Forecasting the Realized Variance of Oil-Price Returns Using Machine-Learning: Is there a Role for U.S. State-Level Uncertainty?

Oğuzhan Çepni[a], Rangan Gupta[b], Daniel Pienaar[c], Christian Pierdzioch[d]

January 2022

Abstract

Predicting the variance of oil-price returns is of paramount importance for policymakers and investors. Recent research has focused on whether disaggregate measures of economic-policy uncertainty provide better forecasts. Given that the United States (U.S.) is a major player in the international oil market, we extend this line of research by exploring by means of machine-learning techniques whether accounting for U.S. state-level measures of economic-policy uncertainty results in more accurate forecasts. We find improvements in forecast accuracy, especially when we study intermediate and long forecast horizons. This finding is robust to various changes in the model configuration (realized variance vs. realized volatility, sample period, recursive vs. rolling-estimation window, loss function of forecast consumers). Understandably, our findings have important implications for oil traders and policy authorities.

JEL classification: C22, C53, D8, Q02

Keywords: Oil Price; Realized variance; Forecasting; Machine Learning; Aggregate and Regional Uncertainties

[a] Copenhagen Business School, Department of Economics, Porcelænshaven 16A, Frederiksberg DK-2000, Denmark; Central Bank of the Republic of Turkey, Haci Bayram Mah. Istiklal Cad. No:10 06050, Ankara, Turkey. Email address: oce.eco@cbs.dk.

[b] Department of Economics, University of Pretoria, Private Bag X20, Hatfield 0028, South Africa. Email address: Rangan.Gupta@up.ac.za

[c] Department of Economics, University of Pretoria, Private Bag X20, Hatfield 0028, South Africa. Email address: u14033705@tuks.co.za.

[d] Department of Economics, Helmut Schmidt University, Holstenhofweg 85, P.O.B. 700822, 22008 Hamburg, Germany. Email address: macroeconmoics@hsu-hh.de.

1 Introduction

The severity of the Global Financial Crisis (GFC), followed by the prolonged European sovereign debt crisis, the Brexit, and, of course, the ongoing COVID-19 pandemic, have all highlighted the risks associated with portfolios comprised of conventional financial assets (Balcilar et al., 2017, 2020; Muteba Mwamba et al., 2017). These extreme events have, in turn, triggered an interest in alternative investment opportunities, particularly in the commodity market, since investors search for diversification opportunities by supplementing their traditional portfolios with positions in commodities, and oil in particular (Bampinas and Panagiotidis, 2015, 2017). The recent financialization of the commodity market (Bahloul et al., 2018; Bonato, 2019) has been reflected in increased participation of hedge funds, pension funds, and insurance companies in commodity investments, and crude oil is now considered a profitable alternative instrument in the portfolio decisions of financial institutions (Degiannakis and Filis, 2017). Given this, the market size of crude oil investments is \$1.7 trillion per year at current spot prices, with 34 billion barrels produced each year and over 1.7 trillion barrels of crude oil in remaining reserves (U.S. Energy Information Administration (EIA); BP Statistical Review of World Energy), making it by far the most actively traded commodity.[1] Since the variance of price movements is an essential input to investment decisions and portfolio choices (Poon and Granger, 2003), accurate forecasts of the realized variance of oil-price returns are of paramount importance to oil traders. At the same time, the variability oil-price returns have been shown to have predictive value for slowdowns in worldwide economic growth (van Eyden et al., 2019; Salisu et al., 2021), and, hence, policy-makers need precise estimates of the future movements in oil-price variability to conceptualize macroeconomic policies ahead of time to prevent economic recessions, or at least reduce the depth of the same.

Given the importance of oil-price variability for both investors and policy authorities, and the increased economic and financial uncertainties, following a series of crises over the last decade

[1]Gold comes in a distant second, valued at \$170 billion per year at current spot prices with a production of over 3,200 tons per annum (World Gold Council).

and a half, several contributions to the academic literature have delved into the role of uncertainty for forecasting the variability of oil-price movements (see, for example, Bonaccolto et al. (2018), Bouri et al. (2020), Li et al. (2020), Liang et al. (2020), Dutta et al. (2021), Gupta and Pierdzioch (2021, forthcoming)). From a theoretical perspective, the relationship between uncertainty and the variability of oil-price returns (and commodities in general) is based on the 'Theory of Storage' (Bakas and Triantafyllou, 2018; 2020). This theory posits that increases in uncertainty would cause the future path of aggregate demand of commodities, and hence, aggregate production to be less predictable. Due to this increased unpredictability, risk-averse commodity producers would raise their holding of physical inventory. As a result of which, the convenience yield would rise, and this, in turn, would heighten fluctuations of commodity, including oil prices.

