Forecasting oil prices with penalized regressions, variance risk premia and Google data

Fantazzini, Dean and Kurbatskii, Alexey and Mironenkov, Alexey and Lycheva, Maria

2022

M. Lycheva, A. Mironenkov, A. Kurbatskii, D. Fantazzini [1]

Forecasting oil prices with penalized regressions, variance risk premia and Google data

This paper investigates whether augmenting models with the variance risk premium (VRP) and Google search data improves the quality of the forecasts for real oil prices. We considered a time sample of monthly data from 2007 to 2019 that includes several episodes of high volatility in the oil market. Our evidence shows that penalized regressions provided the best forecasting performances across most of the forecasting horizons. Moreover, we found that models using the VRP as an additional predictor performed best for forecasts up to 6-12 months ahead forecasts, while models using Google data as an additional predictor performed better for longer-term forecasts up to 12-24 months ahead. However, we found that the differences in forecasting performances were not statistically different for most models, and only the Principal Component Regression (PCR) and the Partial least squares (PLS) regression were consistently excluded from the set of best forecasting models. These results also held after a set of robustness checks that considered model specifications using a wider set of influential variables, a Hierarchical Vector Auto-Regression model estimated with the LASSO, and a set of forecasting models using a simplified specification for Google Trends data.

Keywords: Oil price; Variance Risk Premium; Google Trends; VAR; LASSO; Ridge; Elastic Net; Principal components, Partial least squares.

JEL classification: C22; C32; C52; C53; C55; C58; G17; O13; Q47.

1. Introduction

The real price of oil plays an important role almost in every economic sector. Accurate forecasting of this macroeconomic variable provides an opportunity for oil-importing and exporting countries, investors, and other economic agents to develop more efficient business strategies and plan more balanced economic activity. Moreover, the forecast is also important for energy policy modelling, energy system planning, and carbon emission regulations, see Baumeister and Hamilton (2019), Bhattacharyya (2019), Fantazzini et al. (2011), Fantazzini (2016), Hamilton (2008, 2009, 2013), Kilian (2008, 2009, 2016), Kilian and Zhou (2022), and Schwarz (2017) for a broader discussion.

The first contribution of this work is the introduction of two new additional predictors based on the Variance Risk Premium (VRP) originally proposed by Bollerslev et al. (2009) and Bollerslev et al. (2014), which is defined as the difference between the implied volatility (IV) and the realized volatility (RV). There is a large literature that showed that the VRP can be an effective variable to predict future equity returns, see Atmaz (2022) and references

[1] **Lycheva, Maria** — ICEF-HSE, Moscow.
Alexey Mironenkov — Moscow School of Economics, Moscow State University.
Alexey Kurbatskii — Moscow School of Economics, Moscow State University.
Fantazzini, Dean — Moscow School of Economics, Moscow State University, Moscow; fantazzini@mse-msu.ru (corresponding author).

This is the working paper version of the paper *Forecasting oil prices with penalized regressions, variance risk premia and Google data*, forthcoming in the journal ***Applied Econometrics***.

therein for more details. The intuition behind this evidence is that the increase (decrease) of risk aversion leads to asset prices being discounted, and later this discount is accrued, thus resulting in future higher (lower) returns. Similar patterns can also be found for exchange rates see, for example, Londono and Zhou (2017) and Ornelas (2019), and for commodities, see Ornelas and Mauad (2019). The VRP has been rarely used to forecast oil prices and, to our knowledge, only Chevalier and Sevi (2014) and Ornelas and Mauad (2019) employed this variable for in-sample analyses to model oil price dynamics without any out-of-sample forecasting. Following Chevalier and Sevi (2014), we computed the VRP by using the CBOE Crude Oil Volatility Index (OVX) as a proxy for the implied volatility[2], while the monthly realized volatility was computed using the sum of the daily realized volatilities over a given month. More recently, Bazhenov, Fantazzini (2019), Fantazzini and Shangina (2019), and Basistha et al. (2020) compared the role of implied volatility and Google's online search data for forecasting the realized volatility and risk measures of several financial assets, and they found that models with the implied volatility performed generally better than models with Google data. These authors suggested that the informational content included in Google data is also present in the implied volatility, but the opposite is not true: the implied volatility is a forward-looking measure based on the expectations of large investors with premium and insider information, while Google search data are based on the expectations of small investors and uninformed traders. Motivated by this evidence, we decided to build a second predictor mimicking the variance risk premium, but where the squared log-returns of the monthly Google Trends data were used in place of the implied volatility to try to measure the sentiment of retail investors. We justify this choice because behavioral factors might influence and improve oil price forecasts, see Qadan and Nama (2018) and references therein for a broad discussion. Therefore, the traditional VRP computed using the implied volatility can be considered a measure representing the risk aversion of large institutional investors, while the second predictor that uses Google data as a measure of the risk aversion of small retail investors.

The second contribution of this work is the proposal of a set of univariate penalized regressions that includes the main predictors proposed by the past literature devoted to oil price forecasting, together with predictors based on the implied volatility and Google search data. We proposed this kind of univariate models due to recent empirical evidence provided by Miao et al. (2017) and Zhang et al. (2019) who showed that univariate regressions with direct forecasts outperformed several competing models widely used nowadays for oil price forecasting (but without considering vector autoregressive models).

The third contribution of this paper is a forecasting exercise where we considered different forecast horizons up to 24-month ahead and different competing model specifications, including common vector autoregressive models that were not considered in the recent literature proposing univariate penalized regressions with direct forecasts.

The fourth contribution of the paper is a set of robustness checks to verify that our results also hold when considering model specifications using a wider set of influential variables, a Hierarchical Vector Auto-Regression model estimated with the LASSO, and a set of forecasting models using a simplified specification for Google Trends data.

The empirical analysis provided evidence that penalized regressions (particularly the Ridge and the Elastic net models) had the best forecasting performances across most forecasting horizons. Moreover, we found that models using the VRP as an additional predictor performed best for forecasts up to 6-12 months ahead forecasts, while

[2] The OVX index measures the market's expectation of the 30-day volatility of crude oil prices. See *https://www.cboe.com/us/indices/dashboard/ovx* for more details.

models using Google data as an additional predictor performed better for longer-term forecasts up to 12-24 months ahead. The original model by Kilian and Murphy (2014) using the full sample starting in 1973 performed well for short-term forecasts, but it became less competitive with longer-term horizons and, in the case of 24-month ahead forecasts, it was even excluded from the model confidence set. The traditional benchmark represented by the no-change forecast performed well across all forecast horizons, whereas approaches for dimensionality reduction such as the principal component regression (PCR) and the partial least squares (PLS) regression performed poorly across all horizons. However, the empirical evidence also showed that the differences in forecasting performances were not statistically different for most models, and only the PCR and PLS models were consistently excluded from the set of best forecasting models.

The remainder of the paper is organized as follows. Section 2 reviews the literature devoted to oil price forecasting, while Section 3 discusses the data and the methods proposed to model and forecast the oil price. Section 4 describes the empirical results, while robustness checks are discussed in Section 5. Section 6 concludes.

2. Literature review

After the oil price spike in 2008 and the subsequent crash, there has been a large number of works dedicated to oil price forecasting. It is important to note that the oil market has changed a lot since then due to the advent of shale oil in the US and the massive financialization of the oil markets that allowed trading also to small traders, see e.g. Hamilton (2011), Alquist et al. (2013), Kilian (2016), Baumeister and Hamilton (2019), and Kilian and Zhou (2022) for a broad discussion. Therefore, we focused in this work only on papers dealing with the forecasting of oil prices published after that event.

The literature has become large in the last 15 years but there are at least two models that can be considered the main benchmarks for forecasting the real price of oil: the simple no-change forecast (that is, the random walk model without drift) and the VAR model proposed by Kilian and Murphy (2014). The latter model consists of a VAR(24) model with four variables (the global crude oil production, a global real activity measure, a proxy for the global above-ground crude oil inventories, and the real price of oil), together with centered seasonal dummies to take care of monthly seasonality. We remark that the issue of seasonality in the oil market is still somewhat controversial. For example, Coleman (2011) found no evidence of significant seasonality and suggested that the effects attributed to seasonality can be attributed to changes in the oil market, such as changes in OPEC policies, seasonality in terrorist attacks, and speculative activity. Instead, Quayyoum et al. (2019) found significant positive Monday and Thursday effects while, on a monthly basis, oil returns appear to be significantly negative in November and December.

A recent strand of the literature proposed univariate models with LASSO and direct forecasts to model the dynamics of oil prices. Miao et al. (2017) considered univariate LASSO models with different sets of predictors included in six broad classes such as supply, demand, political factors, speculative factors, as well as variables connected with commodity and financial markets. They showed the superiority of LASSO models for forecasting oil prices against all competing models, including also the no-change random walk forecast and a full factor VAR model. However, the original VAR model by Kilian and Murphy (2014) was not considered. Zhang et al. (2019) employed a wide set of models, ranging from a univariate LASSO-based model, to predictive regressions based on principal components, partial least squares, and to several types of forecast combination approaches. They showed that the univariate LASSO model provided the better out-of-sample forecasts according to several metrics. Moreover, they also showed that investors obtain higher economic gains when using the LASSO model compared to any other approach. Similar to Miao et al. (2017), the no-change random walk forecast was included among the competing models, whereas the original VAR model by Kilian and Murphy (2014) was not.

A couple of papers used the variance risk premium to model the oil price dynamics. Chevallier and Sevi (2013) examined the predictive power of the variance risk premia for WTI light sweet crude oil excess returns, and the VRPs were computed as the difference between model-free implied and realized volatility measures following Bollerslev et al. (2009). They also considered additional predictors related to macroeconomic, financial, and oil-specific variables. They found that the explanatory power of the VRP on oil excess returns reaches up to 25% for the adjusted R-squared across several regressions and can complement other financial factors. Ornelas and Mauad (2019) studied the importance of VRP in modeling commodity prices, and they showed that there is a positive relation between commodity VRP and its future returns for most commodities. The general idea behind these papers is to consider the VRP as a measure for an investor's risk aversion: when it increases, the risky asset is discounted, so this discount will be accrued over time and will lead to higher returns in the future.

Fantazzini and Fomichev (2014) were the first to use Google search data to forecast the real price of oil. Their forecasting exercises showed that multivariate models with economic and energy aggregates outperformed the competing models only up to 3 steps ahead, whereas low dimensional models with Google data outperformed the other models up to 24 steps ahead. Moreover, the forecasting power of the best models using Google data increased with the length of the forecast horizon, particularly with forecast horizons higher than 12 steps ahead. Qadan and Nama (2018) showed that Google online data can be used as a representation of investors' sentiment. They highlighted that not only real economic factors but also behavioral factors should be used to predict oil price movements, and they showed that these factors have a significant effect on oil prices. Interestingly, they found that their results are most significant during and after the early 2000s when financial products based on oil commodities started to gain popularity among investors. Afkhami et al. (2017) searched for the best keywords in Google Trends and their combinations that could be used as a proxy for investors' attention in energy commodities markets. They showed that these new predictors significantly improved the in-sample model fitting of oil price volatility beyond conventional GARCH models. We used their approach to select the best keywords combinations to capture investors' attention when building a predictor mimicking the variance risk premium based on Google search data. In this regard, we remark that some papers compared the use of implied volatility and Google data for forecasting the volatility and the risk measures of several financial and commodity markets (including oil), see Fantazzini and Shangina (2019) and Basistha et al. (2020). They found that models with implied volatility performed better than models with Google Trends and they showed that the informational content included in Google search data is also present in the implied volatility, but the opposite is not true. They suggested that this is probably because the implied volatility is a forward-looking measure based on the expectations of large investors who have access to premium and insider information, while Google Trends data are mainly based on the expectations of small investors and uninformed traders. In our work, we built predictors based on both implied volatility and Google data and we examined which of them is more useful for forecasting real oil prices.

3. Methodology

The goal of this paper is to verify whether augmenting models with predictors based on the implied volatility and Google search data improves the quality of the forecasts for real oil prices. Before presenting the main empirical results, we review the models employed in our forecasting exercise.

3.1. Benchmark models: Vector Auto-Regressive models (VARs) and the no-change forecast

Oil market *VAR models* have become one of the main tools for modelling the dynamics of the real price of oil and understanding its macroeconomic effects, see Kilian and Zhou (2022) and references therein for a large discussion. We considered several VAR model specifications in our empirical work.

First, we considered the VAR(24) model with centered seasonal dummies and four variables proposed by Kilian and Murphy (2014), which is one of the benchmark models to forecast the real price of oil:

$$\mathbf{Y}_t = \boldsymbol{\nu} + \sum_{i=1}^{p} \mathbf{\Phi}_{\mathrm{i}} \mathbf{Y}_{t-i} + \boldsymbol{u}_t, \quad \boldsymbol{u}_t \sim WN\left(0, \mathbf{\Sigma}\right)$$

where $\mathbf{Y}_t$ is a 4 × 1 vector containing the log of the real price of oil, the percent change in world crude oil production, the first differences of the proxy for the global above-ground crude oil inventories, and the global real activity measure (REA index). This model was estimated using the full available data sample from 1973 till 2019.

Second, we also considered the same VAR model but augmented with the following predictors:

- the VRP based on the model-free implied volatility;
- the proxy VRP based on Google Trends data;
- the VRP based on the model-free implied volatility and the proxy VRP based on Google Trends data.

These augmented VAR models were estimated using the smaller sample from 2007 till 2019, while the optimal VAR lag were chosen using the AIC criterion. To decrease the model dimensionality, the simple average of all Google search keywords was employed when using these augmented VAR models[3].

The *no-change forecast* is another traditional benchmark model for oil price forecasting, see Baumeister and Kilian (2012, 2015) for a large discussion. The no-change forecast can be interpreted as optimal under a random walk model without drift,

$$y_t = y_{t-1} + \varepsilon_t$$

so the *h*-period oil price forecast is simply equal to today's oil price y_t.

