Kilian (2012).

2.1 Forecast Based on a VAR Model of the Global Oil Market

The first model is a reduced-form VAR model that includes the key variables relevant to the determination of the real price of oil in global markets:

$$B(L)y_t = v + u_t ,$$

where $y_t = \left[\Delta prod_t, rea_t, r_t^{oil}, \Delta inv_t \right]'$ refers to a vector including the per cent change in global crude oil production, a measure of global real activity, the log of the U.S. refiners' acquisition cost for crude oil imports deflated by the log of the U.S. CPI, and the change in global crude oil inventories; v denotes the intercept; and u_t is a white noise innovation. The inventory data are constructed by multiplying U.S. crude oil inventories by the ratio of OECD petroleum inventories to U.S. petroleum inventories. Petroleum inventories are defined to include both stocks of crude oil and stocks of refined products. The inventory data are from the EIA's *Monthly Energy Review*, which also provides data on global oil production and the refiners' acquisition cost. The global real activity index is constructed from freight rate data for global dry cargo ocean shipping as described in Kilian (2009).

This VAR model can be viewed as the reduced-form representation of the structural global oil market model developed in Kilian and Murphy (2013). Its forecast accuracy has been examined extensively in the literature (see, e.g., Alquist, Kilian and Vigfusson 2013; Baumeister and Kilian 2012, 2013a). Throughout the paper, we estimate the unrestricted VAR model with 12 autoregressive lags by the method of least squares.[4] Forecasts $\hat{r}_{t+h|h}$ of the log of the real price of oil are constructed iteratively from the estimated VAR model conditional on the most recent data and converted to levels, resulting in the forecast

---

[4] Similar results would be obtained by imposing standard Bayesian priors in estimation.

$$\hat{R}_{t+h|t}^{oil} = \exp(\hat{r}_{t+h|t}^{oil,VAR}) . \tag{1}$$

Forecasts of the real WTI price are constructed from the same VAR model by assuming that the

most recent spread between the log WTI price and the log of the U.S. refiners' acquisition cost remains

unchanged in the future. By rescaling the forecasts of the U.S. refiners' acquisition cost in this manner,

we allow the relationship between the two oil price measures to evolve as a random walk. This approach

has been shown to be more accurate than the simpler approach of replacing the U.S. real refiners'

acquisition cost in the VAR model by the real WTI price (see Baumeister and Kilian 2013a). The WTI

spot price data are from the FRED database of the Federal Reserve Bank of St. Louis, which also provides

the real-time U.S. CPI data used to deflate the two measures of the nominal price of oil.

2.2 Forecast Based on the Price of Non-Oil Industrial Raw Materials

Much of the empirical success of VAR forecasting models of the real price of oil can be traced to the use

of measures of global real economic activity that help capture fluctuations in the demand for industrial

commodities. A much simpler forecasting method – based on the same intuition that there are broad-

based predictable shifts in the demand for globally traded commodities – exploits real-time information

from recent cumulative changes in non-oil industrial commodity price indices. As discussed in

Baumeister and Kilian (2012), a forecast of the real price of oil can be constructed as follows:

$$R_{t+h|t}^{oil} = R_t^{oil} \left(1 + \pi_t^{h,industrial\ raw\ materials} - E_t(\pi_t^h) \right), \tag{2}$$

where $\pi_t^{h,industrial\ raw\ materials}$ stands for the per cent change of an index of the spot price of industrial raw

materials (other than oil) over the preceding h months. This index is available in real time from the

Commodity Research Bureau. The term $E_t(\pi_t^h)$ is the expected inflation rate over the next h periods. In

practice, this expectation is proxied by recursively constructed averages of past U.S. inflation data,

starting in July 1986.[5]

---

[5] Undoubtedly, the inflation forecast could be refined further, but there is little loss in generality in our approach, given that fluctuations in the nominal price of oil dominate the evolution of the real price of oil.

7

2.3 No-Change Forecast

Baumeister and Kilian (2012) show that forecasting models based on economic fundamentals such as the VAR model described above perform best during times of persistent and hence predictable fluctuations in economic fundamentals. That is why the VAR forecasting model does particularly well before, during and after the Great Recession of 2008. In contrast, during other times the model is only about as accurate as the no-change forecast. Indeed, there are indications that, since 2011, we have once again entered a period during which VAR models offer at best minimal gains relative to the no-change forecast, and that this situation will persist as long as the global economy stagnates. This observation raises the question of whether we would be better off at times if we replaced the VAR forecast by the no-change forecast, or at least downweighted the VAR forecast relative to the no-change forecast. This line of reasoning suggests that we want to allow for the forecast combination to put positive weight on the no-change forecast:

$$\hat{R}^{oil}_{t+h|t} = R^{oil}_t, \tag{3}$$

where R^{oil}_t denotes the real price of oil in levels (as opposed to logs).

2.4 Forecast Based on Oil Futures Prices

Yet another approach is to exploit information from oil futures markets. Many practitioners rely on the price of oil futures contracts in generating forecasts of the nominal price of oil. This forecast can then be converted to a forecast of the real price of oil by subtracting expected inflation. This approach is embodied in the forecasting model

$$R^{oil}_{t+h|t} = R^{oil}_t \left(1 + f^h_t - s_t - E(\pi^h_t)\right), \qquad h = 1,3,6,9,12,15,18, \tag{4}$$

where R_t denotes the current level of the real price of oil, f^h_t is the log of the current WTI oil futures price for maturity h, s_t is the corresponding WTI spot price and $E_t(\pi^h_t)$ is again the expected inflation rate over the next h periods. Both f^h_t and s_t are available in real time. The oil futures prices are obtained from Bloomberg. Although forecasts based on (4) are not significantly more accurate than the no-change forecast at horizons of 1, 3 or 6 months (and sometimes less accurate), especially in recent years the

accuracy of futures-based forecasts at horizons of 9 and 12 months has improved. In this paper, we use the monthly WTI oil futures price data up to a horizon of 18 months, which is the maximum horizon for which the construction of continuous monthly time series is feasible, given our evaluation period. This means that for horizons beyond 18 months the futures-based forecast receives zero weight in the forecast combinations we construct in sections 3 through 6.

