Australian National University

Crawford School of Public Policy

# CAMA

Centre for Applied Macroeconomic Analysis

# Real-time Forecast Combinations for the Oil Price

**CAMA Working Paper 38/2018**
**August 2018**

**Anthony Garratt**
University of Warwick

**Shaun P. Vahey**
University of Warwick
Centre for Applied Macroeconomic Analysis, ANU

**Yunyi Zhang**
University of Warwick

## Abstract

Baumeister and Kilian (2015) combine forecasts from six empirical models to predict real oil prices. In this paper, we broadly reproduce their main economic findings, employing their preferred measures of the real oil price and similar real-time variables. Mindful of the importance of Brent crude oil as a global price benchmark, we extend consideration to the North Sea based measure and update the evaluation sample to 2017:12. We model the oil price futures curve using a factor-based Nelson-Siegel specification to fill in missing values of oil price futures in the source data. We find that the combined forecasts for Brent are as effective as for other oil price measures. The extended sample using the oil price measures adopted by Baumeister and Kilian (2015) yields similar results to those reported in their paper. And the futures-based model improves forecast accuracy at longer horizon forecasts. The real-time data set is available for download from shaunvahey.com.

**Keywords**

Real oil price forecasting, Brent crude oil, Forecast combination

**JEL Classification**

**Address for correspondence:**

(E) cama.admin@anu.edu.au

**ISSN 2206-0332**

# Real-time Forecast Combinations for the Oil Price*

Anthony Garratt†, Shaun P. Vahey‡, Yunyi Zhang§

July 27, 2018

**Summary:** Baumeister and Kilian (2015) combine forecasts from six empirical models to predict real oil prices. In this paper, we broadly reproduce their main economic findings, employing their preferred measures of the real oil price and similar real-time variables. Mindful of the importance of Brent crude oil as a global price benchmark, we extend consideration to the North Sea based measure and update the evaluation sample to 2017:12. We model the oil price futures curve using a factor-based Nelson-Siegel specification to fill in missing values of oil price futures in the source data. We find that the combined forecasts for Brent are as effective as for other oil price measures. The extended sample using the oil price measures adopted by Baumeister and Kilian (2015) yields similar results to those reported in their paper. And the futures-based model improves forecast accuracy at longer horizon forecasts. The real-time data set is available for download from *shaunvahey.com*.

**Keywords:** Real oil price forecasting, Brent crude oil, Forecast combination

*We thank Lutz Kilian for helpful conversations. Data, database documentation and not for publication appendices are available from `shaunvahey.com`.

†University of Warwick

‡University of Warwick and CAMA (ANU)

§University of Warwick

# 1 Introduction

Notable recent features of the three real oil prices measures illustrated in Figure 1 include: from 2011, the divergence between Brent crude, the U.S. Refiners' Acquisition Cost (RAC), and the West Texas Intermediate (WTI); the relative convergence during 2014; and the lower conditional means of all these measures post-2014.

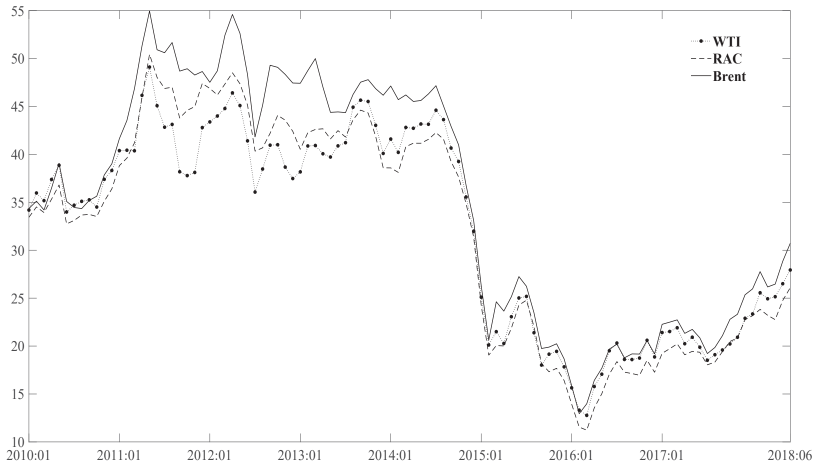


**Figure 1: Real Oil Price Measures**

Baumeister and Kilian (2015) compare the forecasting performance of six econometric models for the real oil price, individually and in combination relative to a no-change benchmark model. Their analysis is restricted to a sample ending in 2012:9 – excluding much of the more recent data plotted in Figure 1 – and neglects the Brent crude oil price. Arguably, the Brent measure represents an increasingly important benchmark for the world oil price; see discussions by (among others), Morana (2001), Alberola, Chevallier, and Chèze (2008), and Baumeister and Kilian (2016).

In this paper, we consider three extensions to their analysis. First, the robustness of the results reported by Baumeister and Kilian (2015) to utilising the real Brent measure (as well as the WTI and RAC measures). Second, the sensitivity of their findings to a longer evaluation sample, ending in 2017:12, rather than 2012:9. Third, the consideration of futures-based forecasts at longer forecast horizons. We find evidence of similar predictability across real oil price measures and over the extended evaluation sample, confirming the general findings of Baumeister and Kilian (2015), but with stronger forecasting performance at longer horizons over the extended sample. This last feature of our results arises from our use of factor-based estimation of the oil price futures curve using the specification of Nelson and Siegel (1987).

We provide a multivariate database vital for subsequent real-time research on the oil market. The database provides real-time measurements by data vintage for variables similar to those described by Baumeister and Kilian (2012), updated so that 2018:06 represents the last time series observation for all variables. We provide detailed data descriptions in the database documentation, together with the real-time data, on *shaunvahey.com*.

The remainder of this paper is structured as follows. The next section summarises the real-time oil market data set, together with the forecast combination methods of Baumeister and Kilian (2015). The subsequent section describes the results and the final section concludes.

## 2 Real-time Data and Model Space

When compiling the monthly real-time data for the oil market, we broadly followed the nowcast and backcast methods described in Baumeister and Kilian (2012). The main differences between our approach and Baumeister and Kilian's being: (1) the inclusion of Brent crude prices; (2) the extended monthly sample with the last observation of 2018:06; and, (3) the use of crude oil price futures data for longer horizons, over the period 1991:12-2018:06.

We collected real-time data for the U.S. CPI, the real world economic activity index and the following variables provided by the Energy Information Association (EIA): the RAC, world crude oil production, U.S. crude oil inventories, U.S. petroleum inventories, and OECD petroleum inventories. The EIA publications provided real-time measurements over a variable window, up to three years prior to the most recent observation.

Following the conventional terminology in the real-time macroeconomic forecasting literature, we defined a "vintage" of data as the historical time series observed by forecasters at a specific point in time (sometimes known as the "vintage date"). For example, the 2018:06 vintage includes observations only available at the end of June 2017. There are 319 vintages in total in the database, summarised in the database documentation available together with the data from *shaunvahey.com*.