3.2. Penalized regressions: LASSO, Ridge, and Elastic net

The *LASSO method* is a variable selection method originally introduced by Tibshirani (1996), where the choice of the predictive variables is performed using an algorithmic procedure. The idea behind the LASSO is to exclude insignificant variables, so the final model is more efficient compared to the full model. The LASSO method adds a penalty term to the cost function to keep the estimated value of the regression coefficients small, thus reducing

[3] We did not pre-test the data for unit roots and cointegration, but we stuck to the original VAR model specification proposed by Kilian and Murphy (2014) due to its past forecasting success. Moreover, it is well known that unit root and cointegration tests tend to work poorly with small and medium datasets, see Gospodinov et al. (2013) for a large discussion. Furthermore, cointegrated models also suffer from estimation problems in the case of small-medium samples, particularly when noisy data such as Google data or proxy variables (like the global crude oil inventories data) are used, see section 4.4 in Fantazzini and Toktamysova (2015) and references therein for more details.

the problem of inflated standard errors caused by multicollinearity. More specifically, the LASSO minimizes the residual sum of squares subject to the sum of the absolute value of the coefficients being less than a constant. In other terms, it minimizes the least-squares cost function subject to a penalty term λ, which implies that the coefficients that are smaller than λ are set to zero. The parameter λ regulates the strength of the regularization, and the higher λ, the stronger is the penalty and the higher the probability that more variables will be considered insignificant and excluded from the model. The motivation to use the LASSO model is to improve the accuracy of the estimated parameters because it reduces overfitting (thanks to the exclusion of insignificant variables), but it does not increase the bias. Statistically, the LASSO model can be expressed as,

$$\hat{\beta}^L = \arg\min\left\{\sum_{t=1}^{T}\left(y_t - \beta_0 - \sum_{j=1}^{p}\beta_j x_{t,j}\right)^2\right\},\ s.t.\sum_{j=1}^{k}\left|\hat{\beta}_j\right| \le c$$

which is equivalent to,

$$\hat{\beta}^L = \arg\min\left\{\sum_{i=1}^{n}\left(y_t - \beta_0 - \sum_{j=1}^{p}\beta_j x_{t,j}\right)^2 + \lambda\sum_{j=1}^{k}\left|\beta_j\right|\right\},\ \lambda > 0$$

where y_t is real oil price and $x_{t,j}$ is a set of lagged predictors, which in our case consists of the following variables: world crude oil stocks, global crude oil production, REA index, variance risk premia and centered seasonal dummies. To make the univariate models comparable to VAR models, we considered up to 24 lags for each predictor.

The *ridge regression* was originally proposed by Hoerl and Kennard (1970), and it is similar to the LASSO approach. This method adds a penalty term to the usual minimization of squared errors, which results in biased but more efficient estimates compared to ordinary least squares. Moreover, this method is particularly useful in the presence of multicollinearity. Statistically, the ridge model can be expressed as follows:

$$\hat{\beta}^R = \arg\min\left\{\sum_{t=1}^{T}\left(y_t - \beta_0 - \sum_{j=1}^{p}\beta_j x_{t,j}\right)^2 + \lambda\sum_{j=1}^{k}\beta_j^2\right\},\ \lambda > 0$$

The *elastic net regularization* proposed by Zou and Hastie (2005) is a generalization of the LASSO and ridge penalizations, and it is the minimizer of this equation

$$\hat{\beta}^{EN} = \arg\min\left\{\sum_{i=1}^{n}\left(y_t - \beta_0 - \sum_{j=1}^{p}\beta_j x_{t,j}\right)^2 + \lambda_1\sum_{j=1}^{k}\left|\beta_j\right| + \lambda_2\sum_{j=1}^{k}\beta_j^2\right\}$$

This approach wants to combine the strengths of both penalizations in a computationally tractable way: the L^1 norm penalty allows the generation of a sparse model, while the L^2 norm penalty eliminates the limitation on the number of selected variables. Moreover, this approach encourages a grouping effect because the coefficients of a group of highly correlated variables tend to be equal, see Zou and Hastie (2005) for more details.

3.3 Approaches for dimensionality reduction: the Principal Component Regression (PCR) and the Partial least squares (PLS) regression

The *principal component regression (PCR)* first employs principal component analysis to transform the original set of lagged predictors into a few new variables known as principal components (PC), which are a linear combination of the original lagged data. The principal components are then used to estimate a linear regression model:

$$y_t = \alpha + \sum_{k=1}^{K}\beta_k PC_{k,t} + \varepsilon_t$$

Following Neely et al. (2014) and Zhang et al. (2019), the optimal number K of principal components is selected by using the adjusted R^2. Similar to the LASSO and the ridge regression, the PCR is worth using when the data set contains highly correlated variables.

A possible limitation of PCR is that the selected principal components may not be associated with the dependent variable because their selection does not depend on the outcome variable. Instead, *the Partial Least Squares (PLS)* selects the new principal components using not only the original predictors but also the dependent variable. These components are then employed to estimate a linear regression model. Similar to PCR, PLS is a convenient approach in the case of highly-correlated regressors. There are several algorithms to implement the partial least squares regression, see, for example, Hastie et al. (2009) and Vinzi et al. (2010) and for a large discussion at the textbook level.

3.4. Forecasting model evaluation

The forecasting performance of the different models was checked by comparing the forecasted values of the real oil price with the actual oil price for each month, and then computing the traditional forecast evaluation statistics such as the mean squared error (MSE), and the Mean Absolute Error (MAE).

The MSE loss function was subsequently employed with the *Model Confidence Set* (MCS) by Hansen et al. (2011) to select the best forecasting models at a specified confidence level. Given the difference between the MSEs (or the MAEs) of models i and j at time t (for example, $d_{i,j,t} = MSE_{i,t} - MSE_{j,t}$), the MCS approach is used to test the following hypothesis of equal predictive ability, $\mathbf{H_{0,M}}$: $\mathrm{E}(d_{i,j,t}) = 0$, for all $i,j \in M$, where M is the set of forecasting models. First, the following t-statistics are computed, $t_{i\cdot} = \bar{d}_{i\cdot}/\widehat{var}(\bar{d}_{i\cdot})$ for $i \in M$, where $\bar{d}_{i\cdot} = m^{-1}\sum_{j\in M}\bar{d}_{ij}$ is the simple loss of the i-th model relative to the average losses across models in the set M, $\bar{d}_{ij} = T^{-1}\sum_{t=1}^{T} d_{ij,t}$ measures the sample loss differential between model i and j, and $\widehat{var}(\bar{d}_{i\cdot})$ is an estimate of $var(\bar{d}_{i\cdot})$. Secondly, the following test statistic is computed to test for the null hypothesis: $T_{max} = \max_{i\in M}(t_{i\cdot})$. This statistic has a non-standard distribution, so the distribution under the null hypothesis is computed using bootstrap methods with 5000 replications and a minimum block length equal to 12. If the null hypothesis is rejected, one model is eliminated from the analysis and the testing procedure starts from the beginning.

4. Empirical analysis

4.1. Data

We used monthly oil market data, Google Trends data, and macro variables for the 2007-2019 period to capture the shocks in oil demand and supply that affected the real price of oil. We chose this specific time sample because earlier data was not available for some variables, such as the implied volatility or Google data[4]. Data from 2020 onwards were not considered in this work because the Covid-19 pandemic represented a major structural break in the oil market and would require separate modelling. This is why we leave this issue as an avenue for further research.

The forecasted variable is the real price of oil, which we obtained by deflating the nominal oil price by the consumer price index (CPI). The nominal price of oil that we used in our analysis was the US crude oil imported acquisition cost by refiners published by the US Energy Information Administration, while the CPI was obtained from the Federal Reserve Economic Data (FRED). A plot of the real price of oil is reported in Figure 1.

Fig. 1. Real oil price ($/barrel).

As for the explanatory variables, we employed those originally proposed by Kilian and Murphy (2014). The first one is an estimate of the world crude oil stocks (see Figure 2): due to the lack of data for other countries, we constructed a proxy for global crude oil inventories following the Kilian and Murphy (2014) methodology, that is we took the values for the US crude oil inventories and scaled them by the ratio of the OECD petroleum stocks over US petroleum stocks. Similar to Kilian and Murphy (2014), we also transformed these data into first differences to achieve stationarity. All these data were obtained from the US Energy Information Administration.

[4] However, we remark that we employed the full available data sample from 1973 till 2019 when we used the VAR(24) model with centered seasonal dummies proposed by Kilian and Murphy (2014).

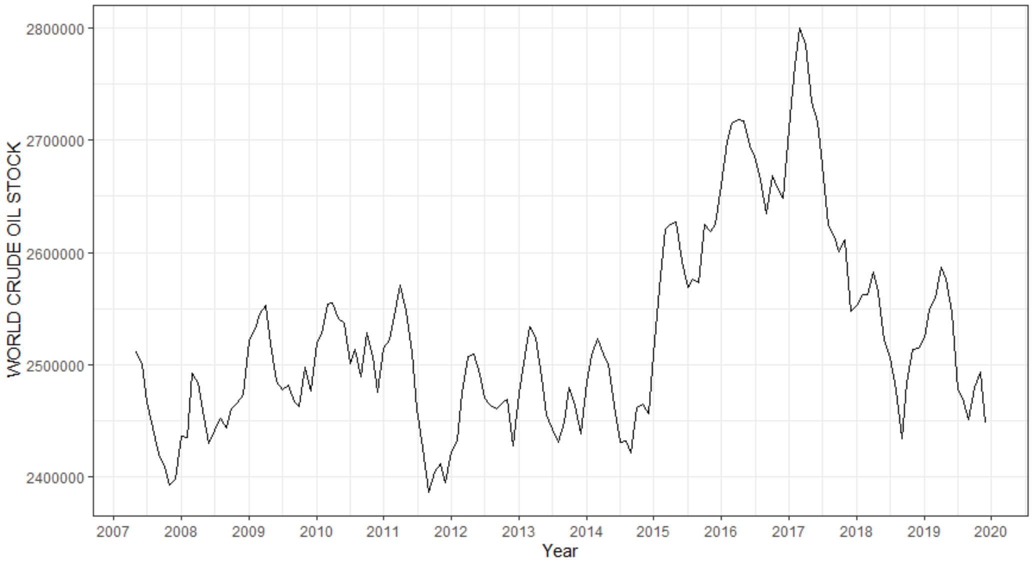

Fig. 2. World crude oil stocks (billion barrels).

Another variable is the Real Economic Activity (REA) index, which captures shifts in global economic activity and was originally developed by Kilian (2009), see Figure 3. The idea of this index is to consider freight rates as a measure of economic activity, due to their connection to demand factors. This index is stationary by construction, and it can be obtained from the Federal Reserve Bank of Dallas website. Since the index is based on various bulk dry cargoes consisting of grain, oilseeds, coal, iron, ore, scrap metals and fertilizers, an increase in freight rates indicates an increase in world demand.

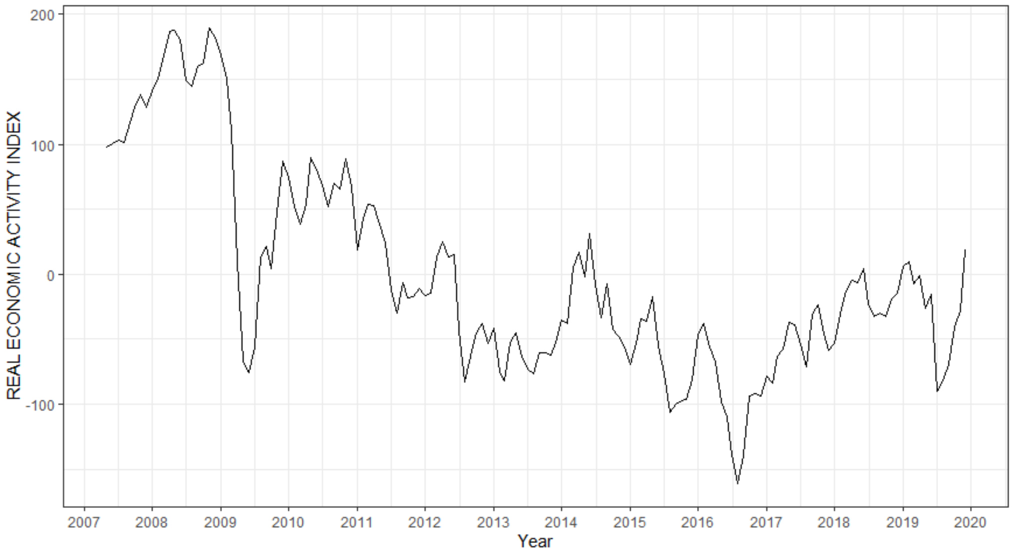

Fig. 3. Real Economic Activity Index.

The fourth variable in our analysis is the world crude oil production which can be downloaded from the US Energy Information Administration, see Figure 4. Following Kilian and Murphy (2014), we transformed the data into percentage changes to achieve stationarity. This variable is used to capture crude oil supply shocks.

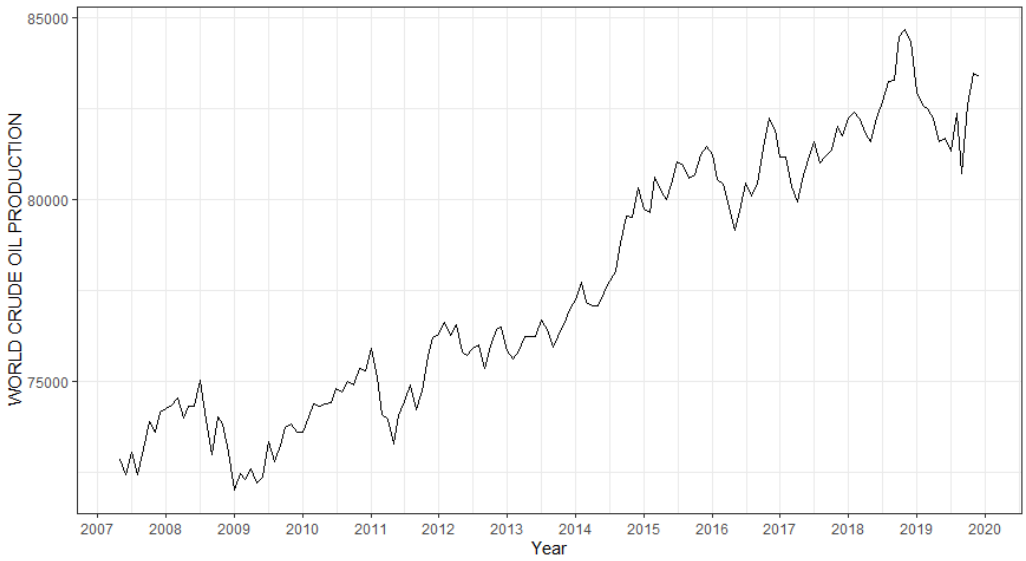

Fig. 4. World crude oil production (million barrels/day).

Besides the previous traditional drivers for the oil market, we also considered the variance risk premium (VRP) originally proposed by Bollerslev et al. (2009), which was later applied to the oil market for the first time by Chevalier and Sevi (2014). The VRP can be defined as the difference between the ex-ante risk-neutral expectation of the future return volatility and the ex-post realized return volatility:

$$VRP_t = E^Q[\sigma_{t,t+T}] - E^P[\sigma_{t,t+T}]$$

In practice, this variable is computed as the difference between a model-free estimate of the implied volatility (IV_t) and the realized volatility (RV_t) for a given WTI oil futures contract[5]. Following Chevalier and Sevi (2014), we used the CBOE crude oil volatility index (also known as OVX) as a proxy for the implied volatility, which estimates the market's expectation of the 30-day volatility of crude oil prices[6]. The monthly realized volatility was computed as the sum of the daily realized volatilities over a given month, and the VRP was calculated on a monthly basis as the difference between the IV and the RV series, see Bollerslev et al. (2009), Chevalier and Sevi (2014), and references therein for more details. A plot of the OVX index and the VRP is reported in Figure 5.