2.5 Spread between the Spot Prices of Gasoline and Crude Oil

Another promising class of oil price forecasting models involves the use of product spreads. Many market practitioners believe that a rising spread between the price of gasoline and the price of crude oil signals upward pressures on the price of crude oil. For example, Goldman Sachs in April 2013 cut its oil price forecast, citing significant pressure on product spreads, which it interpreted as an indication of reduced demand for products (see Strumpf 2013). Such a forecasting model was derived in Baumeister, Kilian and Zhou (2013):

$$\hat{R}_{t+h|t}^{oil} = R_t^{oil} \exp\left\{\hat{\beta}\left[s_t^{gas} - s_t\right] - E_t(\pi_t^h)\right\}, \tag{5}$$

where s_t^{gas} is the log of the nominal U.S. spot price of gasoline, s_t is the log of the spot price of WTI crude oil as defined earlier and $\hat{\beta}$ is obtained from estimating the model

$$\Delta s_{t+h|t} = \beta\left[s_t^{gas} - s_t\right] + \varepsilon_{t+h}$$

recursively by the method of least squares. It can be demonstrated that imposing an intercept of zero, as shown in (5), greatly enhances the out-of-sample accuracy of this model. This gasoline spread model greatly improves on the accuracy of a no-change forecast, especially at horizons beyond one year, making it a natural complement to models based on economic fundamentals, which are most accurate at shorter horizons. The gasoline spot price data are readily available in real time from the EIA. For further details see Baumeister, Kilian and Zhou (2013).

2.6 Time-Varying Model of the Gasoline and Heating Oil Spreads

The simplicity of the forecast based on the gasoline spread is appealing, yet there are reasons to be wary.
One concern is that the price of crude oil is likely to be determined by the refined product in highest
demand. According to Verleger (2011), traditionally, in the United States this product has been gasoline,
but more recently it has been heating oil (which is nearly equivalent to diesel fuel), suggesting a
forecasting model that allows for both a gasoline spread and a heating oil spread with time-varying
coefficients. Another concern is that crude oil supply shocks, local capacity constraints in refining,
changes in environmental regulations or other market turmoil may all temporarily undermine the
predictive power of product spreads. These considerations motivate the following generalization of model
(5), introduced in Baumeister, Kilian and Zhou (2013).

We first recursively estimate the time-varying regression model

$$\Delta s_{t+h|t} = \beta_{1t} \left[s_t^{gas} - s_t \right] + \beta_{2t} \left[s_t^{heat} - s_t \right] + \varepsilon_{t+h},$$

where the additional variable s_t^{heat} is the log of the nominal U.S. spot price of heating oil. The product
prices are from the EIA. For details on the data sources, the reader is referred to Baumeister, Kilian and
Zhou (2013). In estimating the model, we postulate that $\varepsilon_{t+h} \sim NID(0,\sigma^2)$, while the time-varying
coefficients $\theta_t = \begin{bmatrix} \beta_{1t} & \beta_{2t} \end{bmatrix}'$ evolve according to a random walk as $\theta_t = \theta_{t-1} + \xi_t$, and ξ_t is independent
Gaussian white noise with variance Q. The intercept has again been restricted to zero, following
Baumeister, Kilian and Zhou (2013), who show that this restriction greatly improves the out-of-sample
accuracy. This state-space model is estimated using a Gibbs sampling algorithm. The conditional
posterior of θ_t is normal, and its mean and variance can be derived via standard Kalman filter recursions
(see Kim and Nelson 1999). Conditional on an estimate of θ_t, the conditional posterior distribution of
σ^2 is inverse Gamma and that of Q is inverse Wishart.

Given the TVP estimates, we then construct the TVP model forecast

$$\hat{R}_{t+h|t}^{oil} = R_t^{oil} \exp\left\{\hat{\beta}_{1t}\left[s_t^{gas} - s_t\right] + \hat{\beta}_{2t}\left[s_t^{heat} - s_t\right] - E_t(\pi_t^h)\right\} \tag{6}$$

by Monte Carlo integration as the mean of the forecasts simulated based on 1,000 Gibbs iterations

conditional on the most recent data. Our forecasts take into account that the model parameters continue to

drift over the forecast horizon according to their law of motion. The first 30 observations of the initial

estimation period are used as a training sample to calibrate the priors and to initialize the Kalman filter.

This TVP product spread model has been shown to be systematically more accurate than the no-

change forecast, especially at horizons beyond one year. At some horizons, it produces forecasts even

more accurate than the gasoline spread forecast (5). Hence, there is reason to believe that this approach

may have additional predictive information not already contained in the simpler gasoline spread model.

3. Baseline Results

Knowing ex post that one or the other forecasting method would have been more accurate is not of much

use to applied forecasters. The challenge is to be able to detect in real time when one model should be

downweighted compared to another. A natural approach to measuring the real-time forecast accuracy of

competing models is to construct inverse MSPE weights based on the recent forecasting performance of

each model. These weights may then be used to construct a suitable weighted average of the forecasting

models in question. This forecast combination approach has a long tradition in econometrics (see, e.g.,

Diebold and Pauly 1987; Stock and Watson 2004). The smaller the MSPE of a model is at date t, the

larger the weight that this model receives in the combination forecast:

$$\hat{R}_{t+h|t}^{oil} = \sum_{k=1}^{6} \omega_{k,t} \hat{R}_{t+h|t}^{oil,k}, \qquad \omega_{k,t} = \frac{m_{k,t}^{-1}}{\sum_{j=1}^{6} m_{j,t}^{-1}},$$

where $m_{k,t}$ is the recursive MSPE of model k in period t. In practice, the MSPE estimates must be

initialized. We proceed by assigning equal weight to each model when entering the evaluation period. For

subsequent periods, we then recursively update the MSPE of each model. The advantage of inverse

MSPE weights is that they allow the forecast combination to adjust according to the recent MSPE of each

model. As we illustrate in section 5, these weights may also be used to monitor the real-time performance of each model, given that they are bounded between 0 and 1 for each model and add up to 1.