Following Baumeister and Kilian (2015), we used a point forecast combination methodology to mitigate issues of model misspecification. They combined point forecasts from six specifications using equal weights and inverse mean squared predictive error (MSPE) weights. The six specifications include: an unrestricted global oil market vector autoregression (VAR), a commodity price model, an oil price futures spread model, a gasoline spread model, a time-varying parameter (TVP) product spread, and a no-change benchmark model.

Baumeister and Kilian (2015) forecast the nominal crude oil price deflated by the U.S. CPI, based on information at time $t$ for period $t+h$, $\widehat{R}^{oil}_{t+h|t}$, where $h$ is the forecast horizon. We examine point forecast combinations based on these six different specifications:

$$\widehat{R}^{oil}_{t+h|t} = \sum_{k=1}^{6} w_{k,t} \widehat{R}^{k}_{t+h|t} \tag{1}$$

where the weights, $w_{k,t}$, are assigned to model $k$ at time $t$. Equal weights, $w_{k,t} = \frac{1}{6}$ and rolling and recursively estimated mean squared predictive error (MSPE)-based weights are used, where the latter are defined as:

$$w_{k,t} = \frac{m_{k,t}^{-1}}{\sum_{j=1}^{6} m_{j,t}^{-1}}$$

where $m_{k,t}^{-1}$ is the inverse MSPE of model $k$ calculated with respect to observed outcomes available at time $t$.

With $\widehat{R}^{k}_{t+h|t}$ denoting the forecast from the $k^{th}$ specification, the six models are as described below.

1. An unrestricted global oil market vector autoregression (VAR):

$$\widehat{R}^{1}_{t+h|t} = exp(\widehat{r}^{\text{VAR}}_{t+h|t}) \tag{2}$$

   where $\widehat{r}^{VAR}_{t+h|t}$ is the forecast of the log and the VAR has four variables: the percentage change in global crude oil production, the business cycle index of global real activity (rea), the log of the RAC oil price deflated by the log of CPI, and the change in global crude oil inventories. The WTI (Brent) forecasts are constructed using the (current) spread with RAC and the RAC forecast.

2. A commodity price based model:

$$\widehat{R}^{2}_{t+h|t} = R^{oil}_{t}(1 + \pi_t^{h,\text{raw}} - E_t(\pi^h_{t+h})) \tag{3}$$

   where $\pi_t^{h,\text{raw}}$ is the difference between the log price of non-oil industrial raw materials at $t$ and $t-h$, and $R^{oil}_t$ is the real oil price measure at time $t$. Following Baumeister and Kilian (2015), $E_t(\pi^h_{t+h})$ is expected U.S. inflation, based on the historical average for CPI inflation from 1986:07.

3. A futures-based model:

$$\widehat{R}^{3}_{t+h|t} = R^{oil}_{t}(1 + f^{h}_{t} - s_t - E_t(\pi^{h}_{t+h})) \tag{4}$$

where $s_t$ is the log of monthly WTI spot price, and $f^h_t$ is the log of oil price futures for maturity $h$ observed at $t$. WTI and RAC forecasts are based on WTI futures; Brent forecasts are based on Brent futures. The monthly oil futures prices for WTI are the average of daily futures closed prices collected from Bloomberg. There are missing values in the Bloomberg source for our evaluation sample, for monthly WTI oil futures at horizons greater than 17 months and for Brent futures at horizons beyond 8 months.

Baumeister and Kilian (2015, pp.341) gave zero weight to the futures-based forecasts at long horizons due to the missing values. In order to avoid having to drop futures-based forecasts in the combinations at these horizons, given only six models are available in the first instance, we fill in missing data by estimating a factor-based model for crude oil price futures. Following Hevia et al. (2016) and Garratt and Petrella (2018), we assume that futures prices are a function of two factors, the level and slope, and impose Nelson and Siegel's (1987) parametric restrictions to the loadings. A VAR(1) is assumed for the dynamics, estimation exploits the Kalman filter, and we use the estimated model to fill in the missing values. See the database documentation for further details.

4. A gasoline spread based model:

$$\widehat{R}^{4}_{t+h|t} = R^{oil}_{t} exp\{\widehat{\beta}[s^{gas}_{t} - s_t] - E_t(\pi^{h}_{t+h})\} \tag{5}$$

where $s^{gas}_t$ is the log spot price of gasoline and $\widehat{\beta}$ is estimated from the regression $\Delta s_{t+h} = \beta[s^{gas}_t - s_t] + \varepsilon_{t+h}$, employing ordinary least squares. $\Delta s_{t+h} = s_{t+h} - s_t$ is the h-period ahead log-difference of spot WTI prices.

5. A time-varying parameter (TVP) product spread model:

$$\widehat{R}^{5}_{t+h|t} = R^{oil}_{t} exp\{\widehat{\delta_{1t}}[s^{gas}_{t} - s_t] + \widehat{\delta_{2t}}[s^{heat}_{t} - s_t] - E_t(\pi^{h}_{t+h})\}. \tag{6}$$

The parameters $\widehat{\delta_{1t}}$ and $\widehat{\delta_{2t}}$ are estimated from:

$$\Delta s_{t+h} = \delta_{1t}[s^{gas}_{t} - s_t] + \delta_{2t}[s^{heat}_{t} - s_t] + e_{t+h}$$

where $s^{heat}_t$ is the log spot price of heating oil with the error term $e_{t+h} \sim NID(0, \sigma^2)$. The TVP model Bayesian estimation of gasoline and heating oil spreads employs an independent Normal-Wishart prior and the Gibbs sampler.

6. No change forecast, from the random walk model:

$$\widehat{R}^{5}_{t+h|t} = R^{oil}_{t}. \tag{7}$$

   The no-change forecast is included in the forecast combinations, and is the used as the benchmark.

Each specification is estimated over different samples, following Baumeister and Kilian (2015), to maximise the number of observations for parameter estimation.

# 3 Results

Baumeister and Kilian (2015) evaluate the forecasts for the WTI and RAC real oil price measures from 1992:01 to 2012:09 using the 2013:03 data vintage as the target variable. Here we focus on the broader replication with the evaluation extended to 2017:12 for the monthly real Brent measure and also consider combinations which include the futures-based model. Results for the RAC and WTI measures (for the extended evaluation sample), the shorter evaluation time frame (for Brent), and the sensitivity of the inclusion of futures-based forecasts at longer horizons are presented in Appendix A (not for publication).

Table 1 reports the MSPE and success ratios of the point combination forecasts for various forecasting horizons (shown in the first column), evaluated on observations from 1992:01 to 2017:12, with the 2018:06 vintage data as the target real Brent price. The upper panel presents the end-evaluation MSPE ratios, relative to the no-change forecasts. If the MSPE ratio is below 1, the forecast is more accurate than the benchmark. The lower panel presents end-evaluation success ratios. These describe the directional accuracy, with a success ratio higher than 0.5 indicating an improvement over the benchmark. The results for point forecast combinations with equal, recursive MSPE and rolling MSPE weights (with window sizes of 36, 24, and 12 months, respectively) are reported in the columns.

