Energy Economics 70 (2018) 472–483

Contents lists available at ScienceDirect

Energy Economics

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

Forecasting the prices of crude oil: An iterated combination approach

Yaojie Zhang, Feng Ma *, Benshan Shi, Dengshi Huang

School of Economics & Management, Southwest Jiaotong University, Chengdu, China

ARTICLE INFO

Article history:
Received 28 October 2017
Received in revised form 18 January 2018
Accepted 25 January 2018
Available online 1 February 2018

JEL classification:
C32
C53
Q43
Q47
G11
G17

Keywords:
Oil price predictability
Iterated combination
Out-of-sample forecasts
Asset allocation

ABSTRACT

In this paper, we employ an iterated combination approach to examine oil price predictability with a large set of predictors, including 18 macroeconomic variables and 18 technical indicators. The empirical results show that iterated combination approach outperforms the standard combination approach for both in- and out-of-sample. Specifically, the iterated combination forecasts always yield significantly larger out-of-sample R^2 values and higher success ratios than the corresponding standard combination forecasts. Furthermore, we document that the results are robust to various settings, including alternative proxies of crude oil prices, three predictor sets, different forecasting windows, and various standard combination approaches. From an asset allocation perspective, we measure the economic value of the iterated combination approaches, where the leverage of oil futures trading is considered. The results suggest that the more accurate forecasts of the iterated combination approaches can generate substantially larger certainty equivalent return (CER) gains for a mean-variance investor in practice.

1. Introduction

Oil price shocks have crucial impacts on the financial markets and the real economy (Apergis and Miller, 2009; Hou et al., 2016; Kilian and Park, 2009; Kilian and Vigfusson, 2013, 2017; Park and Ratti, 2008; Wang et al., 2013). The managers and policy makers from oil-related industries, central banks, and governments are thus interested in the prediction of crude oil prices. Accordingly, numerous academic studies investigate oil price predictability (see, e.g., Baumeister and Kilian, 2012, 2015; Baumeister et al., 2014, 2017; Drachal, 2016; Han et al., 2017; Naser, 2016; Wang et al., 2015, 2017; Yin and Yang, 2016; Zhang et al., 2015).

Similar with this paper, many studies improve the out-of-sample predictive performance of oil prices based on various advanced forecasting models, including the vector autoregressive (VAR) model (Baumeister and Kilian, 2012, 2014; Kilian, 2009), predictive regressions with both economic and statistical restrictions (Wang et al., 2015), forecast combinations (Baumeister and Kilian, 2015; Baumeister et al., 2014; Wang et al., 2017), mixed-frequency data sampling model (Baumeister et al., 2015), and dynamic model averaging (Drachal, 2016; Naser, 2016). Compared with the extant studies, this paper makes three primary contributions to the literature on oil price predictability as follows.

First, we employ a novel combination approach to forecast crude oil prices. Lin et al. (forthcoming) propose an iterated combination approach to improve the predictability of bond return. To the best of our knowledge, there is, however, no study that uses this method to forecast the returns of oil prices previously. A prominent strength of the iterated combination approach is that it can capture useful information content from a large set of predictors; therefore, we consider 36 predictors, including 18 macroeconomic variables and 18 technical indicators. Our empirical results suggest that compared with the conventional combination approaches used by Rapach et al. (2010), the corresponding iterated combination approaches exhibit relatively good in-sample performance (that is, large in-sample R^2 values). More importantly, the iterated combination approaches always generate more accurate oil price forecasts than the conventional combination approaches during the out-of-sample period, which is evaluated by both out-of-sample R^2 statistics and success ratios.

Second, we further use 5 additional combination approaches used by Stock and Watson (2004) and Wang et al. (2016) to extend the existing iterated combination approaches. The results are similar for the new iterated combination approaches, indicating the robustness and compatibility of the iterated combination approach. Furthermore, we document that the results are robust to various settings including alternative proxies of crude oil prices, three predictor sets, different forecasting

* Corresponding author.
E-mail address: mafeng2016@swjtu.edu.cn (F. Ma).

https://doi.org/10.1016/j.eneco.2018.01.027
0140-9883/

windows, and various evaluation criteria. More specifically, we consider two popular proxies of crude oil prices, the US refiner's acquisition cost for imported crude oil and the spot price of West Texas Intermediate crude oil. With respect to predictor sets, this paper constructs three sets comprising 18 macroeconomic variables, 18 technical indicators, and all of those 36 variables, respectively. To avoid the problem caused by the arbitrary choices of forecasting window sizes (see, e.g., Rossi and Inoue, 2012), we carefully choose three different reasonable expanding windows to predict oil prices. Finally, from both statistical and economic perspectives, we examine various forecasting models' out-of-sample performances using the out-of-sample R^2, success ratio, and certainty equivalent return (CER) gain. Fortunately, we obtain robust results for all these settings.

Third, this paper measures the economic value of oil price predictability of the iterated combination approach from an asset allocation perspective. The economic value is examined by voluminous literature on stock and bond return predictability (see, e.g., Campbell and Thompson, 2008; Huang et al., 2015; Jiang et al., 2017; Lawrenz and Zorn, 2017; Neely et al., 2014; Rapach et al., 2010, 2016; Zhu and Zhu, 2013). In contrast, there are, however, scant studies that measure the economic value of oil price predictability. The work of Yin and Yang (2016) appears to be the only one that provides an asset allocation exercise using the forecasts of oil futures returns. An important feature of futures trading is that an investor does not need to pay all of the money for traded futures and just maintains a margin account at a required level that is substantially smaller than the entire value of the futures. Consequently, the actual gains (losses) of the investor are several times larger than the increase (decline) of futures price, which is the so-called leverage that enables gains and losses to be multiplied. Given this, we further provide the optimal portfolio weight of oil futures for a mean-variance investor who employs leverage to allocate between oil futures and risk-free bills. To the best of our knowledge, this paper is the first to fill this research gap. Furthermore, the results of asset allocation document that the iterated combination forecasts of oil futures returns always generate considerable larger certainty equivalent return (CER) gains than the corresponding standard combination forecasts. In other words, the iterated combination forecasts help the investor to make more money.