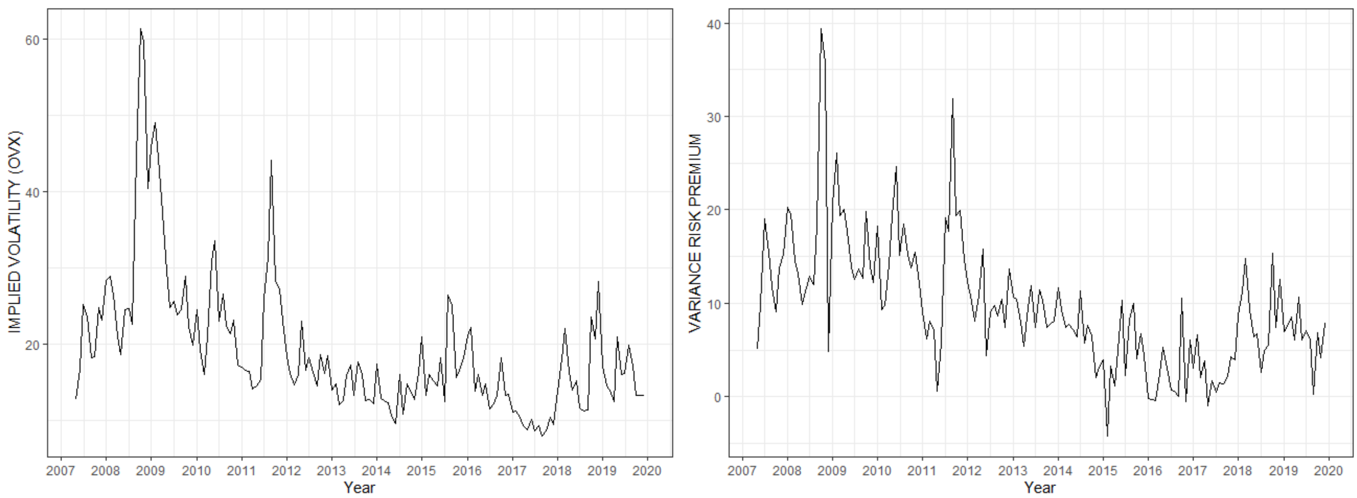

Fig. 5. CBOE OVX index (left) and estimated VRP for the oil price (right).

[5] Note that the VRP can be defined in different ways, see Carr and Wu (2009), Ornelas and Mauad (2019) and references therein for more details.

[6] See https://www.cboe.com/us/indices/dashboard/ovx/ for more details.

Bazhenov, Fantazzini (2019), Fantazzini and Shangina (2019), and Basistha et al. (2020) argued that the implied volatility is a forward-looking measure based on the expectations of large investors with premium and insider information, while Google search data are based on the expectations of small investors and uninformed traders. Motivated by this evidence, we decided to build a second predictor mimicking the previous variance risk premium, but where the squared log-returns of the monthly Google Trends data were used in place of the implied volatility to try to measure the sentiment of retail investors[7]. It goes without saying that this is not a proper variance risk premium, but only an attempt to have a (rough) proxy for the risk aversion of small retail investors and uninformed traders. We justify this choice because behavioral factors might influence and improve oil price forecasts, see e.g. Qadan and Nama (2018) for a detailed discussion. Following Afkhami et al. (2017), we used the following Google search keywords to best represent the oil market sentiment among small investors: "*crude oil*", "*petroleum*", "*Brent crude*". A plot of the Google Trends data and the proxy VRP based on these data is reported in Figure 6. Interestingly, the dynamics of the VRP computed with the model-free implied volatility reported in Figure 5 and the proxy VRP based on Google data in Figure 6 are rather similar.

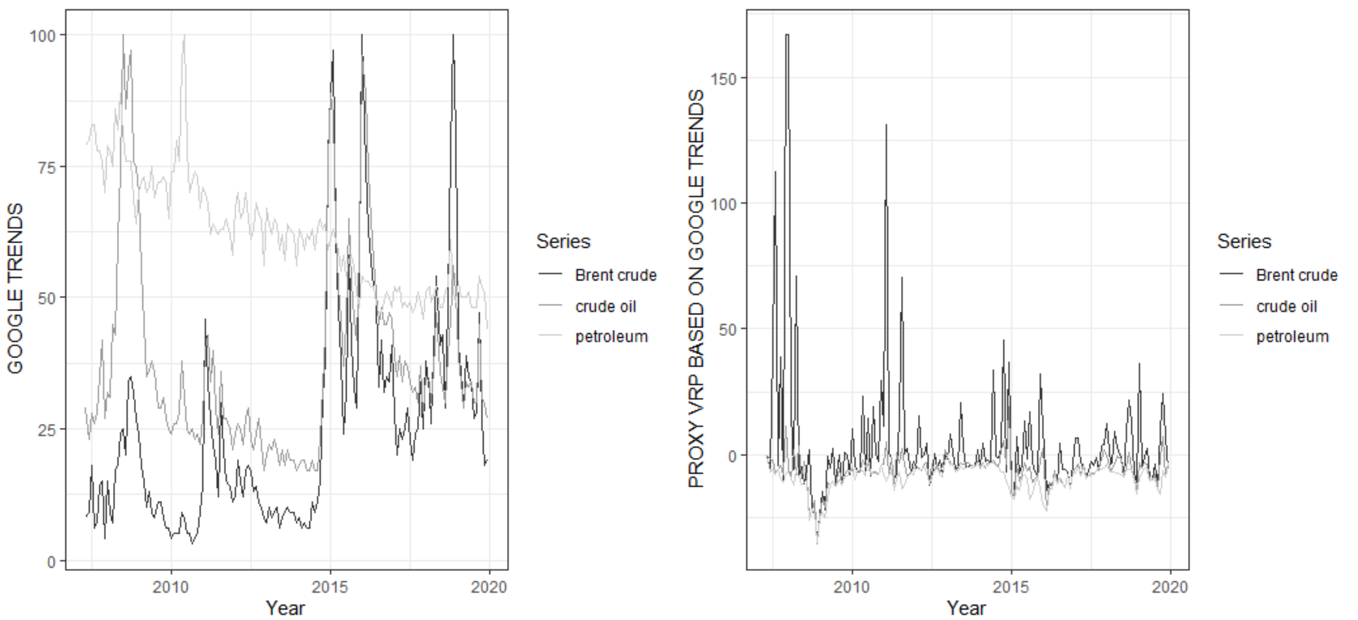

Fig. 6. Google Trends data (left) and proxy VRP based on Google Trends (right).

4.2. Out-of-sample forecasting

We computed forecasts for the real oil price at different forecast horizons (h=1, 6, 12, 24), and compared them by using the mean squared error (MSE) and the Mean Absolute Error (MAE) loss functions. Following Hansen et al. (2011), we then employed the Model Confidence Set (MCS) to select the best forecasting models at a $1-\alpha$ confidence level, with α = 10%. Following a common 67%/33% data split, the data in May 2007 – July 2015 were used as the first training sample for almost all models' estimation, while the data for August 2015 - December 2019 were left for out-of-sample forecasting using an expanding estimation window. The only exception was represented

[7] Google Trends considers the number of search queries for a topic or a keyword over a specific period and a specific region performed in Google, and creates a time-series reporting the relative popularity of the searched queries. More specifically, the amount of searches is divided by the total amount of searches for the same period and region, and the resulting time series is divided by its highest value and multiplied by 100. See https://support.google.com/trends for more details.

by the VAR(24) model proposed by Kilian and Murphy (2014), which employed the data from January 1973 till July 2015 as the first training sample.

Three classes of models were considered for a total of 25 models:

1. *Benchmark models* (5 models):
   - VAR(24) model by Kilian and Murphy (2014). Sample: 1973-2019;
   - VAR(p) model with the four variables used in the VAR model by Kilian and Murphy (2014) + VRP. Sample: 2007-2019;
   - VAR(p) model with the four variables used in the VAR model by Kilian and Murphy (2014) + proxy VRP with Google data. Sample: 2007-2019;
   - VAR(p) model with the four variables used in the VAR model by Kilian and Murphy (2014) + VRP + proxy VRP with Google data. Sample: 2007-2019;
   - No-change forecast.

2. *Penalized regressions* (12 models):
   - Ridge/LASSO/Elastic Net with the four variables used in the VAR(24) model by Kilian and Murphy (2014). Sample: 2007-2019;
   - Ridge/LASSO/Elastic Net with the four variables used in the VAR(24) model by Kilian and Murphy (2014) + VRP. Sample: 2007-2019;
   - Ridge/LASSO/Elastic Net with the four variables used in the VAR(24) model by Kilian and Murphy (2014) + proxy VRP with Google data. Sample: 2007-2019;
   - Ridge/LASSO/Elastic Net with the four variables used in the VAR(24) model by Kilian and Murphy (2014) + VRP + proxy VRP with Google data. Sample: 2007-2019.

3. *Principal Component Regression (PCR) and the Partial least squares (PLS) regression* (8 models):
   - PCR/PLS with the four variables used in the VAR(24) model by Kilian and Murphy (2014). Sample: 2007-2019;
   - PCR/PLS with the four variables used in the VAR(24) model by Kilian and Murphy (2014) + VRP. Sample: 2007-2019;
   - PCR/PLS with the four variables used in the VAR(24) model by Kilian and Murphy (2014) + proxy VRP with Google data. Sample: 2007-2019;
   - PCR/PLS with the four variables used in the VAR(24) model by Kilian and Murphy (2014) + VRP + proxy VRP with Google data. Sample: 2007-2019.

Additional models could surely be added, but this selection already gave important indications of whether predictors based on the implied volatility and Google search data are useful for forecasting the real oil price.

A summary of the models' performances according to the mean squared error (MSE) and the mean absolute error (MAE) for 1-month ahead, 6-month ahead, 12-month ahead, and 24-month ahead forecasts, and the potential inclusion in the model confidence set (MCS) are reported in Tables 1-4, respectively.

Table 1: Models' performances according to the mean squared error (MSE) and the mean absolute error (MAE) for ***1-month ahead forecasts***, as well as the inclusion in the Model Confidence Set. The smallest values are reported in bold font. KM means that the four variables originally considered by Kilian and Murphy (2014) were used in the model estimation.

MODELS	**MAE**	**MSE**	**MCS(MAE)**	**MCS(MSE)**
Random Walk (no-change forecast)	3.25	18.98	INCLUDED	INCLUDED
VAR (KM, 1973-2019)	3.55	19.40	INCLUDED	INCLUDED
VAR (KM + VRP, 2007-2019)	3.33	**18.60**	INCLUDED	INCLUDED
VAR (KM + GOOGLE, 2007-2019)	3.40	18.82	INCLUDED	INCLUDED
VAR (KM + VRP + GOOGLE, 2007-2019)	3.72	34.49	INCLUDED	INCLUDED
PCR (KM, 2007-2019)	5.18	50.35	INCLUDED	INCLUDED
PLS (KM, 2007-2019)	6.99	82.36		INCLUDED
RIDGE (KM, 2007-2019)	9.03	177.04		INCLUDED
LASSO (KM, 2007-2019)	3.31	20.03	INCLUDED	INCLUDED
ELASTIC NET (KM, 2007-2019)	3.61	22.45	INCLUDED	INCLUDED
PCR (KM + VRP, 2007-2019)	9.35	138.89		
PLS (KM + VRP, 2007-2019)	12.43	259.85		
RIDGE (KM + VRP, 2007-2019)	8.38	146.14	INCLUDED	INCLUDED
LASSO (KM + VRP, 2007-2019)	**3.24**	19.81	INCLUDED	INCLUDED
ELASTIC NET (KM + VRP, 2007-2019)	3.71	24.13	INCLUDED	INCLUDED
PCR (KM + GOOGLE, 2007-2019)	76.07	13754.97		
PLS (KM + GOOGLE, 2007-2019)	57.83	7118.25		
RIDGE (KM + GOOGLE, 2007-2019)	8.69	162.40		INCLUDED
LASSO (KM + GOOGLE, 2007-2019)	3.49	21.12	INCLUDED	INCLUDED
ELASTIC NET (KM + GOOGLE, 2007-2019)	3.79	25.13	INCLUDED	INCLUDED
PCR (KM + VRP + GOOGLE, 2007-2019)	14.15	333.42		
PLS (KM + VRP + GOOGLE, 2007-2019)	13.58	315.49		
RIDGE (KM + VRP + GOOGLE, 2007-2019)	7.74	133.61	INCLUDED	INCLUDED
LASSO (KM + VRP + GOOGLE, 2007-2019)	3.42	20.70	INCLUDED	INCLUDED
ELASTIC NET (KM + VRP + GOOGLE, 2007-2019)	3.75	24.06	INCLUDED	INCLUDED

Table 2: Models' performances according to the mean squared error (MSE) and the mean absolute error (MAE) for ***6-month ahead forecasts***, as well as the inclusion in the Model Confidence Set. The smallest values are reported in bold font. KM means that the four variables originally considered by Kilian and Murphy (2014) were used in the model estimation

MODELS	**MAE**	**MSE**	**MCS(MAE)**	**MCS(MSE)**
Random Walk (no-change forecast)	8.65	107.35	INCLUDED	INCLUDED
VAR (KM, 1973-2019)	10.77	172.87	INCLUDED	INCLUDED
VAR (KM + VRP, 2007-2019)	8.61	103.63	INCLUDED	INCLUDED
VAR (KM + GOOGLE, 2007-2019)	8.73	105.91	INCLUDED	INCLUDED
VAR (KM + VRP + GOOGLE, 2007-2019)	8.65	103.78	INCLUDED	INCLUDED
PCR (KM, 2007-2019)	21.37	623.53		INCLUDED
PLS (KM, 2007-2019)	27.19	1095.14		
RIDGE (KM, 2007-2019)	10.71	241.15	INCLUDED	INCLUDED
LASSO (KM, 2007-2019)	8.84	129.22	INCLUDED	INCLUDED
ELASTIC NET (KM, 2007-2019)	8.60	122.58	INCLUDED	INCLUDED
PCR (KM + VRP, 2007-2019)	17.67	424.86		INCLUDED
PLS (KM + VRP, 2007-2019)	37.83	5060.37		INCLUDED
RIDGE (KM + VRP, 2007-2019)	10.34	202.77	INCLUDED	INCLUDED
LASSO (KM + VRP, 2007-2019)	7.89	96.04	INCLUDED	INCLUDED
ELASTIC NET (KM + VRP, 2007-2019)	**7.56**	**81.72**	INCLUDED	INCLUDED
PCR (KM + GOOGLE, 2007-2019)	91.21	20453.62		
PLS (KM + GOOGLE, 2007-2019)	64.73	9549.37		
RIDGE (KM + GOOGLE, 2007-2019)	10.22	235.78	INCLUDED	INCLUDED
LASSO (KM + GOOGLE, 2007-2019)	9.73	162.92	INCLUDED	INCLUDED
ELASTIC NET (KM + GOOGLE, 2007-2019)	9.22	151.46	INCLUDED	INCLUDED
PCR (KM + VRP + GOOGLE, 2007-2019)	23.42	1046.20		INCLUDED
PLS (KM + VRP + GOOGLE, 2007-2019)	22.48	912.92		INCLUDED
RIDGE (KM + VRP + GOOGLE, 2007-2019)	10.26	207.82	INCLUDED	INCLUDED
LASSO (KM + VRP + GOOGLE, 2007-2019)	9.38	145.33	INCLUDED	INCLUDED
ELASTIC NET (KM + VRP + GOOGLE, 2007-2019)	9.17	138.17	INCLUDED	INCLUDED

Table 3: Models' performances according to the mean squared error (MSE) and the mean absolute error (MAE) for ***12-month ahead forecasts***, as well as the inclusion in the Model Confidence Set. The smallest values are reported in bold font. KM means that the four variables originally considered by Kilian and Murphy (2014) were used in the model estimation.