Echoing the WTI and RAC results on the shorter evaluation period reported by Baumeister and Kilian (2015), we find evidence of significant predictability from forecast combinations for the Brent measure. The second column of Table 1 displays MSPE and success ratios consistent with improved accuracy (relative to the benchmark) for the equal weight combination at all forecast horizons from 1 to 24 months. The results using MSPE weights are similar to those for equal weights, regardless of whether the combinations are recursive (third column) or rolling (remaining columns).

Table 1: Forecast Accuracy for Brent, Evaluation 1992:01 to 2017:12

| MH | Real Brent price | | | | |
|---|---|---|---|---|---|
| | | | Rolling weights based on windows of length | | |
| MH | Equal weight | Recursive weights | 36 | 24 | 12 |
| | | | Recursive MSPE ratios | | |
| 1 | **0.941****(0.029) | **0.925****(0.003) | **0.927****(0.003) | **0.931****(0.005) | **0.947****(0.044) |
| 3 | **0.935****(0.006) | **0.935****(0.004) | **0.937****(0.004) | **0.940****(0.006) | **0.943****(0.007) |
| 6 | **0.975***(0.077) | **0.982**(0.122) | **0.988**(0.203) | **0.987**(0.174) | **0.984**(0.147) |
| 9 | **0.959****(0.006) | **0.966****(0.014) | **0.969****(0.020) | **0.968****(0.020) | **0.966****(0.024) |
| 12 | **0.926****(0.000) | **0.935****(0.000) | **0.939****(0.000) | **0.933****(0.000) | **0.920****(0.000) |
| 15 | **0.920****(0.000) | **0.933****(0.000) | **0.940****(0.000) | **0.944****(0.000) | **0.936****(0.000) |
| 18 | **0.933****(0.000) | **0.952****(0.001) | **0.964****(0.014) | **0.975***(0.060) | **0.979**(0.164) |
| 21 | **0.952****(0.001) | **0.975****(0.046) | **0.985**(0.165) | **0.990**(0.260) | **0.992**(0.346) |
| 24 | **0.946****(0.001) | **0.976***(0.056) | **0.975***(0.058) | **0.977***(0.063) | 1.007(0.631) |
| | | | Success ratios | | |
| 1 | **0.532****(0.040) | **0.542****(0.034) | **0.551****(0.019) | **0.542****(0.038) | **0.554****(0.011) |
| 3 | **0.548****(0.042) | **0.545****(0.049) | **0.555****(0.022) | **0.558****(0.016) | **0.577****(0.003) |
| 6 | **0.547****(0.047) | **0.524**(0.182) | **0.515**(0.309) | **0.528**(0.128) | **0.547****(0.034) |
| 9 | **0.563****(0.009) | **0.556****(0.013) | **0.536***(0.065) | **0.526**(0.107) | **0.549****(0.016) |
| 12 | **0.625****(0.000) | **0.615****(0.000) | **0.608****(0.000) | **0.611****(0.000) | **0.638****(0.000) |
| 15 | **0.634****(0.000) | **0.621****(0.000) | **0.621****(0.000) | **0.614****(0.000) | **0.631****(0.000) |
| 18 | **0.586****(0.000) | **0.583****(0.000) | **0.563****(0.001) | **0.549****(0.008) | **0.546****(0.011) |
| 21 | **0.568****(0.001) | **0.545****(0.010) | **0.538****(0.022) | **0.524***(0.070) | **0.524***(0.070) |
| 24 | **0.547****(0.030) | 0.488(0.553) | **0.543****(0.038) | **0.516**(0.221) | **0.519**(0.170) |

NOTES: MH represents monthly forecast horizons. Boldface indicates improvements relative to the no-change forecast. As a rough guide, p-values of a Harvey, Leybourne, and Newbold (1997) small-sample adjustment of the Diebold and Mariano (1995) test are reported in brackets after recursive MSPE ratios. We also report p-values for the Pesaran and Timmermann (2009) test for the null hypothesis of no directional accuracy in brackets after success ratios. ∗ denotes significance at the 10% level and ∗∗ at the 5% level.

Table 2: Forecast Accuracy for Brent, Equal Weight Combinations, Excluding and Including Futures-based Forecasts (FUTURES), Evaluation 1992:01 to 2017:12

| | **Recursive MSPE ratios** | | **Success ratios** | |
|---|---|---|---|---|
| MH | Excluding FUTURES | Including FUTURES | Excluding FUTURES | Including FUTURES |
| 9 | **0.971***(0.078) | **0.959****(0.006) | **0.549****(0.028) | **0.563****(0.009) |
| 10 | **0.963****(0.024) | **0.949****(0.000) | **0.571****(0.004) | **0.587****(0.000) |
| 11 | **0.949****(0.003) | **0.934****(0.000) | **0.599****(0.000) | **0.632****(0.000) |
| 12 | **0.945****(0.001) | **0.926****(0.000) | **0.595****(0.000) | **0.625****(0.000) |
| 13 | **0.947****(0.001) | **0.923****(0.000) | **0.550****(0.038) | **0.610****(0.000) |
| 14 | **0.947****(0.001) | **0.919****(0.000) | **0.565****(0.011) | **0.615****(0.000) |
| 15 | **0.951****(0.002) | **0.920****(0.000) | **0.584****(0.002) | **0.634****(0.000) |
| 16 | **0.955****(0.005) | **0.922****(0.000) | **0.566****(0.011) | **0.626****(0.000) |
| 17 | **0.966****(0.023) | **0.929****(0.000) | **0.561****(0.012) | **0.591****(0.000) |
| 18 | **0.973***(0.058) | **0.933****(0.000) | **0.573****(0.002) | **0.586****(0.000) |
| 19 | **0.978**(0.103) | **0.937****(0.000) | **0.558****(0.008) | **0.582****(0.000) |
| 20 | **0.990**(0.277) | **0.946****(0.000) | **0.580****(0.001) | **0.573****(0.000) |
| 21 | **1.000**(0.493) | **0.952****(0.001) | **0.538***(0.091) | **0.568****(0.001) |
| 22 | **0.998**(0.456) | **0.949****(0.001) | **0.526**(0.208) | **0.546****(0.023) |
| 23 | **0.995**(0.395) | **0.945****(0.000) | **0.517**(0.297) | **0.559****(0.008) |
| 24 | **0.998**(0.447) | **0.946****(0.001) | **0.522**(0.253) | **0.547****(0.030) |

NOTES: MH represents monthly forecast horizons. Boldface indicates improvements relative to the no-change forecast. As a rough guide, p-values of a Harvey et al. (1997) small-sample adjustment of the Diebold and Mariano (1995) test are reported in brackets after recursive MSPE ratios. We also report p-values for the Pesaran and Timmermann (2009) test for the null hypothesis of no directional accuracy in brackets after success ratios. ∗ denotes significance at the 10% level and ∗∗ at the 5% level.