All forecasting models are evaluated for the same evaluation period of 1992.1–2012.9. Using such a long evaluation period reduces the odds of spurious fits. The initial estimation period ends in 1991.12. Some forecasting models such as the VAR models are estimated on data back to 1973.2. For other forecasting models, the estimation period starts much later, reflecting the availability of the data. For example, monthly spot prices for gasoline and heating oil are available starting only in 1986. It is important to stress that our data are in many cases subject to real-time data constraints. Where appropriate, we rely on an updated and extended version of the real-time data base developed at the Bank of Canada for the purpose of forecasting oil prices (see Baumeister and Kilian 2012, 2013a,b). We use the real price of oil in the March 2013 vintage up to September 2012 as a proxy for the ex-post revised data, against which all forecasts are evaluated.

We assess the accuracy of various forecast combinations based on their recursive MSPEs over the evaluation period (expressed as a ratio relative to the MSPE of the no-change forecast). MSPE ratios below 1 mean that the forecast in question is more accurate than the no-change forecast. We also examine the directional accuracy of the forecast combinations. Under the null hypothesis of no directional accuracy, the model should be no more successful at predicting the direction of change in the price of oil than would be tossing a fair coin with success probability 0.5, so any success ratio higher than 0.5 indicates an improvement over the no-change forecast. Tests of the null of no directional accuracy are conducted using the test of Pesaran and Timmermann (2009). There are, to our knowledge, no tests for the statistical significance of MSPE reductions for estimated forecast combinations. Standard tests in the literature are based on the premise that we compare the same two models at each point in time. Given that the model weights used in the combination evolve over time, that premise appears to be violated.

It is useful to start with the evidence for the real U.S. refiners' acquisition cost for crude oil imports. The first column of Table 1 shows that the forecast combination based on recursively estimated inverse MSPE weights systematically reduces the recursive MSPE at all horizons but horizon 21; in the

latter case, it is almost as accurate as the no-change forecast. The reductions in the recursive MSPE are as high as 10 per cent at some horizons. This forecast combination also has directional accuracy at all horizons with success ratios as high as 66 per cent. Most success ratios are statistically significant at the 10 per cent level or even the 5 per cent level. The gains in forecast accuracy extend to horizons between 12 and 24 months. It is worth pointing out that some models included in the forecast combination may achieve even larger gains at some horizons, but no individual model produces gains as uniformly large across horizons. This pattern of results is not specific to the real U.S. refiners' acquisition cost for crude oil imports. The fifth column in Table 1 illustrates that similar results hold when forecasting the real WTI price. The reductions in the recursive MSPE are as high as 10 per cent and the success ratios as high as 65 per cent, and mostly statistically significant.

While these results are encouraging, it is not clear that recursive estimates of the inverse MSPE are consistent with the notion of smooth structural change. In the latter case, a more natural approach would be to use rolling windows in estimating the weights. An obvious question concerns the length of these windows. The more pronounced the structural change (or the less stable the individual forecasting models), the shorter the window length should be. At the same time, the window cannot be too short without the estimates of the weights becoming too noisy.

In Table 1 we experiment with three window lengths: 36, 24 and 12 months. Columns 2 through 4 of Table 1 relate to the real U.S. refiners' acquisition cost for crude oil imports. They show that, while the results are remarkably robust across specifications, the most accurate forecasts overall are obtained with a rolling window based on the 12 most recent observations. The latter specification reduces the recursive MSPE ratios below 1 at all horizons, with gains ranging from 5 per cent to 15 per cent, and has directional accuracy between 57 per cent and 73 per cent that is statistically significant at all but one horizon. The extent of the improvements in accuracy even at horizons 15 through 24 is impressive. Much the same pattern applies to the real WTI price in columns 6 through 8. Again, the most accurate forecast combination is obtained with rolling windows based on the last 12 observations. The MSPE reductions range from 5 per cent to 17 per cent and the directional accuracy ranges from 52 per cent to 75 per cent.

In general, the longer the length of the rolling window, the closer the results are to those for the recursively estimated weights.

4. Sensitivity Analysis

Our baseline combination involves six forecasting models. A question of practical interest is whether all of these models are required or whether the set of models may be reduced further. This question may be addressed by recomputing the accuracy of the forecast combination, having eliminated one model at a time from the forecast combination, as shown in Table 2. Our point of departure is the result in Table 1 with weights based on rolling windows consisting of the last 12 observations. Evidence that dropping one of the six models systematically lowers the recursive MSPE ratio would be an indication that this model ought to be eliminated from the forecast combination, if we care about the MSPE outcomes. We illustrate this approach for the case of the real U.S. refiners' acquisition cost for oil imports in Table 2.

The first column of Table 2 shows that leaving out the VAR model more often than not raises the MSPE ratio compared with column 4 in Table 1, reaffirming our decision to include this model in the forecast combination. The second and third columns of Table 2 provide similar evidence for the futures-based forecast and for the non-oil commodity price model. In sharp contrast, the fourth column shows that the recursive MSPE may be lowered at all forecast horizons, if we eliminate the no-change model from the forecast combination. This result contradicts the rationale we provided in section 2 for including the no-change forecast in the forecast combination. A similar systematic improvement in the recursive MSPE can be observed after dropping the gasoline spread model in the fifth column, but not for the TVP product spread model in the last column. This evidence suggests that the gasoline spread model adds nothing beyond the predictive information in the TVP product spread model, and should be eliminated from the forecast combination. It is also worth noting that the VAR model and the commodity price model are particularly useful at short horizons, according to Table 2, whereas the oil futures spread model contributes mainly at medium-term horizons and the TVP product spread model at medium and longer horizons.