At this stage, to provide the motivation behind our research and highlight our contribution, we present a brief discussion of the current literature on uncertainty and its role as a predictor of the future volatility of oil-price returns. In this regard, Bonaccolto et al., (2018) analyze the role of (newspapers-based) metrics of economic policy and equity market uncertainties of the United States (U.S.) as predictors of the conditional quantiles of crude-oil returns and its volatility through a non-parametric *k*-th order causality-in-quantiles framework. They find that these uncertainty measures have heterogeneous effects across different quantiles, but in general, are relevant during periods of market stress, which is basically when oil risk is of paramount importance for market participants. Bouri et al., (2020), also using the daily newspaper-based index of U.S. uncertainty but related to infectious diseases (EMVID), report that adding EMVID to a heterogeneous autoregressive model of (realized) volatility (HAR-RV) significantly improves forecast accuracy across a range of short-, medium-, and long-run forecast horizons. Pursuing a different route, Li et al., (2020) apply the mixed-data-sampling generalized autoregressive conditional heteroskedastic (MIDAS-GARCH) model to shed light on the contribution of monetary-policy uncertainty besides overall economic-policy uncertainty of the U.S. in forecasting oil-market volatility. Dutta et al., (2021), based on a quantiles-based modeling framework, however, show that unlike overall uncertainty of the U.S. related to policies, equity-market volatility of the U.S., and the same due to commodity-market movements and crises, have a stronger predictive value for oil-market volatility. Along similar lines, Liang et al., (2020) also emphasize the importance of overall equity-market volatility indexes of the U.S. using a standard predictive-regression model, model combination, and shrinkage techniques. Unlike these papers, Li et al., (2020) find no evidence of a metric of global uncertainty in predicting oil-market volatility. Taking a global perspective, Gupta and Pierdzioch (2021a), using shrinkage estimators, show that text-based uncertainty indexes for the group of G7 (Canada, France, Germany, Italy, Japan, the United Kingdom (U.K.), and the U.S.) countries and China, as well as their spillovers to the rest of the world, contain valuable forecasting information for the volatility of West Texas Intermediate (WTI) and Brent crude oil-price movements at intermediate- and long-run horizons. Finally, Gupta and Pierdzioch (forthcoming) compare the role of aggregate versus several disaggregated metrics of policy-related and equity market-based uncertainties of the U.S. (as well as the same for geopolitical risks) in forecasting the realized volatility of WTI oil-price returns. Relying on machine-learning techniques, they find that adding the disaggregated uncertainty metrics to the vector of predictors improves the accuracy of forecasts at intermediate and long forecast horizons. This is especially so when a nonlinear machine-learning method (random forests) is used to estimate the forecasting model.

Our research is motivated by the need to disaggregate measures of uncertainty in producing more accurate forecasts of oil-price volatility, as shown by Gupta and Pierdzioch (forthcoming). Against this backdrop, we aim to add to this growing empirical literature on uncertainty and forecasting of oil-market variance, by comparing the role of aggregate versus state-level metrics of economic and policy-related uncertainties of the U.S., in predicting the future realized variance (and its square root, that is, realized volatility) of WTI oil-price returns over the monthly period from 1984/01 to 2019/12. In our forecasting experiment, we not only compare the performance of the overall uncertainty versus all the state-level versions of the same, but also consider the scenario when we complement the information content of the former with the latter. The intuition to look at state-level uncertainties in forecasting the realized variance of oil-price returns is straightforward and can be explained as follows: The U.S. states are exceptionally heterogeneous in terms of their oil dependency, calculated as oil consumed minus oil produced as a percentage of oil consumed, and in the process the strengths of their status as oil suppliers and demanders (De

Michelis et al., 2021). At the same time, studies have shown that demand and supply conditions of the oil market is reflected in not only the aggregate measures of U.S. uncertainty (Degiannakis et al., 2018; Shahzad et al., 2021), but also at the state-level (Gupta and Sheng, 2021). In other words, given their underlying heterogeneity in terms of oil dependency, the regional uncertainties provide a proxy for the local demand and supply conditions for oil (as outlined by Gupta and Pierdzioch (2021) at the global level involving the G7 and China), which, in turn, are expected to drive oil-market volatility (Pan et al., 2017; Demirer et al., 2020).[2] Naturally, if state-level uncertainty measures produce additional forecasting gains relative to the aggregate counterpart, then this is a finding of considerable value to both investors and policymakers, as well as for academics investigating the possibility of new factors that drive volatility of oil-price.