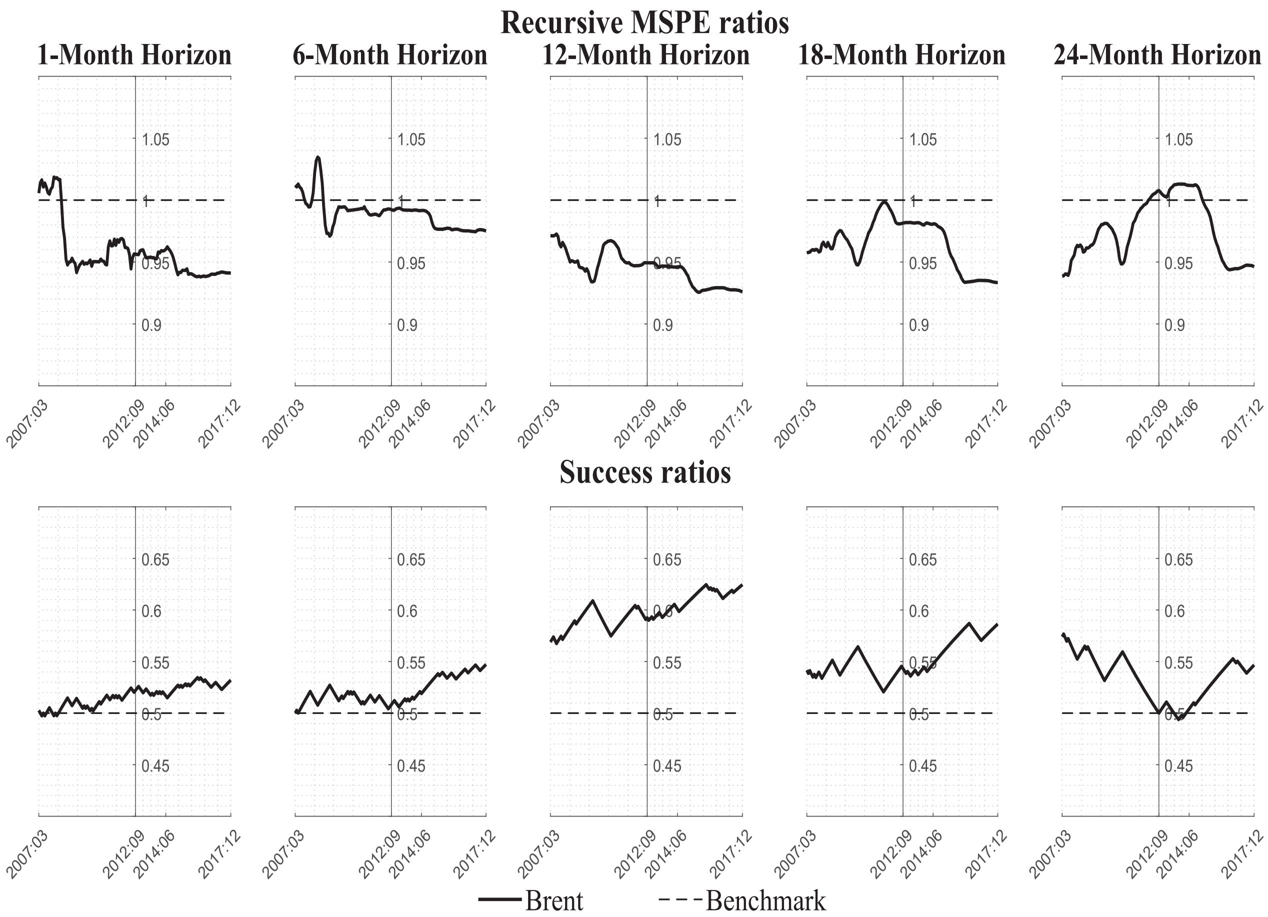


**Figure 2: Recursive Forecast Accuracy of Equal Weighted Combinations**

In Table 2 we compare the equal-weight combinations' forecast accuracy with and without futures-based forecasts for the 1992:01-2017:12 evaluation sample at horizons beyond 8 months for Brent (see Section 2 and the Appendix for further details.). Prior to this horizon there are no missing values on futures for Brent. The inclusion of futures-based forecasts reduces MSPE ratios, and raises the success ratios, with stronger statistical significance based on a Harvey et al. (1997) small-sample adjustment of the Diebold and Mariano (1995) test and the Pesaran and Timmermann (2009) test (see appendix further details).

Digging a little deeper into the real-time properties of forecast combinations, Figure 2 plots the recursive MSPE ratios (top panel) and the recursive success ratios (bottom panel) of the equal weight combinations for selected horizons (1, 6, 12, 18 and 24 months), with end-evaluation dates from 2007:03 to 2017:12. The 2012:09 end evaluation considered by Baumeister and Kilian (2015) sits in the middle of the x-axis for each cell. The equal weight combination is preferred if the line lies below 1 for the upper panel and above 0.5 for the lower. The recursive MSPE and success ratios consistently indicate that equal weight combinations dominate the benchmark before and after 2012:09 (with the exception of the 24-month horizon case for MSPE between 2012 and 2015).

## 4 Conclusions

In this paper, we have broadly replicated the findings of Baumeister and Kilian using point forecast combinations to predict the real oil price, evaluating real-time forecasts for 1992:01 to 2017:12. We found the accuracy of their point forecast combinations to be robust across different measures of the oil price and over various evaluation samples. We have also found that including futures-based information improves the longer horizon forecasts. Subsequent researchers will find the real-time data set for this study particularly helpful when investigating new candidate models and methods for both point and density forecast combinations.

## References


Alberola, E., Chevallier, J., & Chèze, B. (2008). Price drivers and structural breaks in European carbon prices 2005–2007. *Energy Policy*, *36*(2), 787–797.

Baumeister, C., & Kilian, L. (2012). Real-time forecasts of the real price of oil. *Journal of Business & Economic Statistics*, *30*(2), 326–336.

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

Baumeister, C., & Kilian, L. (2016). Understanding the decline in the price of oil since June 2014. *Journal of the Association of Environmental and Resource Economists*, *3*(1), 131–158.

Diebold, F. X., & Mariano, R. S. (1995). Comparing predictive accuracy. *Journal of Business & Economic Statistics*, *13*(3).

Garratt, A., & Petrella, I. (2018). The informational content of commodity prices for probabilistic inflation forecasts. *Unpublished manuscript*.

Harvey, D., Leybourne, S., & Newbold, P. (1997). Testing the equality of prediction mean squared errors. *International Journal of forecasting*, *13*(2), 281–291.

Hevia, C., Petrella, I., & Sola, M. (2016). Risk premia and seasonality in commodity futures. *Journal of Applied Econometrics*.

Morana, C. (2001). A semiparametric approach to short-term oil price forecasting. *Energy Economics*, *23*(3), 325–338.

Nelson, C. R., & Siegel, A. F. (1987). Parsimonious modeling of yield curves. *Journal of business*, 473–489.

Pesaran, M. H., & Timmermann, A. (2009). Testing dependence among serially correlated multicategory variables. *Journal of the American Statistical Association*, *104*(485), 325–337.