MODELS	**MAE**	**MSE**	**MCS(MAE)**	**MCS(MSE)**
Random Walk (no-change forecast)	**9.99**	**139.57**	INCLUDED	INCLUDED
VAR (KM, 1973-2019)	16.91	360.28	INCLUDED	INCLUDED
VAR (KM + VRP, 2007-2019)	10.23	157.59	INCLUDED	INCLUDED
VAR (KM + GOOGLE, 2007-2019)	10.47	160.15	INCLUDED	INCLUDED
VAR (KM + VRP + GOOGLE, 2007-2019)	10.52	174.88	INCLUDED	INCLUDED
PCR (KM, 2007-2019)	16.00	344.92	INCLUDED	INCLUDED
PLS (KM, 2007-2019)	19.39	512.58		INCLUDED
RIDGE (KM, 2007-2019)	11.48	223.69	INCLUDED	INCLUDED
LASSO (KM, 2007-2019)	13.72	242.10	INCLUDED	INCLUDED
ELASTIC NET (KM, 2007-2019)	12.93	222.45	INCLUDED	INCLUDED
PCR (KM + VRP, 2007-2019)	26.12	1456.01		INCLUDED
PLS (KM + VRP, 2007-2019)	34.12	2328.52		INCLUDED
RIDGE (KM + VRP, 2007-2019)	10.87	212.47	INCLUDED	INCLUDED
LASSO (KM + VRP, 2007-2019)	14.40	267.73	INCLUDED	INCLUDED
ELASTIC NET (KM + VRP, 2007-2019)	14.46	272.04	INCLUDED	INCLUDED
PCR (KM + GOOGLE, 2007-2019)	67.80	8913.46		
PLS (KM + GOOGLE, 2007-2019)	52.85	5849.85		
RIDGE (KM + GOOGLE, 2007-2019)	11.56	209.95	INCLUDED	INCLUDED
LASSO (KM + GOOGLE, 2007-2019)	13.19	228.80	INCLUDED	INCLUDED
ELASTIC NET (KM + GOOGLE, 2007-2019)	13.46	239.84	INCLUDED	INCLUDED
PCR (KM + VRP + GOOGLE, 2007-2019)	14.66	363.41	INCLUDED	INCLUDED
PLS (KM + VRP + GOOGLE, 2007-2019)	13.82	315.86	INCLUDED	INCLUDED
RIDGE (KM + VRP + GOOGLE, 2007-2019)	11.11	207.29	INCLUDED	INCLUDED
LASSO (KM + VRP + GOOGLE, 2007-2019)	14.73	275.70	INCLUDED	INCLUDED
ELASTIC NET (KM + VRP + GOOGLE, 2007-2019)	14.25	265.60	INCLUDED	INCLUDED

Table 4: Models' performances according to the mean squared error (MSE) and the mean absolute error (MAE) for ***24-month ahead forecasts***, as well as the inclusion in the Model Confidence Set. The smallest values are reported in bold font. KM means that the four variables originally considered by Kilian and Murphy (2014) were used in the model estimation.

MODELS	**MAE**	**MSE**	**MCS(MAE)**	**MCS(MSE)**
Random Walk (no-change forecast)	14.31	289.05	INCLUDED	INCLUDED
VAR (KM, 1973-2019)	22.32	581.09		
VAR (KM + VRP, 2007-2019)	14.60	352.55	INCLUDED	INCLUDED
VAR (KM + GOOGLE, 2007-2019)	14.69	357.63	INCLUDED	INCLUDED
VAR (KM + VRP + GOOGLE, 2007-2019)	13.52	302.72	INCLUDED	INCLUDED
PCR (KM, 2007-2019)	22.29	620.29	INCLUDED	INCLUDED
PLS (KM, 2007-2019)	21.64	614.87	INCLUDED	INCLUDED
RIDGE (KM, 2007-2019)	**6.88**	**103.69**	INCLUDED	INCLUDED
LASSO (KM, 2007-2019)	15.88	347.20	INCLUDED	INCLUDED
ELASTIC NET (KM, 2007-2019)	15.92	343.15	INCLUDED	INCLUDED
PCR (KM + VRP, 2007-2019)	28.35	867.14		
PLS (KM + VRP, 2007-2019)	31.42	1099.40		
RIDGE (KM + VRP, 2007-2019)	9.01	131.37	INCLUDED	INCLUDED
LASSO (KM + VRP, 2007-2019)	15.46	329.55	INCLUDED	INCLUDED
ELASTIC NET (KM + VRP, 2007-2019)	15.59	334.67	INCLUDED	INCLUDED
PCR (KM + GOOGLE, 2007-2019)	31.17	1283.35		
PLS (KM + GOOGLE, 2007-2019)	30.83	1262.67		
RIDGE (KM + GOOGLE, 2007-2019)	7.26	107.84	INCLUDED	INCLUDED
LASSO (KM + GOOGLE, 2007-2019)	14.70	303.93	INCLUDED	INCLUDED
ELASTIC NET (KM + GOOGLE, 2007-2019)	14.40	293.32	INCLUDED	INCLUDED
PCR (KM + VRP + GOOGLE, 2007-2019)	31.99	1142.86		
PLS (KM + VRP + GOOGLE, 2007-2019)	30.87	1091.59		
RIDGE (KM + VRP + GOOGLE, 2007-2019)	8.96	129.71	INCLUDED	INCLUDED
LASSO (KM + VRP + GOOGLE, 2007-2019)	14.65	319.14	INCLUDED	INCLUDED
ELASTIC NET (KM + VRP + GOOGLE, 2007-2019)	14.28	293.93	INCLUDED	INCLUDED

In general, penalized regressions provided the best forecasting performances across most of the forecasting horizons, thus confirming the empirical evidence reported by Miao et al. (2017) and Zhang et al. (2019). In the case of forecasts up to 6-month ahead, the penalized regressions employing the four variables originally considered by Kilian and Murphy (2014) together with the VRP provided the lowest MAE and MSE. However, the other model specifications including the proxy VRP with Google data performed very similarly, and they were also included in the model confidence set together with the VAR models and the no-change forecast. Instead, the PCR and PLS models performed very poorly, and they were rarely included in the MCS for all forecasts horizons: most likely, the presence of noisy data such as Google data and proxy variables for the oil market was not appropriate for these models and created computational problems. In the case of forecasts up to 12-month ahead, the no-change forecast was the best, thus confirming its past prowess, followed closely by the VAR models with the VRP and Google data, and by the Ridge penalized regression models. In the case of longer-term forecasts up to 24-month ahead, the Ridge model employing the four variables originally considered by Kilian and Murphy (2014) was the best, followed by other penalized regression models including the proxy VRP with Google search data. Interestingly, we noted that models using the VRP as an additional predictor performed pretty well for forecasts up to 6-12 months ahead forecasts. Instead, models using Google data as an additional predictor performed better for longer-term forecasts up to 12-24 months ahead, thus confirming similar evidence reported in Fantazzini and Fomichev (2014) for oil price forecasting, and in Fantazzini and Toktamysova (2015) for car sales forecasting. The original model by Kilian and Murphy (2014) using the full sample starting in 1973 performed well for short-term forecasts, but it became less competitive with longer-term horizons and, in the case of 24-month ahead forecasts, it was even excluded from the model confidence set.

5. Robustness checks

We wanted to verify that our previous results also held with different forecasting models and regressors. Therefore, we performed a series of robustness checks considering a wider set of influential variables, a Hierarchical Vector Auto-Regression (HVAR) model estimated with the Least Absolute Shrinkage and Selection Operator (LASSO) as proposed by Nicholson et al. (2020) and Wilms et al. (2021), and a set of forecasting models using a simplified specification for Google Trends data.

5.1. A wider set of predictors

Following Miao et al. (2017), we estimated our models using a wider set of predictors compared to the baseline case. More specifically, we added the Baltic Dirty Tanker Index as an additional demand factor, the capacity utilization rate and the global crude exports as additional supply factors, and the US Federal Funds rate as an additional financial factor.

The Baltic (Exchange) Dirty Tanker Index is published by the Baltic Exchange in London and is an important price index for the worldwide shipping of oil. This index takes into account 17 shipping routes and records the costs for time charter for four ship classes (VLCC, Suezmax, Aframax, and Panamax). The refinery utilization rates (obtained from the US Energy Information Administration) are believed to have a negative relation to the oil price: as discussed by Kaufmann et al. (2008), increasing rates of refinery utilization forces refiners to buy crudes that are less well suited to their refineries. As a consequence, this reduces yield and decreases the value of the products they produce, thus reducing the price they are willing to pay for crude oil. Moreover, as refineries reach full capacity, the demand for crude oil drops, and oil prices fall. The global crude oil exports were obtained from the JODI-Oil Database as another supply indicator, and it usually expected that the higher are oil exports, the lower is the oil price. Finally, the US Federal Funds Rate can be considered an important financial indicator that affects other interest rates such as those of loans and mortgages, and which can indirectly influence the confidence of the economic agents. We refer to Miao et al. (2017) and references therein for a large discussion about the main influential factors in crude oil price forecasting. A summary of the augmented models' performances according to the mean squared error (MSE) and the mean absolute error (MAE) for 1-month ahead, 6-month ahead, 12-month ahead, and 24-month ahead forecasts, and the potential inclusion in the model confidence set (MCS) are reported in Tables 5-8 in the Appendix.

The results are quite similar to the baseline case, with only some marginal improvements in terms of MAE/MSE for long-term forecasts up to 12-24 months ahead. Moreover, all the models included in the MCS in the baseline case were also included in the MCS in this robustness check. Therefore, it is possible to say that adding these new predictors did not statistically change the models' performances that we observed in the baseline case with fewer variables.

5.2. A hierarchical VAR model with LASSO

Our baseline case considered only univariate models with regularization. Unfortunately, it is well known that VAR models can quickly have a wealth of parameters, making them extremely difficult if not impossible to estimate, depending on the sample size. Therefore, we wanted to check how our previous results changed with a multivariate model able to both accommodate a large number of regressors and improve the model estimation and its forecasting performances. To achieve this goal, we employed the *Hierarchical Vector Autoregression* (HVAR) model estimated with the Least Absolute Shrinkage and Selection Operator (LASSO) proposed by Nicholson et al. (2020) and Wilms et al. (2021), which is a special case of a multivariate penalized least squares optimization problem. Let us consider again the previous VAR(24) process,

$$Y_t = \boldsymbol{v} + \sum_{l=1}^{24} \Phi^l Y_{t-l} + \boldsymbol{u}_t, \quad \boldsymbol{u}_t \sim WN(0, \Sigma),$$

where Y_t is a ($n \times 1$)-vector containing the endogenous variables, $\boldsymbol{v}$ is an intercept vector, while Φ^l are the usual coefficient matrices. The HVAR approach proposed adds structured convex penalties to the least-squares VAR problem, so that the optimization problem is given by

$$\min_{\boldsymbol{v},\Phi} \sum_{t=1}^{T} \left\| Y_t - \boldsymbol{v} - \sum_{l=1}^{24} \Phi^l Y_{t-l} \right\|_F^2 + \lambda \left(\mathcal{P}_Y(\Phi) \right),$$

where $\|A\|_F$ denotes the Frobenius norm of matrix *A* (that is, the elementwise 2-norm), $\lambda \geq 0$ is a penalty parameter, while $\mathcal{P}_Y(\Phi)$ is the group penalty structure on the endogenous coefficient matrices. The HVAR class of models solves the problem of an increasing maximum lag order by including the lag order into hierarchical group LASSO penalties, which induce sparsity and a low maximum lag order. For our empirical work, we employed the *elementwise penalty function*

$$\mathcal{P}_Y(\Phi) = \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{l=1}^{24} \left\| \Phi_{ij}^l \right\|_2$$

which is the most general structure, because every variable in every equation is allowed to have its own maximum lag resulting in n^2 possible lag orders. The penalty parameter λ is estimated by sequential cross-validation, see Nicholson et al. (2020) for the full details.

A summary of the HVAR models' performances according to the mean squared error (MSE) and the mean absolute error (MAE) for 1-month ahead, 6-month ahead, 12-month ahead, and 24-month ahead forecasts, and the potential inclusion in the model confidence set (MCS) are reported in Tables 9-12.

The results are somewhat mixed: the HVAR computed using only the four variables considered by Kilian and Murphy (2014) with the full sample starting in 1973 performed much better than the corresponding VAR model in the baseline case, particularly for forecasts up to 6-24 months ahead. Instead, the HVAR models that used a larger number of regressors (and smaller samples) performed mostly worse than the corresponding VAR models in the baseline case and in the previous robustness check. Therefore, it appears that the higher efficiency and computational simplicity of univariate penalized regressions have to be preferred compared to multivariate penalized models, thus confirming similar empirical evidence reported by Miao et al. (2017) and Zhang et al. (2019). Besides, all the models included in the MCS in the baseline case were again included in the MCS in this robustness check, thus highlighting that these forecasting improvements were not statistically significant, given this sample data.