For our forecasting experiment, from an econometric perspective, as in Gupta and Pierdzioch (forthcoming), we use a statistical data-driven approach involving not only linear, but also nonlinear machine-learning methods in deciding on the predictive role of the large number of uncertainty-based predictors associated with the aggregate and disaggregated regions of the U.S.. In terms of the forecasting models, first, we use the least absolute shrinkage and selection operator (Lasso), proposed by Tibshirani (1996). The Lasso belongs to the class of linear regression-analysis methods and performs, in a data-driven way, a regularization of the forecasting model and a selection of predictors so as to improve the interpretability of the model and the accuracy of predictions derived from the regularized model. Second, we move to a nonlinear setting and estimate random forests (Breiman, 2001), which, in turn, is a machine-learning technique designed to operate for a large array of predictors. In essence, random forests automatically capture any potential nonlinear links (as widely depicted in the succinct literature review provided above) between the oil-market volatility and the uncertainties, as well as any interaction effects (as shown to exist across state-level uncertainties by Gupta et al., (2020)) between the predictors.

At this stage, it is important to clarify two additional issues: Firstly, we forecast the monthly realized variance (*RV*) of returns of the WTI oil-price, where we capture the realized variance

[2]Interestingly, Van Robays (2016) has shown that different levels of macroeconomic uncertainty over time can explain time variation in the price elasticity of oil, and therefore in oil-price volatility originating from oil shocks.

as the sum of daily squared returns over a month (following Andersen and Bollerslev (1998)), which, in turn, yields an observable and unconditional measure of volatility – an otherwise latent process. Traditionally,[3] the time-varying volatility of oil-price returns is based on various models belonging to the GARCH family, under which the conditional variance is a deterministic function of model parameters and past data. Alternatively, some researchers have considered stochastic volatility (SV) models in recent works, wherein volatility is depicted as a latent variable that follows a stochastic process. In this regard, whether a researcher uses GARCH or SV models, the resulting volatility estimate is not unconditional or model-free, as it is the case with RV. Secondly, while oil is a global commodity, since our focus is data on the state-level economic policy uncertainty of the U.S., we consider the WTI as our proxy for the world oil-price, but this is permissible, since the U.S. is a major player on both demand- and supply-fronts of the oil market.

To the best of our knowledge, this is the first paper to compare the role of aggregate and regional measures of uncertainties in forecasting the RV of oil-price returns, using linear and nonlinear machine-learning techniques, besides standard Ordinary Least Squares (OLS) applied to the smaller models (in terms of predictors) involving only the national metric. In this manner, our research adds to the already existing significant literature on forecasting of oil-returns volatility by considering the role of overall and state-level uncertainties, with existing studies having considered the predictive content of the information derived from a large number of macroeconomic, financial, behavioral, and climate patterns-related variables based on a broad spectrum of (non-)linear univariate and multivariate models (see, for example, Asai et al., (2019, 2020), Bonato et al., (2020), Demirer et al., (2020, 2021), Gkillas et al., (2020), Bouri et al., (2021), Gupta and Pierdzioch (2021b), Salisu et al., (2022), Luo et al., (forthcoming), and the references cited within these papers).

We organize the remainder of our research as follows: Section 2 contains a description of the data, while Section 3 is devoted to the methodologies. Then, we discuss in Section 4 the forecasting results, and finally we conclude in Section 5.

[3] See the discussions in Chan and Grant (2016) and Lux et al., (2016) for further details.

2 Data

We compute, in a first step, the daily log-returns, i.e., first-differences of the natural logarithm of the WTI oil price, and then obtain, in a second step, the monthly realized variances (RV) by summing up the daily squared log-returns over a specific month. In this regard, note that the underlying daily WTI crude oil nominal price date is derived from Global Financial Data.[4] Figure 1 plots the resulting realized variance.