# Appendix to Garratt, Vahey and Zhang (2018)

## (I) Shorter evaluation sample, 1992:01–2012:09

Our results from the narrow replication, using the evaluation sample examined by Baumeister and Kilian (2015) are shown in Tables A-1a and A-1b. These results use the same WTI and RAC measures considered by those authors. The results confirm the main findings of their paper. Equal weight point combinations have lower MSPE ratios and higher success ratios than inverse MSPE weights for most horizons. The corresponding recursive MSPE and success ratios for the Brent measure, with the same 1992:01 to 2012:09 evaluation sample are displayed in Table A-1c.

Table A-1a. Forecast Accuracy for RAC, Evaluation 1992:01-2012:09

| | Real U.S. refiners' acquisition cost for oil imports | | | | |
|---|---|---|---|---|---|
| | | | Rolling weights based on windows of length | | |
| MH | Equal weight | Recursive weights | 36 | 24 | 12 |
| | | | Recursive MSPE ratios | | |
| 1 | **0.931****(0.032) | **0.936****(0.044) | **0.938****(0.050) | **0.930****(0.036) | **0.928****(0.033) |
| 3 | **0.922****(0.009) | **0.925****(0.010) | **0.925****(0.011) | **0.920****(0.005) | **0.927****(0.005) |
| 6 | **0.983**(0.139) | **0.987**(0.208) | **0.989**(0.229) | **0.990**(0.256) | **0.989**(0.227) |
| 9 | **0.977***(0.082) | **0.982**(0.145) | **0.981**(0.131) | **0.985**(0.185) | **0.986**(0.233) |
| 12 | **0.938****(0.000) | **0.944****(0.001) | **0.948****(0.003) | **0.946****(0.002) | **0.943****(0.001) |
| 15 | **0.928****(0.000) | **0.939****(0.001) | **0.952****(0.011) | **0.952****(0.009) | **0.969***(0.078) |
| 18 | **0.969****(0.036) | **0.990**(0.277) | 1.018(0.809) | 1.026(0.883) | 1.053(0.984) |
| 21 | 1.002(0.543) | 1.023(0.921) | 1.049(0.992) | 1.058(0.995) | 1.101(1.000) |
| 24 | **0.981**(0.133) | **0.991**(0.296) | **0.994**(0.373) | **0.999**(0.488) | 1.039(0.939) |
| | | | Success ratios | | |
| 1 | **0.558***(0.064) | **0.550***(0.093) | **0.550***(0.093) | **0.546**(0.122) | **0.554***(0.084) |
| 3 | **0.587****(0.013) | **0.583****(0.018) | **0.583****(0.018) | **0.587****(0.014) | **0.579****(0.028) |
| 6 | **0.553**(0.135) | **0.533**(0.324) | **0.537**(0.271) | **0.516**(0.531) | **0.520**(0.419) |
| 9 | **0.548**(0.112) | **0.560****(0.046) | **0.564****(0.041) | **0.552***(0.085) | **0.552***(0.077) |
| 12 | **0.660****(0.000) | **0.647****(0.000) | **0.639****(0.000) | **0.639****(0.000) | **0.664****(0.000) |
| 15 | **0.626****(0.000) | **0.591****(0.001) | **0.574****(0.012) | **0.570****(0.016) | **0.574****(0.015) |
| 18 | **0.565****(0.003) | **0.543****(0.012) | **0.513**(0.177) | 0.496(0.381) | 0.496(0.326) |
| 21 | **0.581****(0.001) | **0.502**(0.124) | **0.550****(0.027) | **0.537**(0.114) | **0.502**(0.498) |
| 24 | **0.558****(0.033) | **0.522***(0.088) | **0.549***(0.079) | **0.540**(0.182) | **0.531**(0.264) |

NOTES: MH represents monthly forecast horizons. Boldface indicates improvements relative to the no-change forecast. As a rough guide, p-values of a Harvey et al. (1997) small-sample adjustment of the Diebold and Mariano (1995) test are reported in brackets after recursive MSPE ratios. We also report p-values for the Pesaran and Timmermann (2009) test for the null hypothesis of no directional accuracy in brackets after success ratios. $*$ denotes significance at the 10% level and $**$ at the 5% level.

## Table A-1b. Forecast Accuracy for WTI, Evaluation 1992:01-2012:09

| | Real WTI price | | | | |
|---|---|---|---|---|---|
| | | | Rolling weights based on windows of length | | |
| MH | Equal weight | Recursive weights | 36 | 24 | 12 |
| | | Recursive MSPE ratios | | | |
| 1 | **0.907**$^{**}$(0.008) | **0.911**$^{**}$(0.009) | **0.912**$^{**}$(0.011) | **0.912**$^{**}$(0.012) | **0.909**$^{**}$(0.013) |
| 3 | **0.923**$^{**}$(0.011) | **0.927**$^{**}$(0.012) | **0.929**$^{**}$(0.014) | **0.928**$^{**}$(0.011) | **0.934**$^{**}$(0.013) |
| 6 | **0.986**(0.214) | **0.991**(0.289) | **0.992**(0.309) | **0.994**(0.350) | **0.997**(0.419) |
| 9 | **0.980**(0.127) | **0.986**(0.208) | **0.982**(0.158) | **0.986**(0.217) | **0.989**(0.281) |
| 12 | **0.946**$^{**}$(0.002) | **0.954**$^{**}$(0.007) | **0.953**$^{**}$(0.008) | **0.949**$^{**}$(0.005) | **0.945**$^{**}$(0.003) |
| 15 | **0.939**$^{**}$(0.001) | **0.952**$^{**}$(0.007) | **0.962**$^{**}$(0.042) | **0.963**$^{**}$(0.041) | **0.943**$^{**}$(0.005) |
| 18 | **0.968**$^{**}$(0.035) | **0.993**(0.351) | 1.020(0.843) | 1.038(0.954) | 1.065(0.988) |
| 21 | 1.000(0.496) | 1.029(0.948) | 1.051(0.991) | 1.062(0.994) | 1.081(0.998) |
| 24 | **0.973**$^{*}$(0.080) | **0.994**(0.373) | **0.995**(0.403) | 1.003(0.554) | 1.048(0.958) |
| | | Success ratios | | | |
| 1 | **0.550**(0.135) | **0.554**(0.124) | **0.550**(0.150) | **0.558**(0.102) | **0.558**(0.102) |
| 3 | **0.555**(0.124) | **0.551**(0.150) | **0.543**(0.210) | **0.547**(0.198) | **0.538**(0.306) |
| 6 | **0.529**(0.366) | **0.520**(0.487) | **0.508**(0.635) | **0.508**(0.584) | 0.484(0.843) |
| 9 | **0.544**(0.130) | **0.560**$^{**}$(0.049) | **0.568**$^{**}$(0.031) | **0.556**$^{*}$(0.073) | **0.535**(0.181) |
| 12 | **0.588**$^{**}$(0.004) | **0.571**$^{**}$(0.014) | **0.584**$^{**}$(0.009) | **0.601**$^{**}$(0.001) | **0.601**$^{**}$(0.001) |
| 15 | **0.583**$^{**}$(0.005) | **0.583**$^{**}$(0.003) | **0.566**$^{**}$(0.028) | **0.566**$^{**}$(0.031) | **0.562**$^{**}$(0.024) |
| 18 | **0.573**$^{**}$(0.003) | **0.556**$^{**}$(0.004) | **0.534**$^{*}$(0.060) | **0.513**(0.221) | **0.509**(0.231) |
| 21 | **0.572**$^{**}$(0.002) | **0.502**(0.106) | **0.537**$^{*}$(0.050) | **0.515**(0.261) | **0.511**(0.324) |
| 24 | **0.549**$^{*}$(0.065) | 0.496(0.290) | **0.531**(0.168) | **0.527**(0.325) | **0.527**(0.301) |