5.3. A simplified specification for Google data

In our baseline case, we considered a predictor mimicking the traditional variance risk premium, where we used the squared log-returns of the monthly Google Trends data in place of the implied volatility to try to measure the sentiment of retail investors. Unfortunately, this type of predictor can further increase the noisy nature of Google data. Therefore, we wanted to check how our previous results changed when using the original Google Trends data in place of the proxy VRP that we considered in the baseline case.

A summary of the performances for a selected group of modified models according to the mean squared error (MSE) and the mean absolute error (MAE) for 1-month ahead, 6-month ahead, 12-month ahead, and 24-month ahead forecasts, and the potential inclusion in the model confidence set (MCS) are reported in Tables 13-16.

The performance of the penalized regression models turned out to be similar to the baseline case or only slightly better. Instead, PCR and PLS models significantly improved their forecasting performances, even though they remained inferior to penalized regression models. This outcome was expected because the use of Google data as a simple linear predictor instead of the proxy VRP built with their monthly squared log returns allowed for an easier and more efficient estimation of PCR/PLS models. Finally, we remark again that all the models included in the MCS in the baseline case were also included in the MCS in this robustness check.

6. Conclusions

This paper investigated whether augmenting models with two predictors based on the variance risk premium and Google search data improves the quality of the forecasts for real oil prices. To reach this objective, we first computed the VRP by using the CBOE Crude Oil Volatility Index as a proxy for the implied volatility and the monthly realized volatility. We also built a second predictor mimicking the variance risk premium, where we used the squared log-returns of the monthly Google Trends data in place of the implied volatility to build a proxy VRP to measure the sentiment of retail investors. Secondly, we proposed a set of univariate penalized regression models that included the main predictors proposed by the past literature devoted to oil price forecasting, together with our predictors based on the implied volatility and Google search data. Thirdly, we performed a forecasting exercise with different forecast horizons up to 24-month ahead and different competing model specifications.

We found that penalized regressions provided the best forecasting performances across most of the forecasting horizons, thus confirming the empirical evidence reported by Miao et al. (2017) and Zhang et al. (2019). Interestingly, we noted that models using the VRP as an additional predictor performed pretty well for forecasts up to 6-12 months ahead forecasts, while models using Google data as an additional predictor performed better for longer-term forecasts up to 12-24 months ahead, thus confirming similar evidence reported in Fantazzini and Fomichev (2014) and Fantazzini and Toktamysova (2015). The original model by Kilian and Murphy (2014) using the full sample starting in 1973 performed well for short-term forecasts, but it became less competitive with longer-term horizons and, in the case of 24-month ahead forecasts, it was even excluded from the model confidence set. The no-change forecast performed well across all forecast horizons, whereas the PCR and PLS models performed very poorly across all forecasting horizons. However, we remark that the differences in forecasting performances were not statistically different for most models, and only the Principal Component Regression and the Partial Least Squares models were consistently excluded from the set of best forecasting models across all horizons.

Finally, we performed a set of robustness checks to verify that our results also held with different model specifications considering a wider set of influential variables, a Hierarchical Vector Auto-Regression model estimated with the LASSO, and a set of forecasting models using a simplified specification for Google Trends data. We found that the results were quite similar to the baseline case, with only some marginal improvements when using the original Google Trends data in place of the proxy VRP that we considered in the baseline case, and with univariate penalized regressions generally performing better than multivariate penalized models. Moreover, all the models included in the MCS in the baseline case were also included in the MCS in all robustness checks. Even though the forecasting differences were not statistically different for most models, we think that these results are mainly due to the small forecasting samples involved, so that this empirical evidence can still be useful for financial professionals and researchers alike[8].

The general recommendation that emerged from our analysis is to choose univariate penalized regression models with a limited set of (lagged) predictors, preferably those originally proposed by Kilian and Murphy (2014) augmented with the VRP and/or Google search data, depending on the forecasting horizon of interest.

We remark that data from 2020 onwards were not considered in this work because the Covid-19 pandemic represented a major structural break in the oil market and would require separate modelling. Therefore, we leave this issue as an avenue for further research. Another possibility for future work will be to consider forecast combination methods, following the ideas discussed by Clemen (1989), Timmermann (2006), Hsiao and Wan (2014), and Hyndman and Athanasopoulos (2018).

Acknowledgments. The authors gratefully acknowledge financial support from the grant of the Russian Science Foundation no. 20-68-47030.

[8] In this regard, we note that large scale simulation evidence reported by Hansen et al. (2011) showed that "*it takes about 500 observations to remove all the poor models*", (Hansen et al. 2011, p. 479). Given that our forecasting samples ranged from 30 up to 53 observations (depending on the forecasting horizon), it was not a surprise that many models were included into the MCS despite showing much worse MAE/MSE.

References

Afkhami M., Cormack L. Ghoddusi H. (2017). Google search keywords that best predict energy price volatility. *Energy Economics*, 67, 17–27. DOI: 10.1016/j.eneco.2017.07.014.

Alquist R., Kilian L., Vigfusson R. J. (2013). Forecasting the price of oil. In: *Handbook of economic forecasting*, 2, 427–507. DOI: 10.1016/B978-0-444-53683-9.00008-6.

Atmaz A. (2022). Stock return extrapolation, option prices, and variance risk premium. *The Review of Financial Studies*, 35(3), 1348–1393. DOI: 10.1093/rfs/hhab051.

Bhattacharyya S. C. (2019). *Energy economics: concepts, issues, markets and governance.* Springer, London. DOI: 10.1007/978-1-4471-7468-4.

Basistha A., Kurov A., Wolfe M. H. (2020). Volatility forecasting: the role of internet search activity and implied volatility. *Journal of Risk Model Validation*, 14(1), 35–63. DOI: 10.2139/ssrn.2812387.

Baumeister C., Kilian L. (2012). Real-Time Forecasts of the Real Price of Oil. *Journal of Business and Economic Statistics*, 30, 326–336. DOI: 10.1080/07350015.2011.648859.

Baumeister C., Kilian L. (2015). Forecasting the real price of oil in a changing world: a forecast combination approach. *Journal of Business and Economic Statistics*, 33(3), 338–351. DOI: 10.1080/07350015.2014.949342.

Baumeister C., Hamilton J. D. (2019). Structural interpretation of vector autoregressions with incomplete identification: Revisiting the role of oil supply and demand shocks. *American Economic Review*, 109(5), 1873–1910. DOI: 10.1257/aer.20151569.

Bollerslev T., Tauchen G., Zhou H. (2009). Expected Stock Returns and Variance Risk Premia. *Review Finance Study*, 22(11), 4463–4492. DOI: 10.1093/rfs/hhp008.

Bollerslev T., Marrone J., Xu L., Zhou H. (2014). Stock return predictability and variance risk premia: statistical inference and international evidence. *Journal of Financial and Quantitative Analysis*, 49(3), 633–661. DOI: 10.1017/S0022109014000453.

Bazhenov T., Fantazzini D. (2019). Forecasting realized volatility of Russian stocks using Google Trends and implied volatility. *Russian Journal of Industrial Economics*, 12(1), 79–88.

Chevallier J., Sévi B. (2014). A fear index to predict oil futures returns. *Energy Studies Review*, 20(3), 1–17. DOI: 10.15173/esr.v20i3.552.

Clemen R. T. (1989). Combining forecasts: A review and annotated bibliography. *International Journal of Forecasting*, 5 (4), 559–583. DOI: 10.1016/0169-2070(89)90012-5.

Coleman L. (2011). Explaining crude oil prices using fundamental measures. *Energy Policy*, 40, 318–324. DOI: 10.1016/j.enpol.2011.10.012.

Fantazzini D., Höök M., Angelantoni A. (2011). Global oil risks in the early 21st century. *Energy Policy*, 39(12), 7865–7873. DOI: 10.1016/j.enpol.2011.09.035.

Fantazzini D. (2016). The oil price crash in 2014/15: Was there a (negative) financial bubble? *Energy Policy*, 96, 383–396. DOI: 10.1016/j.enpol.2016.06.020.

Fantazzini D. Fomichev N. (2014). Forecasting the real price of oil using online search data. *International Journal of Computational Economics and Econometrics*, 4(1-2), 4–31. DOI: 10.1504/IJCEE.2014.060284.

Fantazzini D., Shangina T. (2019). The importance of being informed: forecasting market risk measures for the Russian RTS index future using online data and implied volatility over two decades. *Applied Econometrics*, 3 (55), 5–31. DOI: 10.244111/1993-7601-2019-10008.

Fantazzini D., Toktamysova Z. (2015). Forecasting German car sales using Google data and multivariate models. *International Journal of Production Economics*, 170, 97–135. DOI: 10.1016/j.ijpe.2015.09.010.

Gospodinov N., Herrera A.M., Pesavento E. (2013). Unit roots, cointegration, and pretesting in VAR models. *Advances in Econometrics*, 32, 81–115. DOI: 10.1108/S0731-9053(2013)0000031003.

Hamilton J. (2008). Oil and the macroeconomy. In: *The New Palgrave Dictionary of Economics*, Palgrave MacMillan, London. DOI: 10.1057/978-1-349-95121-5_2119-1.

Hamilton J. D. (2009). Understanding crude oil prices. *Energy Journal,* 30, 179–206. DOI: 10.5547/ISSN0195-6574-EJ-Vol30-No2-9.

Hamilton J.D. (2011). Nonlinearities and the macroeconomic effects of oil prices. *Macroeconomic dynamics*, 15(S3), 364–378. DOI: 10.1017/S1365100511000307.

Hamilton J.D. (2013). Historical oil shocks. In: *The Routledge Handbook of Major Events in Economic History*, 258–284, Routledge Taylor and Francis Group, New York.

Hastie T., Tibshirani R., Friedman J. H., Friedman J. H. (2009). *The elements of statistical learning: data mining, inference, and prediction*. Springer, New York. DOI: 10.1007/978-0-387-21606-5.

Hsiao C., Wan S. K. (2014). Is there an optimal forecast combination? *Journal of Econometrics*, 178, 294–309. DOI: 10.1016/j.jeconom.2013.11.003.

Hyndman R. J., Athanasopoulos G. (2018). *Forecasting: Principles and practice*. OTexts.

Kaufmann R. K., Dees S., Gasteuil A., Mann M. (2008). Oil prices: The role of refinery utilization, futures markets and non-linearities. *Energy Economics*, 30(5), 2609–2622. DOI: 10.1016/j.eneco.2008.04.010.

Kilian L. (2008). The economic effects of energy price shocks. *Journal of economic literature*, 46(4), 871–909. DOI: 10.1257/jel.46.4.871.

Kilian L. (2009). Not all oil price shocks are alike: Disentangling demand and supply shocks in the crude oil market. *American Economic Review*, 99(3), 1053–1069. DOI: 10.1257/aer.99.3.1053.

Kilian L. (2016). The impact of the shale oil revolution on US oil and gasoline prices. *Review of Environmental Economics and Policy*, 10(2), 185–205. DOI:10.1093/reep/rew001.

Kilian L. Murphy D. (2014). The role of inventories and speculative trading in the global market for crude oil. *Journal of Applied Econometrics*, 29(3), 454–478. DOI: 10.1002/jae.2322.

Kilian L., Zhou X. (2022). The econometrics of oil market VAR models. *Advances in Econometrics*, forthcoming. DOI: 10.1016/j.eneco.2022.105973.

Londono J. M., Zhou H. (2017). Variance risk premiums and the forward premium puzzle. *Journal of Financial Economics*, 124(2), 415–440. DOI: 10.1016/j.jfineco.2017.02.002.

Miao H., Ramchander S., Wang T. Yang D. (2017). Influential factors in crude oil forecasting. *Energy Economics*, 68, 77–88. DOI: 10.1016/j.eneco.2017.09.010.

Neely C. J., Rapach D. E., Tu J., Zhou G. (2014). Forecasting the equity risk premium: the role of technical indicators. *Management science*, 60(7), 1772–1791. DOI: 10.1287/mnsc.2013.1838.

Nicholson W. B., Wilms I., Bien J., Matteson D. S. (2020). High Dimensional Forecasting via Interpretable Vector Autoregression. *Journal of Machine Learning Research*, 21(166), 1–52.

Ornelas J. R. H. (2019). Expected currency returns and volatility risk premia. *The North American Journal of Economics and Finance*, 49, 206–234. DOI: 10.1016/j.najef.2019.03.015.

Ornelas J. R. H., Mauad R. B. (2019). Volatility risk premia and future commodity returns. *Journal of International Money and Finance*, 96, 341–360. DOI: 10.1016/j.jimonfin.2017.07.008.

Qadan M., Nama H. (2018). Investor sentiment and the price of oil. *Energy Economics*, 69, 42–58. DOI: 10.1016/j.eneco.2017.10.035.

Quayyoum S., Khan M. H., Shah S. Z., Simonetti B., Matarazzo M. (2019). Seasonality in crude oil returns. *Soft Computing*, 24, 13547–13556. DOI: 10.1007/s00500-019-04329-0.

Schwarz P. (2017). Energy and sustainability. In: *The Proceedings of Energy and sustainability*, WIT Press. DOI: 10.4324/9781315114064-15.

Timmermann A. (2006). Forecast combinations. *Handbook of Economic Forecasting*, 1, 135–196. DOI: 10.2139/ssrn.1609530.

Vinzi V. E., Chin W. W., Henseler J., Wang H. (2010). *Handbook of partial least squares*. Springer, Berlin. DOI: 10.1007/978-3-540-32827-8.

Wilms I., Basu S., Bien J., Matteson D. S. (2021). Sparse identification and estimation of large-scale vector autoregressive moving averages. *Journal of the American Statistical Association*, 1–12. DOI: 10.1080/01621459.2021.1942013.

Zhang Y., Ma F., Wang Y. (2019). Forecasting crude oil prices with a large set of predictors: Can LASSO select powerful predictors? *Journal of Empirical Finance*, 54, 97–117. DOI: 10.1016/j.jempfin.2019.08.007.

Zou H., Hastie T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 67(2), 301–320. DOI: 10.1111/j.1467-9868.2005.00503.x.