14

The first column of Table 3 confirms that, after dropping the no-change forecast and the gasoline spread forecast, the accuracy of the forecast combination at most horizons improves further. The resulting forecast provides a new benchmark against which other oil price forecasts must be judged. A forecast combination consisting only of the VAR model, the commodity price model, the oil futures spread model and the TVP product spread model has lower recursive MSPE than the no-change forecast at all horizons from 1 month to 24 months. The reductions in the MSPE range from 6 per cent to 18 per cent. The improvements in directional accuracy are statistically significant at all but one horizon and range from 56 per cent to 74 per cent, depending on the horizon. Similar results hold for the real WTI price, as shown in the second column of Table 3. Because all of these results were achieved in real time, the evidence in Table 3 shows that there is a practical alternative to the construction of judgmental forecasts of the real price of oil.

The most striking result in our analysis is the ability of combination forecasts to outperform the no-change forecast of the real price of oil at horizons between one and two years. An important question is whether the recursive MSPE reductions shown in Table 3 are driven by one or two unusual episodes in the data or whether they are more systematic. Figure 2 addresses this question by plotting the recursive MSPE ratio at the two-year horizon for the evaluation period since 1997. We disregard the earlier MSPE ratios for being based on too short a recursive evaluation period to be considered reliable. For illustrative purposes we focus on the real U.S. refiners' acquisition cost for crude oil imports. The plot shows the evolution of the recursive MSPE ratio over time. The last entry on the right corresponds to the entry for horizon 24 in column 1 of Table 3. Figure 2 shows that the forecast combination in Table 3 has been more accurate than the no-change forecast throughout the entire evaluation period. Similar results hold for other long horizons.

5. Diagnostics

A question of practical interest is how the weights assigned to each model evolve in real time, as more data become available. This question can be answered using some simple diagnostics. The upper panel of

Figure 1 plots the one-month-ahead rolling inverse MSPE weights for the real U.S. refiners' acquisition cost for crude oil imports. By construction the weights sum to 1. All results are based on the forecast combination shown in Table 3. It should be noted that the relative weights of each model may differ considerably by forecast horizon and that the results in Figure 1 are not necessarily representative for longer horizons. They are intended for illustrative purposes only. It is immediately apparent from Figure 1 that the weights show considerable variation over time. For example, the VAR model receives particularly high weight during the Great Recession of 2008 as well as in 1994, 2003 and 2006, but comparatively low weight in 2004/05 and 2012.[6]

Interestingly, the forecast based on industrial commodity prices frequently moves in the opposite direction from the VAR forecast. One possible explanation is that the VAR model provides additional information beyond global real economic activity, including information that is specific to the oil market. It is useful to simplify the analysis further by combining the weights associated with the VAR forecasting model and the model based on non-oil industrial commodity prices. The lower panel of Figure 1 labels this combined weight as "Fundamentals." While fundamentals still receive by far the highest weight in the forecast combination, it is readily apparent that the usefulness of economic fundamentals has declined since 2010. This finding is not unexpected. The global economy has been stagnating in recent years. In the absence of large and predictable variation in economic fundamentals, VAR forecasting models will have no advantage over more parsimonious models, such as the no-change forecast, in determining the direction of the real price of oil, but VAR forecasts will suffer from a much higher variance than the no-change forecast and hence a higher MSPE.

In contrast, the weight attached to forecasts based on oil futures prices and forecasts based on the TVP product spread model recovered after the Great Recession, reaching an all-time high by 2012. There

---

[6] The fact that VAR weights were comparatively low at times does not mean that economic fundamentals did not matter at those times. Rather, it means that predictable variation in the fundamentals was low. Even if we accept that global real economic activity helps forecast the real price of oil, for example, this knowledge is useless for forecasting out of sample, unless we can also predict the evolution of future global real economic activity within the VAR model.

is a general tendency for the weight on the TVP product spread model, in particular, to be inversely related to the weight on economic fundamentals. When one model's performance weakens, the other model receives increasing weight.

6. Extensions to Quarterly Horizons

As mentioned earlier, the EIA forecasts not only the monthly price of oil, but also quarterly averages. The construction of quarterly forecasts has been studied in depth in Baumeister and Kilian (2013a), who show that the most accurate forecasts of the quarterly real price of oil are typically obtained by aggregating forecasts from models estimated at monthly frequency to quarterly frequency. For example, the average of the January, February and March forecasts generated in December of the preceding year would constitute the forecast for the first quarter of the subsequent year.

There are two ways of proceeding. The first method is to construct forecast combinations of the forecasts generated each month for the monthly horizons 1 through 24 (similar to the results shown in Table 3) and then to aggregate the resulting monthly forecasts by quarter. This approach has the advantage that the monthly combination forecasts are fully consistent with the quarterly combination forecasts. An alternative method is to aggregate the monthly forecasts of each individual forecasting method first and then to construct forecast combinations of the resulting quarterly forecasts. We explore both methods in Table 4. The construction of the forecast combinations otherwise follows the approach taken in Table 3. The benchmark is again the no-change forecast based on the most recent monthly real price of oil in each quarter. For the alternative method we rely on a rolling window of length 4, which is the quarterly analogue of the window length for the monthly data. We verify that choosing a length of 8 or 12 produces broadly similar, but less accurate, results.

Given the construction of the quarterly real price of oil as the average of the monthly prices, it is not possible to infer from the results reported in Table 3 how accurate the combination forecasts of the quarterly data will be. The MSPE of the latter also depends on the unknown covariance between the monthly forecasts. Table 4 shows that, nevertheless, the method of first combining the monthly forecasts

17

and then aggregating the monthly combination forecast by quarter (Method 1) performs very well for the real U.S. refiners' acquisition cost for oil imports. Column 1 of Table 4 shows systematic reductions in the recursive MSPE at all horizons ranging from 8 per cent to 18 per cent. The directional accuracy ranges from 60 per cent to 74 per cent and is mostly highly statistically significant. Similar results are obtained for the real WTI price in column 4. Method 2, which first aggregates the individual forecasts by quarter before combining these quarterly forecasts, also performs reasonably well for both oil price measures, but is generally slightly less accurate than Method 1, which has the advantage of ensuring consistency between monthly and quarterly forecasts.