NOTES: MH represents monthly forecast horizons. Boldface indicates improvements relative to the no-change forecast. As a rough guide, p-values of a Harvey et al. (1997) small-sample adjustment of the Diebold and Mariano (1995) test are reported in brackets after recursive MSPE ratios. We also report p-values for the Pesaran and Timmermann (2009) test for the null hypothesis of no directional accuracy in brackets after success ratios. $*$ denotes significance at the 10% level and $**$ at the 5% level.

## Table A-1c. Forecast Accuracy for Brent, Evaluation 1992:01-2012:09

| | Real Brent price | | | | |
|---|---|---|---|---|---|
| | | | Rolling weights based on windows of length | | |
| MH | Equal weight | Recursive weights | 36 | 24 | 12 |
| | | Recursive MSPE ratios | | | |
| 1 | **0.956**(0.107) | **0.935**$^{**}$(0.018) | **0.933**$^{**}$(0.017) | **0.936**$^{**}$(0.022) | **0.948**$^{*}$(0.084) |
| 3 | **0.942**$^{**}$(0.027) | **0.941**$^{**}$(0.021) | **0.939**$^{**}$(0.016) | **0.943**$^{**}$(0.023) | **0.947**$^{**}$(0.026) |
| 6 | **0.992**(0.350) | **0.999**(0.472) | 1.004(0.590) | 1.004(0.587) | 1.007(0.658) |
| 9 | **0.981**(0.152) | **0.988**(0.244) | **0.989**(0.256) | **0.995**(0.373) | **0.995**(0.392) |
| 12 | **0.949**$^{**}$(0.002) | **0.959**$^{**}$(0.009) | **0.963**$^{**}$(0.014) | **0.965**$^{**}$(0.014) | **0.957**$^{**}$(0.009) |
| 15 | **0.943**$^{**}$(0.000) | **0.957**$^{**}$(0.005) | **0.969**$^{**}$(0.036) | **0.976**$^{*}$(0.071) | **0.979**$^{*}$(0.093) |
| 18 | **0.981**(0.119) | 1.005(0.631) | 1.026(0.944) | 1.041(0.993) | 1.075(1.000) |
| 21 | 1.014(0.842) | 1.040(0.996) | 1.058(1.000) | 1.064(1.000) | 1.098(1.000) |
| 24 | 1.008(0.688) | 1.025(0.932) | 1.034(0.971) | 1.041(0.993) | 1.096(1.000) |
| | | Success ratios | | | |
| 1 | **0.522**(0.128) | **0.526**(0.133) | **0.534**$^{*}$(0.094) | **0.534**$^{*}$(0.089) | **0.558**$^{**}$(0.016) |
| 3 | **0.522**(0.263) | **0.518**(0.284) | **0.530**(0.174) | **0.530**(0.163) | **0.538**(0.125) |
| 6 | **0.508**(0.406) | 0.484(0.673) | 0.471(0.832) | 0.471(0.764) | 0.488(0.597) |
| 9 | **0.506**(0.355) | 0.498(0.414) | 0.485(0.540) | 0.465(0.753) | 0.490(0.455) |
| 12 | **0.592**$^{**}$(0.000) | **0.592**$^{**}$(0.001) | **0.584**$^{**}$(0.001) | **0.584**$^{**}$(0.001) | **0.597**$^{**}$(0.000) |
| 15 | **0.604**$^{**}$(0.000) | **0.591**$^{**}$(0.000) | **0.587**$^{**}$(0.000) | **0.587**$^{**}$(0.000) | **0.609**$^{**}$(0.000) |
| 18 | **0.543**$^{**}$(0.001) | **0.539**$^{**}$(0.001) | **0.504**$^{**}$(0.035) | 0.483(0.195) | 0.474(0.284) |
| 21 | **0.541**$^{**}$(0.001) | **0.515**$^{**}$(0.002) | 0.493$^{**}$(0.050) | 0.463(0.427) | 0.450(0.582) |
| 24 | 0.500(0.136) | 0.456(0.461) | **0.504**$^{*}$(0.087) | 0.451(0.782) | 0.456(0.749) |

NOTES: MH represents monthly forecast horizons. Boldface indicates improvements relative to the no-change forecast. As a rough guide, p-values of a Harvey et al. (1997) small-sample adjustment of the Diebold and Mariano (1995) test are reported in brackets after recursive MSPE ratios. We also report p-values for the Pesaran and Timmermann (2009) test for the null hypothesis of no directional accuracy in brackets after success ratios. $*$ denotes significance at the 10% level and $**$ at the 5% level.

## (II) Longer evaluation sample, 1992:01–2017:12, for RAC and WTI measures

We also present the forecast accuracy of RAC and WTI for the extended 1992:01 to 2017:12 evaluation sample in Tables A-2a and A-2b, respectively.

Table A-2a. Forecast Accuracy for RAC, Evaluation 1992:01 to 2017:12

| | Real U.S. refiners' acquisition cost for oil imports | | | | |
|---|---|---|---|---|---|
| | | | Rolling weights based on windows of length | | |
| MH | Equal weight | Recursive weights | 36 | 24 | 12 |
| | | | Recursive MSPE ratios | | |
| 1 | **0.933****(0.017) | **0.937****(0.024) | **0.941****(0.032) | **0.934****(0.022) | **0.930****(0.018) |
| 3 | **0.922****(0.002) | **0.925****(0.003) | **0.930****(0.004) | **0.922****(0.001) | **0.922****(0.001) |
| 6 | **0.975****(0.035) | **0.981***(0.075) | **0.982***(0.080) | **0.978****(0.045) | **0.969****(0.019) |
| 9 | **0.969****(0.011) | **0.977****(0.043) | **0.973****(0.023) | **0.971****(0.023) | **0.968****(0.023) |
| 12 | **0.934****(0.000) | **0.943****(0.000) | **0.940****(0.000) | **0.929****(0.000) | **0.921****(0.000) |
| 15 | **0.929****(0.000) | **0.941****(0.000) | **0.941****(0.000) | **0.933****(0.000) | **0.925****(0.000) |
| 18 | **0.941****(0.000) | **0.959****(0.004) | **0.971****(0.047) | **0.968***(0.051) | **0.949****(0.035) |
| 21 | **0.969****(0.014) | **0.994**(0.324) | **0.994**(0.363) | **0.998**(0.449) | **0.995**(0.425) |
| 24 | **0.966****(0.014) | 1.007(0.692) | **0.985**(0.163) | **0.986**(0.213) | 1.001(0.519) |
| | | | Success ratios | | |
| 1 | **0.558****(0.042) | **0.554****(0.047) | **0.567****(0.017) | **0.561****(0.031) | **0.574****(0.010) |
| 3 | **0.603****(0.001) | **0.600****(0.001) | **0.603****(0.001) | **0.610****(0.000) | **0.594****(0.002) |
| 6 | **0.580****(0.011) | **0.564****(0.047) | **0.564****(0.040) | **0.550***(0.098) | **0.560****(0.044) |
| 9 | **0.559****(0.032) | **0.566****(0.019) | **0.582****(0.004) | **0.576****(0.007) | **0.569****(0.013) |
| 12 | **0.661****(0.000) | **0.645****(0.000) | **0.645****(0.000) | **0.651****(0.000) | **0.681****(0.000) |
| 15 | **0.621****(0.000) | **0.594****(0.001) | **0.587****(0.002) | **0.594****(0.001) | **0.584****(0.002) |
| 18 | **0.576****(0.002) | **0.556****(0.017) | **0.529**(0.142) | **0.512**(0.326) | **0.512**(0.293) |
| 21 | **0.579****(0.003) | **0.507**(0.340) | **0.548***(0.055) | **0.551****(0.042) | **0.514**(0.350) |
| 24 | **0.564****(0.019) | **0.512**(0.341) | **0.550***(0.063) | **0.540**(0.123) | **0.529**(0.181) |