APPENDIX

Table 5: Augmented models' performances according to the mean squared error (MSE) and the mean absolute error (MAE) for ***1-month ahead forecasts***, as well as the inclusion in the Model Confidence Set. The smallest values are reported in bold font. All models were augmented with 4 additional variables: Baltic Dirty Tanker Index, capacity utilization rate, global crude exports, and the US Federal Funds rate

MODELS	MAE	MSE	MCS(MAE)	MCS(MSE)
BASELINE MODES				
Random Walk (no-change forecast)	3.25	18.98	INCLUDED	INCLUDED
VAR (KM, 1973-2019)	3.55	19.40	INCLUDED	INCLUDED
VAR (KM + VRP, 2007-2019)	3.33	**18.60**	INCLUDED	INCLUDED
VAR (KM + GOOGLE, 2007-2019)	3.40	18.82	INCLUDED	INCLUDED
VAR (KM + VRP + GOOGLE, 2007-2019)	3.72	34.49	INCLUDED	INCLUDED
PCR (KM, 2007-2019)	5.18	50.35	INCLUDED	INCLUDED
PLS (KM, 2007-2019)	6.99	82.36		INCLUDED
RIDGE (KM, 2007-2019)	9.03	177.04		INCLUDED
LASSO (KM, 2007-2019)	3.31	20.03	INCLUDED	INCLUDED
ELASTIC NET (KM, 2007-2019)	3.61	22.45	INCLUDED	INCLUDED
PCR (KM + VRP, 2007-2019)	9.35	138.89		INCLUDED
PLS (KM + VRP, 2007-2019)	12.43	259.85		INCLUDED
RIDGE (KM + VRP, 2007-2019)	8.38	146.14	INCLUDED	INCLUDED
LASSO (KM + VRP, 2007-2019)	**3.24**	19.81	INCLUDED	INCLUDED
ELASTIC NET (KM + VRP, 2007-2019)	3.71	24.13	INCLUDED	INCLUDED
PCR (KM + GOOGLE, 2007-2019)	76.07	13754.97		
PLS (KM + GOOGLE, 2007-2019)	57.83	7118.25		
RIDGE (KM + GOOGLE, 2007-2019)	8.69	162.40	INCLUDED	INCLUDED
LASSO (KM + GOOGLE, 2007-2019)	3.49	21.12	INCLUDED	INCLUDED
ELASTIC NET (KM + GOOGLE, 2007-2019)	3.79	25.13	INCLUDED	INCLUDED
PCR (KM + VRP + GOOGLE, 2007-2019)	14.15	333.42		INCLUDED
PLS (KM + VRP + GOOGLE, 2007-2019)	13.58	315.49		INCLUDED
RIDGE (KM + VRP + GOOGLE, 2007-2019)	7.74	133.61	INCLUDED	INCLUDED
LASSO (KM + VRP + GOOGLE, 2007-2019)	3.42	20.70	INCLUDED	INCLUDED
ELASTIC NET (KM + VRP + GOOGLE, 2007-2019)	3.75	24.06	INCLUDED	INCLUDED
ADDITIONAL MODES				
VAR (KM + 4 variables, 2007-2019)	3.56	20.66	INCLUDED	INCLUDED
PCR (KM + 4 variables, 2007-2019)	22.90	905.46		INCLUDED
PLS (KM + 4 variables, 2007-2019)	23.07	1003.25		INCLUDED
RIDGE (KM + 4 variables, 2007-2019)	7.70	134.70	INCLUDED	INCLUDED
LASSO (KM + 4 variables, 2007-2019)	3.55	20.57	INCLUDED	INCLUDED
ELASTIC NET (KM + 4 variables, 2007-2019)	3.95	23.03	INCLUDED	INCLUDED
PCR (KM + VRP + GOOGLE + 4 v., 2007-2019)	22.08	880.58		INCLUDED
PLS (KM + VRP + GOOGLE + 4 v., 2007-2019)	20.76	695.26		INCLUDED
RIDGE (KM + VRP + GOOGLE + 4 v., 2007-2019)	6.87	111.46	INCLUDED	INCLUDED
LASSO (KM + VRP + GOOGLE + 4 v., 2007-2019)	3.61	21.10	INCLUDED	INCLUDED
ELASTIC NET (KM+VRP+ GOOGLE+4 v., 2007-2019)	3.98	24.76	INCLUDED	INCLUDED

Table 6: Augmented models' performances according to the mean squared error (MSE) and the mean absolute error (MAE) for ***6-month ahead forecasts***, as well as the inclusion in the Model Confidence Set. The smallest values are reported in bold font. All models were augmented with 4 additional variables: Baltic Dirty Tanker Index, capacity utilization rate, global crude exports, and the US Federal Funds rate

MODELS	MAE	MSE	MCS(MAE)	MCS(MSE)
BASELINE MODES				
Random Walk (no-change forecast)	8.65	107.35	INCLUDED	INCLUDED
VAR (KM, 1973-2019)	10.77	172.87	INCLUDED	INCLUDED
VAR (KM + VRP, 2007-2019)	8.61	103.63	INCLUDED	INCLUDED
VAR (KM + GOOGLE, 2007-2019)	8.73	105.91	INCLUDED	INCLUDED
VAR (KM + VRP + GOOGLE, 2007-2019)	8.65	103.78	INCLUDED	INCLUDED
PCR (KM, 2007-2019)	21.37	623.53		INCLUDED
PLS (KM, 2007-2019)	27.19	1095.14		
RIDGE (KM, 2007-2019)	10.71	241.15	INCLUDED	INCLUDED
LASSO (KM, 2007-2019)	8.84	129.22	INCLUDED	INCLUDED
ELASTIC NET (KM, 2007-2019)	8.60	122.58	INCLUDED	INCLUDED
PCR (KM + VRP, 2007-2019)	17.67	424.86		INCLUDED
PLS (KM + VRP, 2007-2019)	37.83	5060.37		INCLUDED
RIDGE (KM + VRP, 2007-2019)	10.34	202.77	INCLUDED	INCLUDED
LASSO (KM + VRP, 2007-2019)	7.89	96.04	INCLUDED	INCLUDED
ELASTIC NET (KM + VRP, 2007-2019)	**7.56**	**81.72**	INCLUDED	INCLUDED
PCR (KM + GOOGLE, 2007-2019)	91.21	20453.62		
PLS (KM + GOOGLE, 2007-2019)	64.73	9549.37		
RIDGE (KM + GOOGLE, 2007-2019)	10.22	235.78	INCLUDED	INCLUDED
LASSO (KM + GOOGLE, 2007-2019)	9.73	162.92	INCLUDED	INCLUDED
ELASTIC NET (KM + GOOGLE, 2007-2019)	9.22	151.46	INCLUDED	INCLUDED
PCR (KM + VRP + GOOGLE, 2007-2019)	23.42	1046.20		
PLS (KM + VRP + GOOGLE, 2007-2019)	22.48	912.92		INCLUDED
RIDGE (KM + VRP + GOOGLE, 2007-2019)	10.26	207.82	INCLUDED	INCLUDED
LASSO (KM + VRP + GOOGLE, 2007-2019)	9.38	145.33	INCLUDED	INCLUDED
ELASTIC NET (KM + VRP + GOOGLE, 2007-2019)	9.17	138.17	INCLUDED	INCLUDED
ADDITIONAL MODES				
VAR (KM + 4 variables, 2007-2019)	10.07	151.45	INCLUDED	INCLUDED
PCR (KM + 4 variables, 2007-2019)	16.10	683.76	INCLUDED	INCLUDED
PLS (KM + 4 variables, 2007-2019)	15.60	601.59	INCLUDED	INCLUDED
RIDGE (KM + 4 variables, 2007-2019)	9.52	206.79	INCLUDED	INCLUDED
LASSO (KM + 4 variables, 2007-2019)	9.10	124.82	INCLUDED	INCLUDED
ELASTIC NET (KM + 4 variables, 2007-2019)	9.47	134.77	INCLUDED	INCLUDED
PCR (KM + VRP + GOOGLE + 4 v., 2007-2019)	15.79	662.09	INCLUDED	INCLUDED
PLS (KM + VRP + GOOGLE + 4 v., 2007-2019)	15.23	618.43	INCLUDED	INCLUDED
RIDGE (KM + VRP + GOOGLE + 4 v., 2007-2019)	9.36	194.64	INCLUDED	INCLUDED
LASSO (KM + VRP + GOOGLE + 4 v., 2007-2019)	9.66	154.81	INCLUDED	INCLUDED
ELASTIC NET (KM+VRP+ GOOGLE+4 v., 2007-2019)	8.75	115.00	INCLUDED	INCLUDED

Table 7: Augmented models' performances according to the mean squared error (MSE) and the mean absolute error (MAE) for ***12-month ahead forecasts***, as well as the inclusion in the Model Confidence Set. The smallest values are reported in bold font. All models were augmented with 4 additional variables: Baltic Dirty Tanker Index, capacity utilization rate, global crude exports, and the US Federal Funds rate

MODELS	**MAE**	**MSE**	**MCS(MAE)**	**MCS(MSE)**
BASELINE MODES				
Random Walk (no-change forecast)	9.99	**139.57**	INCLUDED	INCLUDED
VAR (KM, 1973-2019)	16.91	360.28	INCLUDED	INCLUDED
VAR (KM + VRP, 2007-2019)	10.23	157.59	INCLUDED	INCLUDED
VAR (KM + GOOGLE, 2007-2019)	10.47	160.15	INCLUDED	INCLUDED
VAR (KM + VRP + GOOGLE, 2007-2019)	10.52	174.88	INCLUDED	INCLUDED
PCR (KM, 2007-2019)	16.00	344.92	INCLUDED	INCLUDED
PLS (KM, 2007-2019)	19.39	512.58	INCLUDED	INCLUDED
RIDGE (KM, 2007-2019)	11.48	223.69	INCLUDED	INCLUDED
LASSO (KM, 2007-2019)	13.72	242.10	INCLUDED	INCLUDED
ELASTIC NET (KM, 2007-2019)	12.93	222.45	INCLUDED	INCLUDED
PCR (KM + VRP, 2007-2019)	26.12	1456.01	INCLUDED	INCLUDED
PLS (KM + VRP, 2007-2019)	34.12	2328.52		INCLUDED
RIDGE (KM + VRP, 2007-2019)	10.87	212.47	INCLUDED	INCLUDED
LASSO (KM + VRP, 2007-2019)	14.40	267.73	INCLUDED	INCLUDED
ELASTIC NET (KM + VRP, 2007-2019)	14.46	272.04	INCLUDED	INCLUDED
PCR (KM + GOOGLE, 2007-2019)	67.80	8913.46		
PLS (KM + GOOGLE, 2007-2019)	52.85	5849.85		
RIDGE (KM + GOOGLE, 2007-2019)	11.56	209.95	INCLUDED	INCLUDED
LASSO (KM + GOOGLE, 2007-2019)	13.19	228.80	INCLUDED	INCLUDED
ELASTIC NET (KM + GOOGLE, 2007-2019)	13.46	239.84	INCLUDED	INCLUDED
PCR (KM + VRP + GOOGLE, 2007-2019)	14.66	363.41	INCLUDED	INCLUDED
PLS (KM + VRP + GOOGLE, 2007-2019)	13.82	315.86	INCLUDED	INCLUDED
RIDGE (KM + VRP + GOOGLE, 2007-2019)	11.11	207.29	INCLUDED	INCLUDED
LASSO (KM + VRP + GOOGLE, 2007-2019)	14.73	275.70	INCLUDED	INCLUDED
ELASTIC NET (KM + VRP + GOOGLE, 2007-2019)	14.25	265.60	INCLUDED	INCLUDED
ADDITIONAL MODES				
VAR (KM + 4 variables, 2007-2019)	11.16	179.72	INCLUDED	INCLUDED
PCR (KM + 4 variables, 2007-2019)	20.34	747.84	INCLUDED	INCLUDED
PLS (KM + 4 variables, 2007-2019)	21.33	922.41	INCLUDED	INCLUDED
RIDGE (KM + 4 variables, 2007-2019)	9.64	170.31	INCLUDED	INCLUDED
LASSO (KM + 4 variables, 2007-2019)	12.24	193.72	INCLUDED	INCLUDED
ELASTIC NET (KM + 4 variables, 2007-2019)	11.73	180.27	INCLUDED	INCLUDED
PCR (KM + VRP + GOOGLE + 4 v., 2007-2019)	21.53	908.39	INCLUDED	INCLUDED
PLS (KM + VRP + GOOGLE + 4 v., 2007-2019)	22.22	1043.39	INCLUDED	INCLUDED
RIDGE (KM + VRP + GOOGLE + 4 v., 2007-2019)	**9.61**	163.62	INCLUDED	INCLUDED
LASSO (KM + VRP + GOOGLE + 4 v., 2007-2019)	11.80	189.16	INCLUDED	INCLUDED
ELASTIC NET (KM+VRP+ GOOGLE+4 v., 2007-2019)	11.31	179.32	INCLUDED	INCLUDED

Table 8: Augmented models' performances according to the mean squared error (MSE) and the mean absolute error (MAE) for ***24-month ahead forecasts***, as well as the inclusion in the Model Confidence Set. The smallest values are reported in bold font. All models were augmented with 4 additional variables: Baltic Dirty Tanker Index, capacity utilization rate, global crude exports, and the US Federal Funds rate

MODELS	MAE	MSE	MCS(MAE)	MCS(MSE)
BASELINE MODES				
Random Walk (no-change forecast)	14.31	289.05	INCLUDED	INCLUDED
VAR (KM, 1973-2019)	22.32	581.09	INCLUDED	INCLUDED
VAR (KM + VRP, 2007-2019)	14.60	352.55	INCLUDED	INCLUDED
VAR (KM + GOOGLE, 2007-2019)	14.69	357.63	INCLUDED	INCLUDED
VAR (KM + VRP + GOOGLE, 2007-2019)	13.52	302.72	INCLUDED	INCLUDED
PCR (KM, 2007-2019)	22.29	620.29	INCLUDED	INCLUDED
PLS (KM, 2007-2019)	21.64	614.87	INCLUDED	INCLUDED
RIDGE (KM, 2007-2019)	**6.88**	103.69	INCLUDED	INCLUDED
LASSO (KM, 2007-2019)	15.88	347.20	INCLUDED	INCLUDED
ELASTIC NET (KM, 2007-2019)	15.92	343.15	INCLUDED	INCLUDED
PCR (KM + VRP, 2007-2019)	28.35	867.14		
PLS (KM + VRP, 2007-2019)	31.42	1099.40		
RIDGE (KM + VRP, 2007-2019)	9.01	131.37	INCLUDED	INCLUDED
LASSO (KM + VRP, 2007-2019)	15.46	329.55	INCLUDED	INCLUDED
ELASTIC NET (KM + VRP, 2007-2019)	15.59	334.67	INCLUDED	INCLUDED
PCR (KM + GOOGLE, 2007-2019)	31.17	1283.35		
PLS (KM + GOOGLE, 2007-2019)	30.83	1262.67		
RIDGE (KM + GOOGLE, 2007-2019)	7.26	107.84	INCLUDED	INCLUDED
LASSO (KM + GOOGLE, 2007-2019)	14.70	303.93	INCLUDED	INCLUDED
ELASTIC NET (KM + GOOGLE, 2007-2019)	14.40	293.32	INCLUDED	INCLUDED
PCR (KM + VRP + GOOGLE, 2007-2019)	31.99	1142.86		
PLS (KM + VRP + GOOGLE, 2007-2019)	30.87	1091.59		
RIDGE (KM + VRP + GOOGLE, 2007-2019)	8.96	129.71	INCLUDED	INCLUDED
LASSO (KM + VRP + GOOGLE, 2007-2019)	14.65	319.14	INCLUDED	INCLUDED
ELASTIC NET (KM + VRP + GOOGLE, 2007-2019)	14.28	293.93	INCLUDED	INCLUDED
ADDITIONAL MODES				
VAR (KM + 4 variables, 2007-2019)	12.66	297.09	INCLUDED	INCLUDED
PCR (KM + 4 variables, 2007-2019)	36.18	3239.83	INCLUDED	INCLUDED
PLS (KM + 4 variables, 2007-2019)	41.34	7194.72	INCLUDED	INCLUDED
RIDGE (KM + 4 variables, 2007-2019)	7.77	**103.46**	INCLUDED	INCLUDED
LASSO (KM + 4 variables, 2007-2019)	11.53	216.08	INCLUDED	INCLUDED
ELASTIC NET (KM + 4 variables, 2007-2019)	11.50	210.64	INCLUDED	INCLUDED
PCR (KM + VRP + GOOGLE + 4 v., 2007-2019)	37.02	3504.88	INCLUDED	INCLUDED
PLS (KM + VRP + GOOGLE + 4 v., 2007-2019)	43.85	7819.17	INCLUDED	INCLUDED
RIDGE (KM + VRP + GOOGLE + 4 v., 2007-2019)	8.98	122.12	INCLUDED	INCLUDED
LASSO (KM + VRP + GOOGLE + 4 v., 2007-2019)	10.91	182.62	INCLUDED	INCLUDED
ELASTIC NET (KM+VRP+ GOOGLE+4 v., 2007-2019)	10.84	180.22	INCLUDED	INCLUDED

Table 9: Augmented models' performances according to the mean squared error (MSE) and the mean absolute error (MAE) for ***1-month ahead forecasts***, as well as the inclusion in the Model Confidence Set. The smallest values are reported in bold font. The first two HVAR models were also augmented with 4 additional variables: Baltic Dirty Tanker Index, capacity utilization rate, global crude ex-ports, and the US Federal Funds rate.