In summary, the iterated combination approaches employed in this study always outperform the corresponding standard combination approaches in both statistical and economic senses. Moreover, this paper also contributes to the application of asset allocation, in which we take the leverage of futures trading into consideration and accordingly adjust the optimal portfolio weights of oil futures and risk-free bills.

The remainder of the paper is organized as follows. Section 2 provides the econometric methodology including standard combination methods and the iterated combination approaches. Section 3 describes our data. We report the empirical results in Section 4. In Section 5, we make several robustness checks. Section 6 reports the results from an asset allocation perspective. The last section concludes the paper.

2. Methodology

In this section, we introduce the iterated combination approach pioneered by Lin et al. (forthcoming). This new forecasting methodology extracts the information from data-rich indicators and has a close linkage with the partial least squares (PLS) forecasting method, which is recently developed by Kelly and Pruitt (2013) and Kelly and Pruitt (2015).

2.1. Standard combinations

To predict crude oil prices, we begin with a univariate regression model,

$$r_{t+1} = \alpha_i + \beta_i x_{i,t} + \varepsilon_{i,t+1}, \tag{1}$$

where $r_{t+1} = \ln(y_{t+1}) - \ln(y_t)$, denoting the log return of oil price at month $t+1$, y_t is the oil price at month t, $x_{i,t}$ represents the ith predictor available at month t, and $\varepsilon_{i,\,t+1}$ is an error term with a mean equal to zero.

Running the predictive regression of Eq. (1) with the data up to month t, we can obtain the individual forecasts as

$$\hat{r}_{i,t+1} = \hat{\alpha}_i + \hat{\beta}_i x_{i,t}, \tag{2}$$

where $\hat{r}_{i,t+1}$ is the individual forecast using the ith predictor at month $t+1$, and $\hat{\alpha}_i$ and $\hat{\beta}_i$ are the regression coefficients from the univariate regression for the ith predictor.

The standard combination forecasts are weighted averages of the N individual forecasts, which are calculated as

$$\hat{r}_{C,t+1} = \sum_{i=1}^{N} \omega_{i,t} \hat{r}_{i,t+1}, \tag{3}$$

where $\hat{r}_{C,t+1}$ denotes the combination forecasts at month $t+1$ and $\omega_{i,\,t}$ denotes the combining weight for the ith individual forecast formed at t.

Following the influential study of Rapach et al. (2010), we use 5 popular combination approaches in this paper.[1]

- Mean combination (MC). The mean combination forecast is the equal-weighted average of the N individual forecasts, $\{\hat{r}_{i,t+1}\}_{i=1}^{N}$.
- Median combination (MDC). This combination approach uses the median of $\{\hat{r}_{i,t+1}\}_{i=1}^{N}$.
- Trimmed mean combination (TMC). The trimmed mean combination forecast sets $\omega_{i,\,t} = 0$ for the largest and smallest individual forecasts in $\{\hat{r}_{i,t+1}\}_{i=1}^{N}$ and $\omega_{i,\,t} = 1/(N-2)$ for the remaining individual forecasts.
- Discount mean square prediction error (DMSPE) combining method. In the DMSPE method, the combining weight of model i at month t is calculated as $\omega_{i,\,t} = \phi_{i,\,t}^{-1}/\sum_{\ell=1}^{N}\phi_{\ell,\,t}^{-1}$, where $\phi_{i,t} = \sum_{s=m+1}^{t}\theta^{t-s}(r_s - \hat{r}_{i,s})^2$, m is the number of observations in the initial training sample, and θ denotes a discount factor. Following Rapach et al. (2010) and Zhu and Zhu (2013), we consider two values of θ, namely, 1 and 0.9 such that two DMSPE methods, DMSPE(1) and DMSPE(0.9), are employed in this study. Finally, note that the first out-of-sample forecasts of DMSPE(1) and DMSPE(0.9) are simply calculated as the mean combination forecast because there is no past individual forecast used to form the DMSPE weight at this time point.

2.2. Iterated combinations

Lin et al. (forthcoming) propose a further combining method based on both the standard combination forecasts in Eq. (3) and a simple benchmark forecast. This advanced combining method is the so-called iterated combination, which can be expressed as

$$r_{t+1} = (1-\lambda)\hat{r}_{B,t+1} + \lambda \hat{r}_{C,t+1} + \varepsilon_{C,t+1}, \tag{4}$$

where λ is the restricted regression coefficient that is estimated by the restricted least-squares estimation, $\hat{r}_{B,t+1}$ is the benchmark forecast at month $t+1$, and $\hat{r}_{C,t+1}$ is a combination forecast at month $t+1$. Note that Lin et al. (forthcoming) use the historical average as the benchmark model to forecast bond returns, while we use the on-change forecasts as the benchmark model,[2] which is more popular and suitable for

[1] Rapach et al. (2010) also consider other complicated combination methods, in which the combining weights are computed more elaborately by using in-sample model fit such as the Schwarz information criterion. However, Rapach et al. (2010) find that these complicated methods generate poor forecasting performance in contrast to the simple methods, which is consistent with the forecast combination puzzle that the estimated optimal combination forecasts commonly perform poorly, but the arithmetic mean typically performs well. See Claeskens et al. (2016) for further details.

[2] The iterated combination forecasts based on no-change forecasts outperform the iterated combination forecasts using the historical average in our applications below.