NOTES: MH represents monthly forecast horizons. Boldface indicates improvements relative to the no-change forecast. As a rough guide, p-values of a Harvey et al. (1997) small-sample adjustment of the Diebold and Mariano (1995) test are reported in brackets after recursive MSPE ratios. We also report p-values for the Pesaran and Timmermann (2009) test for the null hypothesis of no directional accuracy in brackets after success ratios. $*$ denotes significance at the 10% level and $**$ at the 5% level.

Table A-2b. Forecast Accuracy for WTI, Evaluation 1992:01 to 2017:12

| MH | Real WTI price | | | | |
|---|---|---|---|---|---|
| | | | Rolling weights based on windows of length | | |
| | Equal weight | Recursive weights | 36 | 24 | 12 |
| | | | Recursive MSPE ratios | | |
| 1 | **0.910****(0.002) | **0.912****(0.002) | **0.915****(0.003) | **0.915****(0.004) | **0.912****(0.004) |
| 3 | **0.922****(0.003) | **0.925****(0.003) | **0.930****(0.004) | **0.928****(0.003) | **0.929****(0.002) |
| 6 | **0.977***(0.064) | **0.982**(0.115) | **0.983**(0.114) | **0.982**(0.104) | **0.987**(0.177) |
| 9 | **0.972****(0.023) | **0.979***(0.069) | **0.974****(0.035) | **0.973****(0.039) | **0.978***(0.090) |
| 12 | **0.940****(0.000) | **0.948****(0.000) | **0.942****(0.000) | **0.930****(0.000) | **0.936****(0.000) |
| 15 | **0.934****(0.000) | **0.945****(0.000) | **0.942****(0.000) | **0.932****(0.000) | **0.914****(0.000) |
| 18 | **0.940****(0.000) | **0.959****(0.005) | **0.966****(0.033) | **0.968***(0.060) | **0.978**(0.197) |
| 21 | **0.967****(0.013) | **0.990**(0.240) | **0.987**(0.237) | **0.988**(0.274) | **0.985**(0.276) |
| 24 | **0.952****(0.003) | **0.989**(0.239) | **0.962****(0.018) | **0.960****(0.021) | **0.986**(0.331) |
| | | | Success ratios | | |
| 1 | **0.561****(0.045) | **0.564****(0.040) | **0.554***(0.083) | **0.561***(0.052) | **0.571****(0.024) |
| 3 | **0.571****(0.021) | **0.565****(0.035) | **0.568****(0.026) | **0.571****(0.022) | **0.568****(0.033) |
| 6 | **0.554***(0.087) | **0.544**(0.176) | **0.537**(0.236) | **0.541**(0.162) | **0.511**(0.548) |
| 9 | **0.566****(0.020) | **0.576****(0.009) | **0.579****(0.007) | **0.569****(0.018) | **0.559****(0.032) |
| 12 | **0.621****(0.000) | **0.608****(0.000) | **0.611****(0.000) | **0.631****(0.000) | **0.608****(0.000) |
| 15 | **0.607****(0.000) | **0.604****(0.000) | **0.597****(0.001) | **0.604****(0.000) | **0.594****(0.001) |
| 18 | **0.590****(0.000) | **0.583****(0.001) | **0.580****(0.002) | **0.556****(0.026) | **0.553****(0.028) |
| 21 | **0.592****(0.000) | **0.527**(0.103) | **0.558****(0.020) | **0.548***(0.054) | **0.555****(0.036) |
| 24 | **0.561****(0.025) | **0.502**(0.457) | **0.547***(0.069) | **0.536**(0.136) | **0.536**(0.118) |

NOTES: MH represents monthly forecast horizons. Boldface indicates improvements relative to the no-change forecast. As a rough guide, p-values of a Harvey et al. (1997) small-sample adjustment of the Diebold and Mariano (1995) test are reported in brackets after recursive MSPE ratios. We also report p-values for the Pesaran and Timmermann (2009) test for the null hypothesis of no directional accuracy in brackets after success ratios. $*$ denotes significance at the 10% level and $**$ at the 5% level.

## (III) The inclusion of futures-based forecasts

Analysing the effect of including the futures-based forecasts, in Table A-3a and A-3b we compare the the forecast accuracy of equal weight combinations with and without futures-based forecasts for the 1992:01-2012:09 and 1992:01-2017:12 sample periods at horizons 18 to 24 months for RAC and WTI. As with Brent in the main text, the inclusion of futures-based forecasts at these horizons reduces MSPE ratios and raises the success ratios. Table A-4 additionally presents the effect of including the futures-based forecasts for the Brent measure in the 1992:01-2012:09 evaluation sample.