MODELS	**MAE**	**MSE**	**MCS(MAE)**	**MCS(MSE)**
BASELINE MODES				
Random Walk (no-change forecast)	3.25	18.98	INCLUDED	INCLUDED
VAR (KM, 1973-2019)	3.55	19.40	INCLUDED	INCLUDED
VAR (KM + VRP, 2007-2019)	3.33	**18.60**	INCLUDED	INCLUDED
VAR (KM + GOOGLE, 2007-2019)	3.40	18.82	INCLUDED	INCLUDED
VAR (KM + VRP + GOOGLE, 2007-2019)	3.72	34.49	INCLUDED	INCLUDED
PCR (KM, 2007-2019)	5.18	50.35	INCLUDED	INCLUDED
PLS (KM, 2007-2019)	6.99	82.36		INCLUDED
RIDGE (KM, 2007-2019)	9.03	177.04		INCLUDED
LASSO (KM, 2007-2019)	3.31	20.03	INCLUDED	INCLUDED
ELASTIC NET (KM, 2007-2019)	3.61	22.45	INCLUDED	INCLUDED
PCR (KM + VRP, 2007-2019)	9.35	138.89		
PLS (KM + VRP, 2007-2019)	12.43	259.85		
RIDGE (KM + VRP, 2007-2019)	8.38	146.14	INCLUDED	INCLUDED
LASSO (KM + VRP, 2007-2019)	**3.24**	19.81	INCLUDED	INCLUDED
ELASTIC NET (KM + VRP, 2007-2019)	3.71	24.13	INCLUDED	INCLUDED
PCR (KM + GOOGLE, 2007-2019)	76.07	13754.97		
PLS (KM + GOOGLE, 2007-2019)	57.83	7118.25		
RIDGE (KM + GOOGLE, 2007-2019)	8.69	162.40	INCLUDED	INCLUDED
LASSO (KM + GOOGLE, 2007-2019)	3.49	21.12	INCLUDED	INCLUDED
ELASTIC NET (KM + GOOGLE, 2007-2019)	3.79	25.13	INCLUDED	INCLUDED
PCR (KM + VRP + GOOGLE, 2007-2019)	14.15	333.42		
PLS (KM + VRP + GOOGLE, 2007-2019)	13.58	315.49		
RIDGE (KM + VRP + GOOGLE, 2007-2019)	7.74	133.61	INCLUDED	INCLUDED
LASSO (KM + VRP + GOOGLE, 2007-2019)	3.42	20.70	INCLUDED	INCLUDED
ELASTIC NET (KM + VRP + GOOGLE, 2007-2019)	3.75	24.06	INCLUDED	INCLUDED
ADDITIONAL MODES				
HVAR (KM + 4 variables, 2007-2019)	4.87	43.34	INCLUDED	INCLUDED
HVAR (KM + VRP + GOOGLE + 4 v., 2007-2019)	7.42	81.69		
HVAR (KM + VRP + GOOGLE , 2007-2019)	5.05	45.99	INCLUDED	INCLUDED
HVAR (KM, 1973-2019)	3.79	24.11	INCLUDED	INCLUDED

Table 10: Augmented models' performances according to the mean squared error (MSE) and the mean absolute error (MAE) for ***6-month ahead forecasts***, as well as the inclusion in the Model Confidence Set. The smallest values are reported in bold font. The first two HVAR models were also augmented with 4 additional variables: Baltic Dirty Tanker Index, capacity utilization rate, global crude ex-ports, and the US Federal Funds rate.

MODELS	**MAE**	**MSE**	**MCS(MAE)**	**MCS(MSE)**
BASELINE MODES				
Random Walk (no-change forecast)	8.65	107.35	INCLUDED	INCLUDED
VAR (KM, 1973-2019)	10.77	172.87	INCLUDED	INCLUDED
VAR (KM + VRP, 2007-2019)	8.61	103.63	INCLUDED	INCLUDED
VAR (KM + GOOGLE, 2007-2019)	8.73	105.91	INCLUDED	INCLUDED
VAR (KM + VRP + GOOGLE, 2007-2019)	8.65	103.78	INCLUDED	INCLUDED
PCR (KM, 2007-2019)	21.37	623.53		INCLUDED
PLS (KM, 2007-2019)	27.19	1095.14		
RIDGE (KM, 2007-2019)	10.71	241.15	INCLUDED	INCLUDED
LASSO (KM, 2007-2019)	8.84	129.22	INCLUDED	INCLUDED
ELASTIC NET (KM, 2007-2019)	8.60	122.58	INCLUDED	INCLUDED
PCR (KM + VRP, 2007-2019)	17.67	424.86		INCLUDED
PLS (KM + VRP, 2007-2019)	37.83	5060.37		INCLUDED
RIDGE (KM + VRP, 2007-2019)	10.34	202.77	INCLUDED	INCLUDED
LASSO (KM + VRP, 2007-2019)	7.89	96.04	INCLUDED	INCLUDED
ELASTIC NET (KM + VRP, 2007-2019)	7.56	**81.72**	INCLUDED	INCLUDED
PCR (KM + GOOGLE, 2007-2019)	91.21	20453.62		
PLS (KM + GOOGLE, 2007-2019)	64.73	9549.37		
RIDGE (KM + GOOGLE, 2007-2019)	10.22	235.78	INCLUDED	INCLUDED
LASSO (KM + GOOGLE, 2007-2019)	9.73	162.92	INCLUDED	INCLUDED
ELASTIC NET (KM + GOOGLE, 2007-2019)	9.22	151.46	INCLUDED	INCLUDED
PCR (KM + VRP + GOOGLE, 2007-2019)	23.42	1046.20		
PLS (KM + VRP + GOOGLE, 2007-2019)	22.48	912.92		INCLUDED
RIDGE (KM + VRP + GOOGLE, 2007-2019)	10.26	207.82	INCLUDED	INCLUDED
LASSO (KM + VRP + GOOGLE, 2007-2019)	9.38	145.33	INCLUDED	INCLUDED
ELASTIC NET (KM + VRP + GOOGLE, 2007-2019)	9.17	138.17	INCLUDED	INCLUDED
ADDITIONAL MODES				
HVAR (KM + 4 variables, 2007-2019)	16.32	412.17	INCLUDED	INCLUDED
HVAR (KM + VRP + GOOGLE + 4 v., 2007-2019)	21.88	645.31		INCLUDED
HVAR (KM + VRP + GOOGLE , 2007-2019)	16.70	436.24	INCLUDED	INCLUDED
HVAR (KM, 1973-2019)	**7.35**	83.90	INCLUDED	INCLUDED

Table 11: Augmented models' performances according to the mean squared error (MSE) and the mean absolute error (MAE) for ***12-month ahead forecasts***, as well as the inclusion in the Model Confidence Set. The smallest values are reported in bold font. The first two HVAR models were also augmented with 4 additional variables: Baltic Dirty Tanker Index, capacity utilization rate, global crude ex-ports, and the US Federal Funds rate.

MODELS	**MAE**	**MSE**	**MCS(MAE)**	**MCS(MSE)**
BASELINE MODES				
Random Walk (no-change forecast)	9.99	139.57	INCLUDED	INCLUDED
VAR (KM, 1973-2019)	16.91	360.28	INCLUDED	INCLUDED
VAR (KM + VRP, 2007-2019)	10.23	157.59	INCLUDED	INCLUDED
VAR (KM + GOOGLE, 2007-2019)	10.47	160.15	INCLUDED	INCLUDED
VAR (KM + VRP + GOOGLE, 2007-2019)	10.52	174.88	INCLUDED	INCLUDED
PCR (KM, 2007-2019)	16.00	344.92	INCLUDED	INCLUDED
PLS (KM, 2007-2019)	19.39	512.58	INCLUDED	INCLUDED
RIDGE (KM, 2007-2019)	11.48	223.69	INCLUDED	INCLUDED
LASSO (KM, 2007-2019)	13.72	242.10	INCLUDED	INCLUDED
ELASTIC NET (KM, 2007-2019)	12.93	222.45	INCLUDED	INCLUDED
PCR (KM + VRP, 2007-2019)	26.12	1456.01	INCLUDED	INCLUDED
PLS (KM + VRP, 2007-2019)	34.12	2328.52		INCLUDED
RIDGE (KM + VRP, 2007-2019)	10.87	212.47	INCLUDED	INCLUDED
LASSO (KM + VRP, 2007-2019)	14.40	267.73	INCLUDED	INCLUDED
ELASTIC NET (KM + VRP, 2007-2019)	14.46	272.04	INCLUDED	INCLUDED
PCR (KM + GOOGLE, 2007-2019)	67.80	8913.46		
PLS (KM + GOOGLE, 2007-2019)	52.85	5849.85		
RIDGE (KM + GOOGLE, 2007-2019)	11.56	209.95	INCLUDED	INCLUDED
LASSO (KM + GOOGLE, 2007-2019)	13.19	228.80	INCLUDED	INCLUDED
ELASTIC NET (KM + GOOGLE, 2007-2019)	13.46	239.84	INCLUDED	INCLUDED
PCR (KM + VRP + GOOGLE, 2007-2019)	14.66	363.41	INCLUDED	INCLUDED
PLS (KM + VRP + GOOGLE, 2007-2019)	13.82	315.86	INCLUDED	INCLUDED
RIDGE (KM + VRP + GOOGLE, 2007-2019)	11.11	207.29	INCLUDED	INCLUDED
LASSO (KM + VRP + GOOGLE, 2007-2019)	14.73	275.70	INCLUDED	INCLUDED
ELASTIC NET (KM + VRP + GOOGLE, 2007-2019)	14.25	265.60	INCLUDED	INCLUDED
ADDITIONAL MODES				
HVAR (KM + 4 variables, 2007-2019)	19.65	509.26	INCLUDED	INCLUDED
HVAR (KM + VRP + GOOGLE + 4 v., 2007-2019)	22.84	646.24	INCLUDED	INCLUDED
HVAR (KM + VRP + GOOGLE , 2007-2019)	19.96	529.60	INCLUDED	INCLUDED
HVAR (KM, 1973-2019)	**7.57**	**87.64**	INCLUDED	INCLUDED

Table 12: Augmented models' performances according to the mean squared error (MSE) and the mean absolute error (MAE) for ***24-month ahead forecasts***, as well as the inclusion in the Model Confidence Set. The smallest values are reported in bold font. The first two HVAR models were also augmented with 4 additional variables: Baltic Dirty Tanker Index, capacity utilization rate, global crude ex-ports, and the US Federal Funds rate.

MODELS	MAE	MSE	MCS(MAE)	MCS(MSE)
BASELINE MODES				
Random Walk (no-change forecast)	14.31	289.05	INCLUDED	INCLUDED
VAR (KM, 1973-2019)	22.32	581.09		INCLUDED
VAR (KM + VRP, 2007-2019)	14.60	352.55	INCLUDED	INCLUDED
VAR (KM + GOOGLE, 2007-2019)	14.69	357.63	INCLUDED	INCLUDED
VAR (KM + VRP + GOOGLE, 2007-2019)	13.52	302.72	INCLUDED	INCLUDED
PCR (KM, 2007-2019)	22.29	620.29	INCLUDED	INCLUDED
PLS (KM, 2007-2019)	21.64	614.87	INCLUDED	INCLUDED
RIDGE (KM, 2007-2019)	**6.88**	**103.69**	INCLUDED	INCLUDED
LASSO (KM, 2007-2019)	15.88	347.20	INCLUDED	INCLUDED
ELASTIC NET (KM, 2007-2019)	15.92	343.15	INCLUDED	INCLUDED
PCR (KM + VRP, 2007-2019)	28.35	867.14		
PLS (KM + VRP, 2007-2019)	31.42	1099.40		
RIDGE (KM + VRP, 2007-2019)	9.01	131.37	INCLUDED	INCLUDED
LASSO (KM + VRP, 2007-2019)	15.46	329.55	INCLUDED	INCLUDED
ELASTIC NET (KM + VRP, 2007-2019)	15.59	334.67	INCLUDED	INCLUDED
PCR (KM + GOOGLE, 2007-2019)	31.17	1283.35		
PLS (KM + GOOGLE, 2007-2019)	30.83	1262.67		
RIDGE (KM + GOOGLE, 2007-2019)	7.26	107.84	INCLUDED	INCLUDED
LASSO (KM + GOOGLE, 2007-2019)	14.70	303.93	INCLUDED	INCLUDED
ELASTIC NET (KM + GOOGLE, 2007-2019)	14.40	293.32	INCLUDED	INCLUDED
PCR (KM + VRP + GOOGLE, 2007-2019)	31.99	1142.86		
PLS (KM + VRP + GOOGLE, 2007-2019)	30.87	1091.59		
RIDGE (KM + VRP + GOOGLE, 2007-2019)	8.96	129.71	INCLUDED	INCLUDED
LASSO (KM + VRP + GOOGLE, 2007-2019)	14.65	319.14	INCLUDED	INCLUDED
ELASTIC NET (KM + VRP + GOOGLE, 2007-2019)	14.28	293.93	INCLUDED	INCLUDED
ADDITIONAL MODES				
HVAR (KM + 4 variables, 2007-2019)	19.52	465.14	INCLUDED	INCLUDED
HVAR (KM + VRP + GOOGLE + 4 v., 2007-2019)	21.65	538.21	INCLUDED	INCLUDED
HVAR (KM + VRP + GOOGLE , 2007-2019)	20.02	476.29	INCLUDED	INCLUDED
HVAR (KM, 1973-2019)	10.36	143.33	INCLUDED	INCLUDED

Table 13: Augmented models' performances according to the mean squared error (MSE) and the mean absolute error (MAE) for ***1-month ahead forecasts***, as well as the inclusion in the Model Confidence Set. The smallest values are reported in bold font.

