Energy Economics 46 (2014) S33–S43

Contents lists available at ScienceDirect

Energy Economics

journal homepage: www.elsevier.com/locate/eneco

Are there gains from pooling real-time oil price forecasts?

Christiane Baumeister [a], Lutz Kilian [b,*], Thomas K. Lee [c]

[a] International Economic Analysis Department, Bank of Canada, 234 Laurier Avenue West, Ottawa, ON K1A 0G9, Canada
[b] Department of Economics, University of Michigan, 611 Tappan Street, Ann Arbor, MI 48109-1220, USA
[c] U.S. Energy Information Administration, 1000 Independence Avenue, Washington, DC 20585, USA

ARTICLE INFO

Article history:
Received 11 December 2013
Received in revised form 28 July 2014
Accepted 6 August 2014
Available online 20 August 2014

JEL classifications:
Q43
C53

Keywords:
Forecast combination
Real-time data
WTI
Refiners' acquisition cost
Model selection

ABSTRACT

The answer depends on the objective. The approach of combining five of the leading forecasting models with equal weights dominates the strategy of selecting one model and using it for all horizons up to two years. Even more accurate forecasts, however, are obtained when allowing the forecast combinations to vary across forecast horizons. While the latter approach is not always more accurate than selecting the single most accurate forecasting model by horizon, its accuracy can be shown to be much more stable over time. The MSPE of real-time pooled forecasts is between 3% and 29% lower than that of the no-change forecast and its directional accuracy as high as 73%. Our results are robust to alternative oil price measures and apply to monthly as well as quarterly forecasts. We illustrate how forecast pooling may be used to produce real-time forecasts of the real and the nominal price of oil in a format consistent with that employed by the U.S. Energy Information Administration in releasing its short-term oil price forecasts, and we compare these forecasts during key historical episodes.

1. Introduction

Accurate real-time forecasts of the price of oil are important to firms and consumers as well as state and national governments. Real-time forecasts refer to forecasts based on the data actually available to forecasters at the time a forecast is generated as opposed to information that only becomes available later. Real-time data sets explicitly account for delays and revisions in data releases. For example, data on global oil production are only released with a delay of several months and subject to revisions for several years. Ignoring these data constraints, as has been common in many earlier studies, may result in overly optimistic assessments of the ability to forecast oil prices (see Alquist et al., 2013).

There are many alternative real-time approaches to forecasting oil prices ranging from the use of oil futures prices and survey forecasts to atheoretical time series models and econometric models.[1] Our approach in this paper is to focus on short-term oil price forecasting models that can be motivated based on economic grounds. To date, a large number of alternative forecasting model specifications have been considered in the literature on real-time forecasts of the real price of oil. Among these models we restrict attention to forecasting models that have been shown to produce more accurate real-time forecasts than the random walk benchmark model at least for some forecast horizons. We take the specification of the forecasting models employed in this literature as given. Our objective is to examine the forecast accuracy of weighted averages of these forecasts, as measured by the mean-squared prediction error (MSPE) at monthly and quarterly horizons up to two years. We also report results for the directional accuracy of these combined forecasts.

Forecast combinations (also known as pooled forecasts) have a long tradition in macroeconomic forecasting (see, e.g., Timmermann, 2006). With regard to short-term oil price forecasts, Baumeister and Kilian (forthcoming) established that an equal-weighted combination of four recently proposed oil price forecasting models is systematically more accurate than the no-change forecast as well as forecast combinations based on recursive or rolling inverse MSPE weights. The forecasting models considered in that study included a vector autoregressive (VAR) forecast, forecasts based on the spread between oil futures prices and the spot price of oil, forecasts based on non-oil industrial commodity prices, and forecasts based on a time-varying parameter (TVP) model of the spreads between the U.S. spot prices of gasoline and heating oil and the spot price of crude oil. More recent work by Baumeister et al. (forthcoming), which explored the predictive content of high-frequency data from financial and energy markets, uncovered evidence that an important additional source of real-time information about future oil prices is the cumulative change in U.S. crude oil inventories.

* Corresponding author.
E-mail addresses: CBaumeister@bankofcanada.ca (C. Baumeister), lkilian@umich.edu (L. Kilian), Thomas.Lee@eia.doe.gov (T.K. Lee).

[1] For a comprehensive review the reader is referred to Alquist et al. (2013). Subsequent contributions include Baumeister and Kilian (2014a,b, forthcoming), Baumeister et al. (2013), Chen (2014), Baumeister et al. (forthcoming), and Bernard et al. (2014), among others.

http://dx.doi.org/10.1016/j.eneco.2014.08.008
0140-9883/

In the current paper, we extend the set of models to be combined to include the latter forecast, which performs particularly well at horizons between one and two years.

Baumeister and Kilian (forthcoming) compared the accuracy of equal-weighted forecast combinations to that of individual forecasting models and showed that only pooled forecasts are systematically more accurate than the no-change forecast at all horizons up to 18 months or 6 quarters. In this paper, we show that including in addition forecasts based on U.S. crude oil inventories in the forecast combination substantially improves the accuracy of the pooled forecast at horizons between one and two years.

The use of the same weights for all forecast horizons in constructing these baseline results ensures that there are no discontinuous changes in the forecast path across horizons of the type that would arise if we switched forecasting models or forecast combinations from one horizon to the next. For example, using a VAR forecasting model at horizons up to one year and a no-change forecast for longer horizons, may result in a jump in the path of the oil price forecast at the one-year horizon, as illustrated in Baumeister and Kilian (2014a). Some forecast users find such jumps awkward, perhaps because of the difficulty of rationalizing such jumps from an economic point of view. Insisting on a smooth forecast path comes at the price of higher MSPEs, however. If a low MSPE is all we care about in forecasting, one clearly can improve on equal-weighted combinations of all oil price forecasting models. We illustrate this point by allowing for different forecast combinations to be chosen at each horizon. This strategy takes advantage of the fact that some oil price forecasting models perform well at short horizons, but were never intended for longer horizons, whereas other models perform best at longer horizons. We show that relaxing the constraint of a continuous forecast path substantially reduces the MSPE of the pooled forecast at all horizons, but especially at horizons beyond one year.