7. Conclusion

One of the challenges faced by producers of short-term oil price forecasts such as the EIA is how to generate real-time forecasts of the price of oil that are more accurate than the no-change forecast. We showed how to construct such forecasts in a timely manner without relying on judgment. Our analysis relied on forecast combinations of several forecasting models that by themselves are superior to the no-change forecast at least at some horizons. These models utilize as predictors lags of the real price of oil; current oil spot prices and oil futures prices; current spot prices in the market for refined products; and current and lagged data on economic fundamentals such as oil production, global real activity, other industrial commodity prices, and changes in crude oil stocks. We selected the weights for each model in real time according to the model's recent forecast accuracy. The advantage of such data-based forecast combinations is that the resulting oil price forecasts tend to be more robust across forecast horizons and over time than even the best individual forecasting models. They provide some insurance not only against possible model misspecification, but also against smooth structural change. Both problems are potentially important concerns when forecasting oil prices.

The most accurate forecasts are obtained based on inverse MSPE weights computed from rolling windows consisting of the last 12 observations. We demonstrated that combinations of forecasts from VAR models of the global oil market, forecasts based on non-oil industrial commodity prices, forecasts

18

based on oil futures prices and forecasts from TVP product spread models, in particular, are systematically more accurate than the no-change forecast at all horizons from 1 month to 24 months. We are not aware of any other forecasting method that accomplishes this objective. Depending on the horizon, our forecast combinations lower the MSPE by as much as 18 per cent relative to the no-change forecast, and they have directional accuracy as high as 77 per cent. Very similar results are obtained for measures of the global price of oil such as the U.S. refiners' acquisition cost, and the WTI price of crude oil. The most striking result in our analysis is the ability of combination forecasts to outperform the no-change forecast of the real price of oil at horizons between one and two years. An important question is whether the recursive MSPE reductions at long horizons are driven by one or two unusual episodes in the data or whether they are more systematic. We found that suitably constructed combination forecasts are more accurate than the no-change forecast throughout the entire evaluation period.

While much of our analysis focused on forecasting the monthly real price of oil, we showed that our results also can be generalized to the problem of forecasting the quarterly real price of oil. Although we do not pursue this extension in this paper, we note that it would be straightforward to extend our analysis to the problem of forecasting the nominal price of oil. This would involve only the additional step of scaling our forecast of the real price of oil by the expected inflation rate over the relevant forecast horizon. We also provided an algorithm for deciding which forecasting models to include in the forecast combination. We used this algorithm to establish that not all of the forecasting models included in the baseline forecast combination contribute toward lower MSPEs, allowing us to eliminate some models from consideration. Finally, we suggested diagnostics for assessing the contribution of each forecasting model to the forecast combination in real time. We quantified the extent to which forecasting models based on economic fundamentals, for example, are more useful during times such as the Great Recession of 2008 than during tranquil times in oil markets, and we discussed likely reasons for the evolution of the forecast weights over time.

References

Alquist, R., and L. Kilian (2010), "What Do We Learn from the Price of Crude Oil Futures?" *Journal of Applied Econometrics*, 25, 539-573.

Alquist, R., Kilian, L., and R.J. Vigfusson (2013), "Forecasting the Price of Oil," forthcoming in: G. Elliott and A. Timmermann (eds.), *Handbook of Economic Forecasting*, 2, Amsterdam: North-Holland.

Baumeister, C., and L. Kilian (2012), "Real-Time Forecasts of the Real Price of Oil," *Journal of Business and Economic Statistics*, 30, 326-336.

Baumeister, C., and L. Kilian (2013a), "What Central Bankers Need to Know about Forecasting Oil Prices," forthcoming: *International Economic Review*.

Baumeister, C., and L. Kilian (2013b), "Real-Time Analysis of Oil Price Risks using Forecast Scenarios," mimeo, University of Michigan.

Baumeister, C., Kilian, L., and X. Zhou (2013), "Are Product Spreads Useful for Forecasting? An Empirical Evaluation of the Verleger Hypothesis," mimeo, University of Michigan.

Chen, Y-C., Rogoff, K.S., and B. Rossi (2010), "Can Exchange Rates Forecast Commodity Prices?" *Quarterly Journal of Economics*, 125, 1145-1194.

Chernenko, S.V., Schwarz, K.B. and J.H. Wright (2004), "The Information Content of Forward and Futures Prices," International Finance Discussion Paper No. 808, Board of Governors of the Federal Reserve System.

Chinn, M., and O. Coibion (2013), "The Predictive Content of Commodity Futures," forthcoming: *Journal of Futures Markets*.

Davies, P. (2007), "What's the Value of an Energy Economist?" Speech presented at the International Association for Energy Economics, Wellington, New Zealand.

Diebold, F.X., and P. Pauly (1987), "Structural Change and the Combination of Forecasts," *Journal of Forecasting*, 6, 21-40.

Hamilton, J.D. (2009), "Understanding Crude Oil Prices," *Energy Journal*, 30, 179-206.

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.

Kilian, L., and D.P. Murphy (2013), "The Role of Inventories and Speculative Trading in the Global Market for Crude Oil," forthcoming: *Journal of Applied Econometrics*.

Kim, C.J., and C.R. Nelson (1999), *State Space Models with Regime Switching: Classical and Gibbs Sampling Approaches with Applications*. Cambridge, MA: MIT Press.

Knetsch, T.A. (2007), "Forecasting the Price of Oil via Convenience Yield Predictions," *Journal of Forecasting*, 26, 527-549.

Pesaran, M.H., and A. Timmermann (2009), "Testing Dependence Among Serially Correlated Multicategory Variables," *Journal of the American Statistical Association*, 104, 325-337.