Table A-3a: Forecast Accuracy for RAC, Equal Weight Combinations, Excluding and Including Futures-based Forecasts (FUTURES)

| | Real RAC price | | | |
|---|---|---|---|---|
| | 1992:01-2012:09 | | 1992:01-2017:12 | |
| MH | Excluding FUTURES | Including FUTURES | Excluding FUTURES | Including FUTURES |
| | | Recursive MSPE ratios | | |
| 18 | 1.007(0.615) | **0.969****(0.036) | **0.996**(0.385) | **0.941****(0.000) |
| 19 | 1.023(0.845) | **0.985**(0.180) | 1.005(0.639) | **0.948****(0.000) |
| 20 | 1.038(0.953) | **0.999**(0.482) | 1.022(0.932) | **0.961****(0.004) |
| 21 | 1.039(0.953) | 1.002(0.543) | 1.033(0.986) | **0.969****(0.014) |
| 22 | 1.030(0.892) | **0.996**(0.401) | 1.033(0.983) | **0.968****(0.013) |
| 23 | 1.015(0.733) | **0.986**(0.196) | 1.030(0.966) | **0.965****(0.009) |
| 24 | 1.007(0.605) | **0.981**(0.133) | 1.031(0.964) | **0.966****(0.014) |
| | | Success ratios | | |
| 18 | **0.534**(0.456) | **0.565****(0.003) | **0.522**(0.358) | **0.576****(0.002) |
| 19 | **0.537**(0.366) | **0.589****(0.000) | **0.551***(0.065) | **0.602****(0.000) |
| 20 | **0.561**(0.132) | **0.613****(0.000) | **0.553***(0.082) | **0.618****(0.000) |
| 21 | **0.507**(0.875) | **0.581****(0.001) | 0.486(0.899) | **0.579****(0.003) |
| 22 | **0.518**(0.857) | **0.561****(0.010) | 0.488(0.881) | **0.557****(0.026) |
| 23 | **0.542**(0.650) | **0.555****(0.026) | **0.503**(0.762) | **0.552****(0.044) |
| 24 | **0.535**(0.701) | **0.558****(0.033) | **0.509**(0.715) | **0.564****(0.019) |

NOTES: MH represents monthly forecast horizons. Boldface indicates improvements relative to the no-change forecast. As a rough guide, p-values of a Harvey et al. (1997) small-sample adjustment of the Diebold and Mariano (1995) test are reported in brackets after recursive MSPE ratios. We also report p-values for the Pesaran and Timmermann (2009) test for the null hypothesis of no directional accuracy in brackets after success ratios. $*$ denotes significance at the 10% level and $**$ at the 5% level.

Table A-3b: Forecast Accuracy for WTI, Equal Weight Combinations, Excluding and Including Futures-based Forecasts (FUTURES)

| | Real WTI price | | | |
|---|---|---|---|---|
| | 1992:01-2012:09 | | 1992:01-2017:12 | |
| MH | Excluding FUTURES | Including FUTURES | Excluding FUTURES | Including FUTURES |
| | | Recursive MSPE ratios | | |
| 18 | 1.005(0.583) | **0.968****(0.035) | **0.994**(0.357) | **0.940****(0.000) |
| 19 | 1.016(0.757) | **0.980**(0.129) | 1.004(0.596) | **0.949****(0.001) |
| 20 | 1.029(0.890) | **0.994**(0.351) | 1.017(0.861) | **0.959****(0.004) |
| 21 | 1.034(0.920) | 1.000(0.496) | 1.028(0.958) | **0.967****(0.013) |
| 22 | 1.021(0.793) | **0.990**(0.293) | 1.024(0.921) | **0.963****(0.007) |
| 23 | 1.005(0.571) | **0.979**(0.131) | 1.017(0.829) | **0.956****(0.004) |
| 24 | **0.994**(0.420) | **0.973***(0.080) | 1.012(0.731) | **0.952****(0.003) |
| | | Success ratios | | |
| 18 | **0.517**(0.624) | **0.573****(0.003) | **0.519**(0.456) | **0.590****(0.000) |
| 19 | **0.545**(0.201) | **0.593****(0.000) | **0.551***(0.081) | **0.609****(0.000) |
| 20 | **0.539**(0.289) | **0.587****(0.000) | **0.543**(0.162) | **0.590****(0.000) |
| 21 | 0.493(0.938) | **0.572****(0.002) | **0.503**(0.818) | **0.592****(0.000) |
| 22 | **0.522**(0.758) | **0.561****(0.012) | **0.505**(0.788) | **0.577****(0.004) |
| 23 | **0.511**(0.876) | **0.568****(0.011) | **0.507**(0.742) | **0.586****(0.002) |
| 24 | **0.527**(0.780) | **0.549***(0.065) | **0.505**(0.735) | **0.561****(0.025) |

NOTES: MH represents monthly forecast horizons. Boldface indicates improvements relative to the no-change forecast. As a rough guide, p-values of a Harvey et al. (1997) small-sample adjustment of the Diebold and Mariano (1995) test are reported in brackets after recursive MSPE ratios. We also report p-values for the Pesaran and Timmermann (2009) test for the null hypothesis of no directional accuracy in brackets after success ratios. $*$ denotes significance at the 10% level and $**$ at the 5% level.

Table A-4: Forecast Accuracy for Brent, Equal Weight Combinations, Excluding and Including Futures-based Forecasts (FUTURES), Evaluation 1992:01 to 2012:09

| | **Recursive MSPE ratios** | | **Success ratios** | |
|---|---|---|---|---|
| MH | Excluding FUTURES | Including FUTURES | Excluding FUTURES | Including FUTURES |
| 9 | **0.995**(0.420) | **0.981**(0.152) | 0.494(0.584) | **0.506**(0.355) |
| 10 | **0.989**(0.320) | **0.973***(0.057) | 0.500(0.533) | **0.533***(0.069) |
| 11 | **0.976**(0.142) | **0.959****(0.007) | **0.523**(0.320) | **0.582****(0.001) |
| 12 | **0.970***(0.092) | **0.949****(0.002) | **0.534**(0.214) | **0.592****(0.000) |
| 13 | **0.970***(0.087) | **0.946****(0.001) | **0.511**(0.446) | **0.578****(0.004) |
| 14 | **0.968***(0.080) | **0.942****(0.000) | **0.525**(0.262) | **0.576****(0.003) |
| 15 | **0.971***(0.099) | **0.943****(0.000) | **0.528**(0.230) | **0.604****(0.000) |
| 16 | **0.984**(0.232) | **0.953****(0.002) | **0.534**(0.155) | **0.577****(0.001) |
| 17 | 1.002(0.530) | **0.968****(0.021) | **0.511**(0.306) | **0.545****(0.005) |
| 18 | 1.016(0.766) | **0.981**(0.119) | **0.539****(0.038) | **0.543****(0.001) |
| 19 | 1.030(0.911) | **0.996**(0.395) | **0.524****(0.046) | **0.532****(0.001) |
| 20 | 1.042(0.973) | 1.008(0.695) | **0.548****(0.009) | **0.539****(0.000) |
| 21 | 1.049(0.986) | 1.014(0.842) | **0.511**(0.324) | **0.541****(0.001) |
| 22 | 1.046(0.975) | 1.014(0.820) | **0.509**(0.438) | **0.522****(0.016) |
| 23 | 1.036(0.933) | 1.007(0.681) | 0.493(0.604) | **0.515****(0.043) |
| 24 | 1.034(0.912) | 1.008(0.688) | 0.496(0.660) | 0.500(0.136) |

NOTES: MH represents monthly forecast horizons. Boldface indicates improvements relative to the no-change forecast. As a rough guide, p-values of a Harvey et al. (1997) small-sample adjustment of the Diebold and Mariano (1995) test are reported in brackets after recursive MSPE ratios. We also report p-values for the Pesaran and Timmermann (2009) test for the null hypothesis of no directional accuracy in brackets after success ratios. $*$ denotes significance at the 10% level and $**$ at the 5% level.