This fact raises the question of how forecast pooling by horizon compares with simply selecting for each horizon the individual forecasting model with the lowest MSPE. The latter comparison is the relevant benchmark when evaluating the benefits of pooling in the absence of the continuity constraint. We find that pooled forecasts often, but not always have lower MSPE than the best individual forecast. The superior accuracy of the forecast combination at some horizons is not surprising in that pooled forecasts provide insurance against failures of individual models. Our results show that this insurance has a price in the form of lower forecast accuracy in some dimensions, however. For example, at horizons beyond 18 months, the individual forecasts are clearly more accurate. This drawback is offset by the fact that the accuracy of the pooled forecasts is more stable over time, as revealed by plots of the recursive MSPE ratios over time.

The MSPE of the real-time pooled forecasts is up to 29% lower than that of the no-change forecast even at horizons as high as two years. The pooled forecasts also predict the direction of change in the real price of oil correctly with probabilities as high as 73%. Our qualitative results are robust to alternative oil price measures and apply to monthly as well as quarterly forecasts. In addition to presenting these summary statistics, we use graphical methods to examine how the pooled real-time forecasts performed in recent years when the real price of oil fluctuated substantially. We compare these model-based pooled forecasts to the U.S. Energy Information Administration's (EIA) short-term oil price forecasts, as released in the *Short-Term Energy Outlook*. Finally, we discuss how real-time pooled forecasts of the nominal oil price may be derived from the forecasts of the real price, and we illustrate that both real and nominal oil price forecasts may be presented in a format already used by the EIA.

The remainder of the paper is organized as follows. In Section 2 we review the forecasting models considered. Section 3 evaluates our monthly forecasts of the real U.S. refiners' acquisition cost for oil imports and of the West Texas Intermediate (WTI) price of crude oil. In Section 4 we extend the analysis to quarterly forecasts. Section 5 examines how stable the accuracy of these oil price forecasts is over time. In Section 6 we visually compare the accuracy of our pooled oil price forecasts to that of the EIA oil price forecasts during key episodes. In Section 7 we illustrate how these forecasting tools may be used to produce real-time forecasts of the real and the nominal price of oil in a format consistent with that employed by the EIA in releasing its short-term oil price forecasts. The concluding remarks are in Section 8.

2. The forecasting environment

All forecasting models are estimated at monthly frequency. We consider monthly forecast horizons up to two years. Forecasts at the corresponding quarterly horizons are obtained by averaging the monthly forecasts at quarterly frequency, as recommended in Baumeister and Kilian (2014a). The forecasting models are estimated recursively and subject to real-time data constraints. All data are obtained from the real-time database developed in Baumeister and Kilian (2012, forthcoming) and extended in Baumeister et al. (2013). The reader is referred to the latter references for a detailed description of the data sources and definitions. The evaluation period is January 1992 through September 2012 (or equivalently the first quarter of 1992 through the third quarter of 2012). Our objective is to forecast the ex-post revised real price of oil, as measured by the observations in the March 2013 vintage of the real-time database.

The real-time forecasts are evaluated based on their recursive MSPE and their directional accuracy, as measured by the success ratio. The success ratio is the fraction of times that a method correctly predicts the direction of change in the real price of oil. Success ratios above 0.5 indicate an improvement relative to the no-change forecast. The MSPE results are normalized relative to the no-change forecast, with a ratio below 1 indicating a gain in accuracy. There is no valid test for judging the statistical significance of the MSPE reductions in our context, but we examine the stability of our results across horizons, across specifications and over time.[2] The statistical significance of the success ratios is assessed based on the test proposed in Pesaran and Timmermann (2009).

Building on the comprehensive analysis of forecast combination methods in Baumeister and Kilian (forthcoming), we consider five forecasting models with proven credentials.

2.1. Forecasts based on a VAR model of the global oil market

The first model is a reduced-form VAR model of the form:

$$B(L)y_t = \nu + u_t$$

where $y_t = [\Delta \text{prod}_t, rea_t, r_t^{\text{oil}}, \Delta \text{inv}_t]'$ refers to a vector including the percent change in global crude oil production, a measure of global real economic 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, ν denotes the intercept, $B(L) = I_4 - B_1L - \ldots - B_pL^p$ denotes the autoregressive lag order polynomial, p is the autoregressive lag order, L is the lag operator, and u_t is a white noise innovation.[3] This VAR model may be viewed as the reduced-form

[2] There are four distinct problems in testing the statistical significance of MSPE reductions in our context. First, traditional tests for equal predictive accuracy for nested models are concerned with testing predictability in population. There are only very limited results on testing the equality of out-of-sample MSPEs obtained from recursively estimated models. Second, these results are limited to direct forecasts. No results are available for iterated forecasts of the type considered in our analysis. Third, the tests in question have been designed for comparing pairs of nested forecasting models. They have not been designed for evaluating the accuracy of forecast combinations, some of which are nested in the benchmark model and some of which are not. Fourth, these problems are further compounded by the presence of real-time data constraints, which affect the distribution of the test statistics in question (see Clark and McCracken, 2013).

[3] 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 global real activity index is constructed from data on global dry cargo ocean shipping freight rates as described in Kilian (2009).