Reeve, T.A., and R.J. Vigfusson (2011), "Evaluating the Forecasting Performance of Commodity Futures Prices," International Finance Discussion Paper No. 1025, Board of Governors of the Federal Reserve System.

Sanders, D.R., Manfredo, M.R., and K. Boris (2008), "Evaluating Information in Multiple Horizon Forecasts: The DOE's Energy Price Forecasts," *Energy Economics*, 31, 189-196.

Stock, J.H., and M.W. Watson (2004), "Combination Forecasts of Output Growth in a Seven-Country Data Set," *Journal of Forecasting*, 23, 405-430.

Strumpf, D. (2013), "Goldman Cuts the Near-Term Brent Crude Forecast to $100 a Barrel," *Wall Street Journal*, April 23.

Verleger, P.K. (2011), "The Margin, Currency, and the Price of Oil," *Business Economics*, 46, 71-82.

Table 1: Real-Time Forecast Accuracy of Baseline Forecast Combination Based on Six Forecasting Models

| | Real U.S. Refiners' Acquisition Cost for Oil Imports | | | | Real WTI Price | | | |
	Recursive Weights	Rolling Weights Based on Windows of Length:			Recursive Weights	Rolling Weights Based on Windows of Length:		
		36	24	12		36	24	12
Monthly Horizon	Recursive MSPE Ratios				Recursive MSPE Ratios			
1	**0.918**	**0.918**	**0.907**	**0.901**	**0.909**	**0.907**	**0.904**	**0.901**
3	**0.901**	**0.902**	**0.892**	**0.890**	**0.903**	**0.901**	**0.900**	**0.895**
6	**0.957**	**0.961**	**0.955**	**0.950**	**0.961**	**0.962**	**0.961**	**0.954**
9	**0.946**	**0.945**	**0.941**	**0.928**	**0.940**	**0.947**	**0.945**	**0.934**
12	**0.908**	**0.901**	**0.899**	**0.878**	**0.916**	**0.909**	**0.909**	**0.894**
15	**0.904**	**0.889**	**0.885**	**0.850**	**0.913**	**0.894**	**0.892**	**0.865**
18	**0.946**	**0.921**	**0.904**	**0.867**	**0.947**	**0.917**	**0.902**	**0.865**
21	1.006	**0.993**	**0.980**	**0.914**	1.006	**0.993**	**0.982**	**0.910**
24	**0.962**	**0.946**	**0.933**	**0.858**	**0.958**	**0.944**	**0.928**	**0.835**
	Success Ratios				Success Ratios			
1	**0.578***	**0.574***	**0.590***	**0.586***	0.518	0.526	0.530	0.522
3	**0.591***	**0.591***	**0.579****	**0.583****	**0.579***	**0.575***	**0.571****	**0.567****
6	**0.549**	0.533	0.533	0.566	0.533	0.525	0.549	0.562
9	**0.581****	0.564	0.564	**0.610***	0.560	**0.573****	0.560	**0.602***
12	**0.626***	**0.630***	**0.643***	**0.681***	**0.626***	**0.630***	**0.639***	**0.693***
15	**0.655***	**0.672***	**0.681***	**0.732***	**0.638***	**0.660***	**0.664***	**0.741***
18	**0.638***	**0.668***	**0.681***	**0.724***	**0.647***	**0.677***	**0.694***	**0.754***
21	**0.568**	0.563	**0.585***	**0.668***	**0.563***	0.559	**0.590****	**0.604***
24	**0.611****	**0.602****	**0.611****	**0.677***	**0.602****	0.593	**0.615****	**0.695***

NOTES: The models are described in the text. Boldface indicates improvements relative to the no-change forecast. * denotes significance at the 5 per cent level and ** at the 10 per cent level based on the Pesaran and Timmermann (2009) test for the null hypothesis of no directional accuracy. The statistical significance of the MSPE reductions cannot be assessed because none of the currently available tests of equal predictive accuracy applies in this setting.

Table 2: Real-Time Recursive MSPE Ratios of "Leave-One-Out" Forecast Combination with Rolling Weights Based on Windows of Length 12

			Real U.S. Refiners' Acquisition Cost for Oil Imports			
Model Left Out:	VAR	Oil Futures Spread	Commodity Prices	No Change	Gasoline Spread	TVP Product Spread
Monthly Horizon			Recursive MSPE Ratios			
1	**0.952+**	**0.889 -**	**0.913+**	**0.888 -**	**0.891 -**	**0.894 -**
3	**0.930+**	**0.882 -**	**0.920+**	**0.873 -**	**0.878 -**	**0.877 -**
6	**0.964+**	**0.957+**	**0.947 -**	**0.942 -**	**0.946 -**	**0.958+**
9	**0.927 -**	**0.945+**	**0.920 -**	**0.918 -**	**0.924 -**	**0.952+**
12	**0.884+**	**0.908+**	**0.871 -**	**0.861 -**	**0.870 -**	**0.896+**
15	**0.852+**	**0.894+**	**0.858+**	**0.830 -**	**0.839 -**	**0.860+**
18	**0.858 -**	**0.923+**	**0.888+**	**0.849 -**	**0.852 -**	**0.874+**
21	**0.933+**	**0.914**	**0.915+**	**0.909 -**	**0.911 -**	**0.968+**
24	**0.870+**	**0.858**	**0.847 -**	**0.847 -**	**0.857 -**	**0.959+**

NOTES: The models are described in the text. Boldface indicates improvements relative to the no-change forecast. + and – indicate increases and decreases of the MSPE ratio relative to column 4 of Table 3. A positive sign indicates that the model left out would have improved forecast accuracy if included, whereas a negative sign indicates that it would have worsened forecast accuracy. The statistical significance of the MSPE reductions cannot be assessed because none of the currently available tests of equal predictive accuracy applies in this setting.