MODELS	MAE	MSE	MCS(MAE)	MCS(MSE)
BASELINE MODES				
Random Walk (no-change forecast)	3.25	18.98	INCLUDED	INCLUDED
VAR (KM, 1973-2019)	3.55	19.40	INCLUDED	INCLUDED
VAR (KM + VRP, 2007-2019)	3.33	**18.60**	INCLUDED	INCLUDED
VAR (KM + GOOGLE, 2007-2019)	3.40	18.82	INCLUDED	INCLUDED
VAR (KM + VRP + GOOGLE, 2007-2019)	3.72	34.49	INCLUDED	INCLUDED
PCR (KM, 2007-2019)	5.18	50.35	INCLUDED	INCLUDED
PLS (KM, 2007-2019)	6.99	82.36		INCLUDED
RIDGE (KM, 2007-2019)	9.03	177.04	INCLUDED	INCLUDED
LASSO (KM, 2007-2019)	3.31	20.03	INCLUDED	INCLUDED
ELASTIC NET (KM, 2007-2019)	3.61	22.45	INCLUDED	INCLUDED
PCR (KM + VRP, 2007-2019)	9.35	138.89		
PLS (KM + VRP, 2007-2019)	12.43	259.85		
RIDGE (KM + VRP, 2007-2019)	8.38	146.14	INCLUDED	INCLUDED
LASSO (KM + VRP, 2007-2019)	3.24	19.81	INCLUDED	INCLUDED
ELASTIC NET (KM + VRP, 2007-2019)	3.71	24.13	INCLUDED	INCLUDED
PCR (KM + GOOGLE, 2007-2019)	76.07	13754.97		
PLS (KM + GOOGLE, 2007-2019)	57.83	7118.25		
RIDGE (KM + GOOGLE, 2007-2019)	8.69	162.40	INCLUDED	INCLUDED
LASSO (KM + GOOGLE, 2007-2019)	3.49	21.12	INCLUDED	INCLUDED
ELASTIC NET (KM + GOOGLE, 2007-2019)	3.79	25.13	INCLUDED	INCLUDED
PCR (KM + VRP + GOOGLE, 2007-2019)	14.15	333.42		
PLS (KM + VRP + GOOGLE, 2007-2019)	13.58	315.49		
RIDGE (KM + VRP + GOOGLE, 2007-2019)	7.74	133.61	INCLUDED	INCLUDED
LASSO (KM + VRP + GOOGLE, 2007-2019)	3.42	20.70	INCLUDED	INCLUDED
ELASTIC NET (KM + VRP + GOOGLE, 2007-2019)	3.75	24.06	INCLUDED	INCLUDED
ADDITIONAL MODES				
PCR (KM + VRP + simple GOOGLE, 2007-2019)	4.61	37.09	INCLUDED	INCLUDED
PLS (KM + VRP + simple GOOGLE, 2007-2019)	5.76	49.62	INCLUDED	INCLUDED
RIDGE (KM + VRP + simple GOOGLE,2007-2019)	6.01	74.47	INCLUDED	INCLUDED
LASSO (KM + VRP + simple GOOGLE,2007-2019)	**3.24**	20.20	INCLUDED	INCLUDED
ELASTIC NET (KM+VRP+simple GOOGLE,2007-2019)	3.53	21.48	INCLUDED	INCLUDED

Table 14: Augmented models' performances according to the mean squared error (MSE) and the mean absolute error (MAE) for ***6-month ahead forecasts***, as well as the inclusion in the Model Confidence Set. The smallest values are reported in bold font.

MODELS	MAE	MSE	MCS(MAE)	MCS(MSE)
BASELINE MODES				
Random Walk (no-change forecast)	8.65	107.35	INCLUDED	INCLUDED
VAR (KM, 1973-2019)	10.77	172.87	INCLUDED	INCLUDED
VAR (KM + VRP, 2007-2019)	8.61	103.63	INCLUDED	INCLUDED
VAR (KM + GOOGLE, 2007-2019)	8.73	105.91	INCLUDED	INCLUDED
VAR (KM + VRP + GOOGLE, 2007-2019)	8.65	103.78	INCLUDED	INCLUDED
PCR (KM, 2007-2019)	21.37	623.53		INCLUDED
PLS (KM, 2007-2019)	27.19	1095.14		INCLUDED
RIDGE (KM, 2007-2019)	10.71	241.15	INCLUDED	INCLUDED
LASSO (KM, 2007-2019)	8.84	129.22	INCLUDED	INCLUDED
ELASTIC NET (KM, 2007-2019)	8.60	122.58	INCLUDED	INCLUDED
PCR (KM + VRP, 2007-2019)	17.67	424.86		INCLUDED
PLS (KM + VRP, 2007-2019)	37.83	5060.37		INCLUDED
RIDGE (KM + VRP, 2007-2019)	10.34	202.77	INCLUDED	INCLUDED
LASSO (KM + VRP, 2007-2019)	7.89	96.04	INCLUDED	INCLUDED
ELASTIC NET (KM + VRP, 2007-2019)	**7.56**	**81.72**	INCLUDED	INCLUDED
PCR (KM + GOOGLE, 2007-2019)	91.21	20453.62		
PLS (KM + GOOGLE, 2007-2019)	64.73	9549.37		
RIDGE (KM + GOOGLE, 2007-2019)	10.22	235.78	INCLUDED	INCLUDED
LASSO (KM + GOOGLE, 2007-2019)	9.73	162.92	INCLUDED	INCLUDED
ELASTIC NET (KM + GOOGLE, 2007-2019)	9.22	151.46	INCLUDED	INCLUDED
PCR (KM + VRP + GOOGLE, 2007-2019)	23.42	1046.20		INCLUDED
PLS (KM + VRP + GOOGLE, 2007-2019)	22.48	912.92		INCLUDED
RIDGE (KM + VRP + GOOGLE, 2007-2019)	10.26	207.82	INCLUDED	INCLUDED
LASSO (KM + VRP + GOOGLE, 2007-2019)	9.38	145.33	INCLUDED	INCLUDED
ELASTIC NET (KM + VRP + GOOGLE, 2007-2019)	9.17	138.17	INCLUDED	INCLUDED
ADDITIONAL MODES				
PCR (KM + VRP + simple GOOGLE, 2007-2019)	14.75	318.37		INCLUDED
PLS (KM + VRP + simple GOOGLE, 2007-2019)	14.10	292.80		INCLUDED
RIDGE (KM + VRP + simple GOOGLE,2007-2019)	8.33	141.37	INCLUDED	INCLUDED
LASSO (KM + VRP + simple GOOGLE,2007-2019)	9.39	134.98	INCLUDED	INCLUDED
ELASTIC NET (KM+VRP+simple GOOGLE,2007-2019)	9.05	128.32	INCLUDED	INCLUDED

Table 15: Augmented models' performances according to the mean squared error (MSE) and the mean absolute error (MAE) for ***12-month ahead forecasts***, as well as the inclusion in the Model Confidence Set. The smallest values are reported in bold font.

MODELS	**MAE**	**MSE**	**MCS(MAE)**	**MCS(MSE)**
BASELINE MODES				
Random Walk (no-change forecast)	**9.99**	**139.57**	INCLUDED	INCLUDED
VAR (KM, 1973-2019)	16.91	360.28	INCLUDED	INCLUDED
VAR (KM + VRP, 2007-2019)	10.23	157.59	INCLUDED	INCLUDED
VAR (KM + GOOGLE, 2007-2019)	10.47	160.15	INCLUDED	INCLUDED
VAR (KM + VRP + GOOGLE, 2007-2019)	10.52	174.88	INCLUDED	INCLUDED
PCR (KM, 2007-2019)	16.00	344.92	INCLUDED	INCLUDED
PLS (KM, 2007-2019)	19.39	512.58		INCLUDED
RIDGE (KM, 2007-2019)	11.48	223.69	INCLUDED	INCLUDED
LASSO (KM, 2007-2019)	13.72	242.10	INCLUDED	INCLUDED
ELASTIC NET (KM, 2007-2019)	12.93	222.45	INCLUDED	INCLUDED
PCR (KM + VRP, 2007-2019)	26.12	1456.01		INCLUDED
PLS (KM + VRP, 2007-2019)	34.12	2328.52		INCLUDED
RIDGE (KM + VRP, 2007-2019)	10.87	212.47	INCLUDED	INCLUDED
LASSO (KM + VRP, 2007-2019)	14.40	267.73	INCLUDED	INCLUDED
ELASTIC NET (KM + VRP, 2007-2019)	14.46	272.04	INCLUDED	INCLUDED
PCR (KM + GOOGLE, 2007-2019)	67.80	8913.46		
PLS (KM + GOOGLE, 2007-2019)	52.85	5849.85		
RIDGE (KM + GOOGLE, 2007-2019)	11.56	209.95	INCLUDED	INCLUDED
LASSO (KM + GOOGLE, 2007-2019)	13.19	228.80	INCLUDED	INCLUDED
ELASTIC NET (KM + GOOGLE, 2007-2019)	13.46	239.84	INCLUDED	INCLUDED
PCR (KM + VRP + GOOGLE, 2007-2019)	14.66	363.41	INCLUDED	INCLUDED
PLS (KM + VRP + GOOGLE, 2007-2019)	13.82	315.86	INCLUDED	INCLUDED
RIDGE (KM + VRP + GOOGLE, 2007-2019)	11.11	207.29	INCLUDED	INCLUDED
LASSO (KM + VRP + GOOGLE, 2007-2019)	14.73	275.70	INCLUDED	INCLUDED
ELASTIC NET (KM + VRP + GOOGLE, 2007-2019)	14.25	265.60	INCLUDED	INCLUDED
ADDITIONAL MODES				
PCR (KM + VRP + simple GOOGLE, 2007-2019)	13.89	276.88	INCLUDED	INCLUDED
PLS (KM + VRP + simple GOOGLE, 2007-2019)	13.19	259.72	INCLUDED	INCLUDED
RIDGE (KM + VRP + simple GOOGLE,2007-2019)	10.33	180.31	INCLUDED	INCLUDED
LASSO (KM + VRP + simple GOOGLE,2007-2019)	11.98	189.04	INCLUDED	INCLUDED
ELASTIC NET (KM+VRP+simple GOOGLE,2007-2019)	11.54	183.02	INCLUDED	INCLUDED

Table 16: Augmented models' performances according to the mean squared error (MSE) and the mean absolute error (MAE) for ***24-month ahead forecasts***, as well as the inclusion in the Model Confidence Set. The smallest values are reported in bold font.

MODELS	MAE	MSE	MCS(MAE)	MCS(MSE)
BASELINE MODES				
Random Walk (no-change forecast)	14.31	289.05	INCLUDED	INCLUDED
VAR (KM, 1973-2019)	22.32	581.09		
VAR (KM + VRP, 2007-2019)	14.60	352.55	INCLUDED	INCLUDED
VAR (KM + GOOGLE, 2007-2019)	14.69	357.63	INCLUDED	INCLUDED
VAR (KM + VRP + GOOGLE, 2007-2019)	13.52	302.72	INCLUDED	INCLUDED
PCR (KM, 2007-2019)	22.29	620.29	INCLUDED	INCLUDED
PLS (KM, 2007-2019)	21.64	614.87	INCLUDED	INCLUDED
RIDGE (KM, 2007-2019)	**6.88**	**103.69**	INCLUDED	INCLUDED
LASSO (KM, 2007-2019)	15.88	347.20	INCLUDED	INCLUDED
ELASTIC NET (KM, 2007-2019)	15.92	343.15	INCLUDED	INCLUDED
PCR (KM + VRP, 2007-2019)	28.35	867.14		
PLS (KM + VRP, 2007-2019)	31.42	1099.40		
RIDGE (KM + VRP, 2007-2019)	9.01	131.37	INCLUDED	INCLUDED
LASSO (KM + VRP, 2007-2019)	15.46	329.55	INCLUDED	INCLUDED
ELASTIC NET (KM + VRP, 2007-2019)	15.59	334.67	INCLUDED	INCLUDED
PCR (KM + GOOGLE, 2007-2019)	31.17	1283.35		
PLS (KM + GOOGLE, 2007-2019)	30.83	1262.67		
RIDGE (KM + GOOGLE, 2007-2019)	7.26	107.84	INCLUDED	INCLUDED
LASSO (KM + GOOGLE, 2007-2019)	14.70	303.93	INCLUDED	INCLUDED
ELASTIC NET (KM + GOOGLE, 2007-2019)	14.40	293.32	INCLUDED	INCLUDED
PCR (KM + VRP + GOOGLE, 2007-2019)	31.99	1142.86		
PLS (KM + VRP + GOOGLE, 2007-2019)	30.87	1091.59		
RIDGE (KM + VRP + GOOGLE, 2007-2019)	8.96	129.71	INCLUDED	INCLUDED
LASSO (KM + VRP + GOOGLE, 2007-2019)	14.65	319.14	INCLUDED	INCLUDED
ELASTIC NET (KM + VRP + GOOGLE, 2007-2019)	14.28	293.93	INCLUDED	INCLUDED
ADDITIONAL MODES				
PCR (KM + VRP + simple GOOGLE, 2007-2019)	16.96	334.92		INCLUDED
PLS (KM + VRP + simple GOOGLE, 2007-2019)	16.73	332.81	INCLUDED	INCLUDED
RIDGE (KM + VRP + simple GOOGLE,2007-2019)	8.99	127.23	INCLUDED	INCLUDED
LASSO (KM + VRP + simple GOOGLE,2007-2019)	14.90	294.62	INCLUDED	INCLUDED
ELASTIC NET (KM+VRP+simple GOOGLE,2007-2019)	14.75	284.56	INCLUDED	INCLUDED