Table 3: Real-Time Forecast Accuracy of Forecast Combination with Rolling Weights Based on Windows of Length 12 after Dropping No-Change Forecast and Gasoline Spread Forecast

| Monthly Horizon	Real U.S. Refiners' Acquisition Cost for Oil Imports	Real WTI Price
	Recursive MSPE Ratios	
1	**0.875 -**	**0.872**
3	**0.853 -**	**0.859**
6	**0.935 -**	**0.942**
9	**0.910 -**	**0.918**
12	**0.848 -**	**0.869**
15	**0.814 -**	**0.831**
18	**0.829 -**	**0.829**
21	**0.914+**	**0.908**
24	**0.852+**	**0.823**
	Success Ratios	
1	**0.586**[*]	**0.522**
3	**0.591**[*]	**0.583**[*]
6	**0.557**	**0.570**[**]
9	**0.610**[*]	**0.610**[*]
12	**0.681**[*]	**0.723**[*]
15	**0.745**[*]	**0.770**[*]
18	**0.728**[*]	**0.746**[*]
21	**0.664**[*]	**0.655**[*]
24	**0.677**[*]	**0.695**[*]

NOTES: The models are described in the text. Boldface indicates improvements relative to the no-change forecast. + and – in column 1 indicate increases and decreases relative to column 4 of Table 2. A positive sign indicates that the model left out would have improved forecast accuracy if included, whereas a negative sign indicates that it would have worsened forecast accuracy. [*] denotes significance at the 5 per cent level and [**] at the 10 per cent level based on the Pesaran and Timmermann (2009) test for the null hypothesis of no directional accuracy. The statistical significance of the MSPE reductions cannot be assessed because none of the currently available tests of equal predictive accuracy applies in this setting.

Table 4: Real-Time Forecast Accuracy of Forecast Combinations at Quarterly Horizons after Dropping the No-Change Forecast and the Forecast Based on the Gasoline Spread

| Quarterly Horizon	Real U.S. Refiners' Acquisition Cost for Oil Imports		Real WTI Price	
	Method 1	Method 2	Method 1	Method 2
	Recursive MSPE Ratios			
1	**0.873**	**0.877**	**0.866**	**0.871**
2	**0.920**	**0.930**	**0.926**	**0.938**
3	**0.909**	**0.920**	**0.919**	**0.934**
4	**0.855**	**0.870**	**0.868**	**0.885**
5	**0.814**	**0.841**	**0.840**	**0.870**
6	**0.815**	**0.846**	**0.818**	**0.848**
7	**0.910**	**0.957**	**0.907**	**0.956**
8	**0.870**	**0.903**	**0.847**	**0.885**
	Success Ratios			
1	**0.699**[*]	**0.699**[*]	**0.711**[*]	**0.699**[*]
2	**0.659**[*]	**0.646**[*]	**0.695**[*]	**0.671**[*]
3	**0.704**[*]	**0.704**[*]	**0.691**[*]	**0.679**[*]
4	**0.738**[*]	**0.738**[*]	**0.750**[*]	**0.750**[*]
5	**0.671**[*]	**0.658**[*]	**0.658**[*]	**0.646**[*]
6	**0.705**[*]	**0.654**[*]	**0.705**[*]	**0.654**[*]
7	**0.597**	**0.584**	**0.610**[*]	**0.610**[**]
8	**0.632**[*]	**0.592**	**0.658**[*]	**0.618**[**]

NOTES: The models are described in the text. Method 1 relies on monthly forecast combinations, as in Table 3, that subsequently are aggregated by quarter. Method 2 combines the quarterly forecasts obtained by aggregating the forecasts of each monthly model by quarter. The latter method relies on inverse MSPE weights based on a rolling window of length 4 quarters. Boldface indicates improvements relative to the no-change forecast. [*] denotes significance at the 5 per cent level and [**] at the 10 per cent level based on the Pesaran and Timmermann (2009) test for the null hypothesis of no directional accuracy. The statistical significance of the MSPE reductions cannot be assessed because none of the currently available tests of equal predictive accuracy applies in this setting.

Figure 1: Real-Time Inverse MSPE Weights for Real U.S. Refiners' Acquisition Cost of Oil Imports at 1-Month Horizon
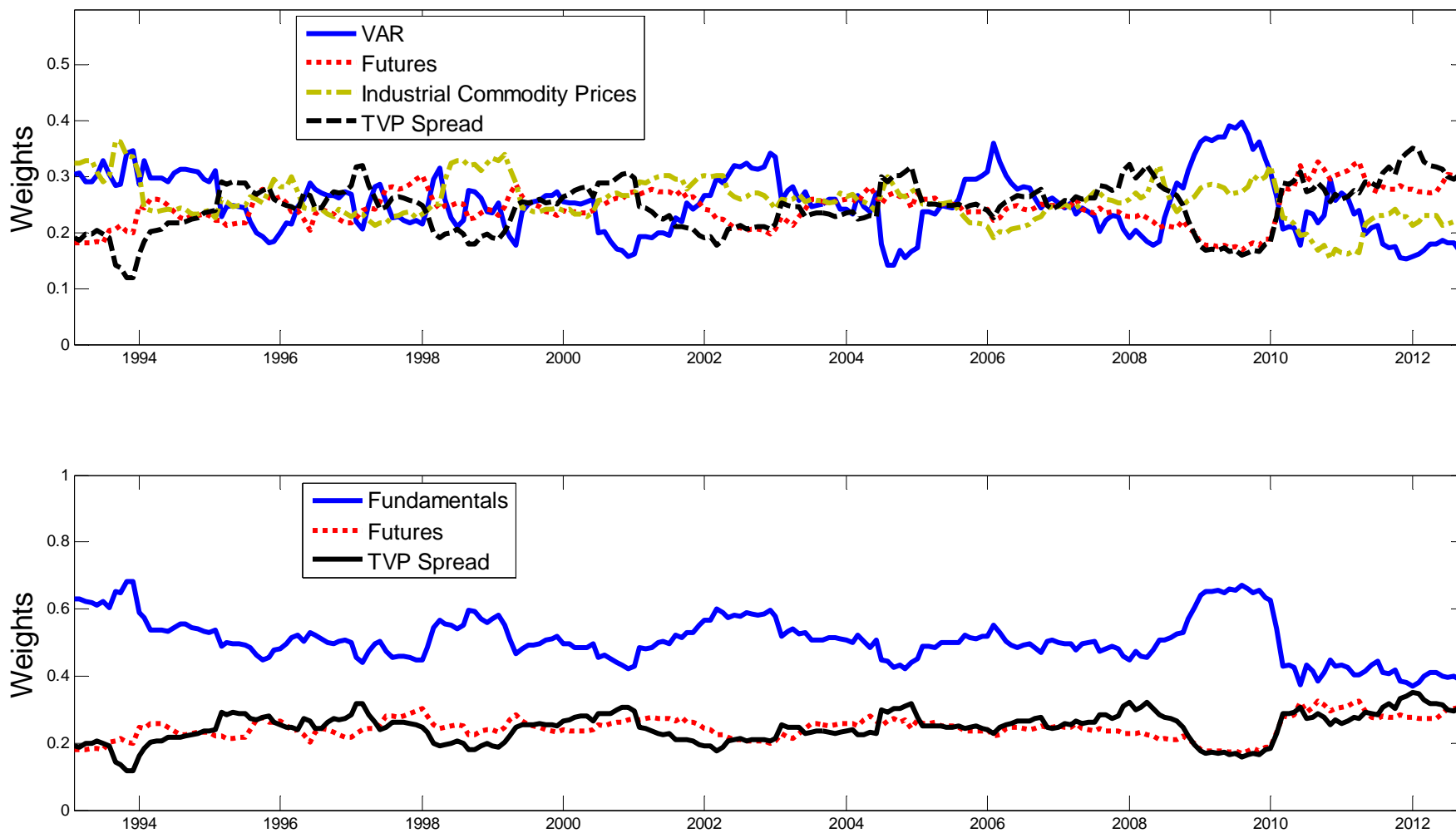
NOTES: The inverse MSPE weights correspond to the weights underlying Table 3. The weight on fundamentals is defined as the sum of the weights for the VAR model forecast and the forecast based on industrial commodity prices.

Figure 2: Real-Time Recursive MSPE Ratio of Forecast Combination Relative to No-Change Forecast of Real U.S. Refiners' Acquisition Cost of Oil Imports at 24-Month Horizon
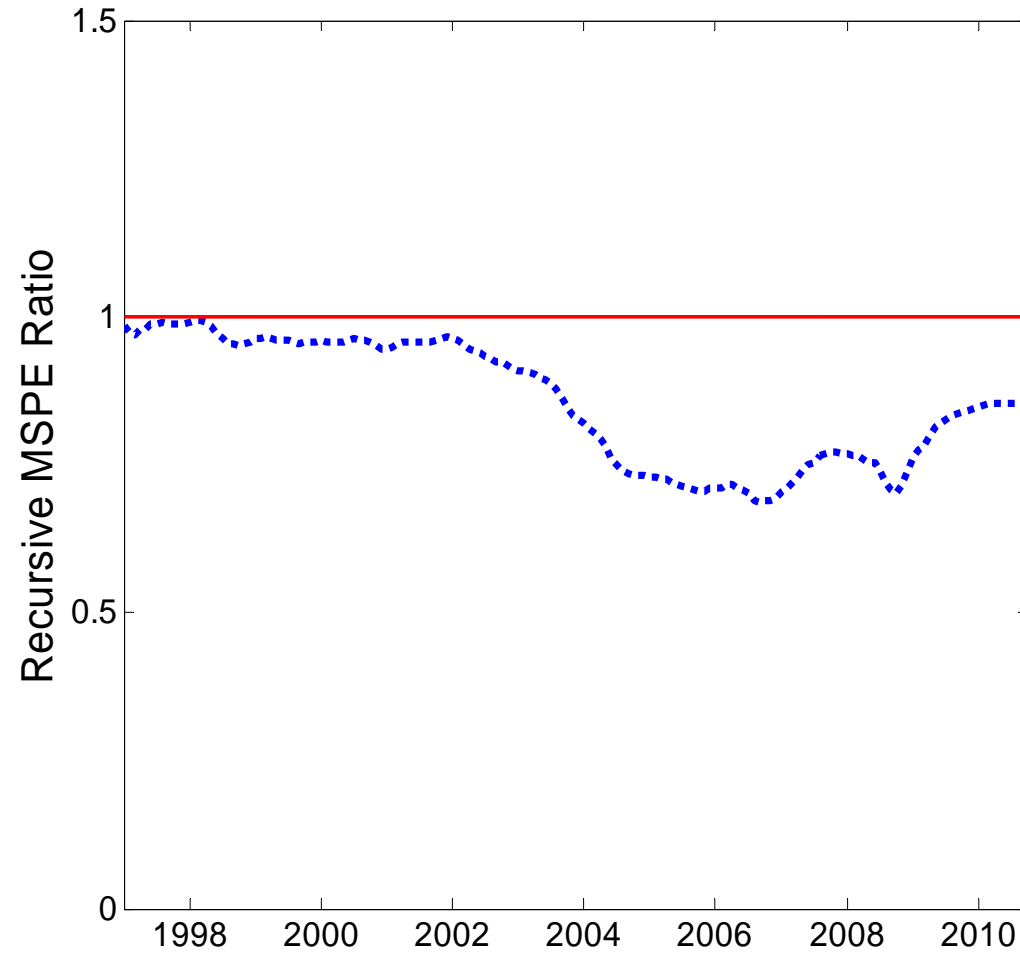
NOTES: Results based on the forecast combination shown in Table 3. A ratio below 1 indicates an improvement relative to the no-change forecast. The plot shows the evolution of the recursive MSPE ratio over time for the forecast evaluation period since 1997. This increases the reliability of the MSPE estimates and allows the MSPE ratio to stabilize.