– Figure 1 about here. –

We next turn to our aggregate and state-level uncertainty data, for which we rely on a text-based approach that constructs indexes of uncertainty from searches of keywords or terms related to (economic and policy) uncertainty (EPU) in major newspapers, following the work of Baker et al., (2016). In this regard, Baker et al., (2016) create an index of search results from 10 large U.S. newspapers (USA Today, the Miami Herald, the Chicago Tribune, the Washington Post, the Los Angeles Times, the Boston Globe, the San Francisco Chronicle, the Dallas Morning News, the Houston Chronicle, and the WSJ), with the searches involving terms related to economic and policy uncertainty. In particular, these authors search for articles containing the term 'uncertainty' or 'uncertain', the terms 'economic' or 'economy' and one or more of the following terms: 'congress', 'legislation', 'white house', 'regulation', 'federal reserve', or 'deficit'.[5]

As far as the state-level measure of EPU is concerned, we rely on the work of Elkamhi et al., (2020),[6] who basically follow the newspapers-based approach of Baker et al., (2016). Elkamhi

[4] `https://globalfinancialdata.com/`.

[5] The data is available for download from: `http://policyuncertainty.com/us_monthly.html`.

[6] We would like to thank the authors of this paper for kindly providing us with the state-level EPU data. Note that state-level uncertainty data have also been developed by Mumtaz (2018), and Mumtaz et al., (2018) based on stochastic-volatility estimates computed using a large-scale structural model involving macroeconomic and financial variables of the U.S. states. Because their metrics are at lower, i.e., quarterly-frequency, and contingent on an econometric framework, we decided to use the higher-frequency measures of uncertainty, based on a model-free approach, which is likely to be more informative to investors and policy authorities in undertaking their respective decisions.

et al., (2020), using news articles from Newslibrary.com,[7] search for the number of articles containing words that are related to the following categories: “State-level", “Economic",“Policy", and “Uncertainty". The authors count an article as related to state-level EPU (SEPU) when it contains at least one word for each of the four categories. Because state newspapers could cover not only local but also nationwide news at the same time, Elkamhi et al., (2020) discard articles that contain a word reflective of nationwide information (such as ‘congress’, ‘white house’, ‘federal reserve’).[8]

– Figures 2 and 3 about here. –

Based on data availability on RV and the predictors, our analysis covers the monthly period from 1984/01 to 2019/12,[9] though our main focus is the period 2000/01 to 2019/12, for which EPU data are available for most states. Figure 2 depicts the aggregate EPU data, while Figure 3 plots, as an example, the state-level EPU data for three states that are among the top oil-producing states (and on which we shall focus in one of the robustness analyses we shall present in Section 4).

3 Methods

Our first forecasting model (Model 1) is given by the following simple autoregressive model:

$$RV_{t+h} = \beta_0 + \beta_1 RV_t + \eta_{t+h}, \quad (1)$$

where we estimate the model coefficients, $\beta_i, i = 0, 1$, by means of the ordinary-least-squares (OLS) technique, the parameter, h, is the forecast horizon, and η_{t+h} is a disturbance term. In

[7]Newslibrary.com covers around 7,000 newspapers with more than 274 million newspaper articles for 50 U.S. states as well as the District of Columbia (DC), Puerto Rico, Guam, U.S. Virgin Islands, and American Samoa.

[8]The reader is referred to Table 1 of Elkamhi et al., (2020) for the complete list of words used to select articles according to their methodology.

[9]For the period of 1984/01 to 1984/12, we use the historical EPU data, which is based on articles from 6 (the Wall Street Journal, the New York Times, the Washington Post, the Chicago Tribune, the LA Times, and the Boston Globe), rather than the 10 above-mentioned newspapers.

order to check whether our results hold for different forecast horizons, we set $h = 1, 3, 6, 12, 24$, where we forecast the the average realized variance whenever we opt for a forecast horizon that exceeds one month, that is, when we set $h > 1$. We structure our data matrix so that its dimension is the same for all forecast horizons.

In the next step, we extend our forecasting model given in Equation (1) to include economic-policy uncertainty (EPU) as an additional predictor. This gives Model 2:

$$RV_{t+h} = \beta_0 + \beta_1 RV_t + \beta_2 EPU_t + \eta_{t+h}. \tag{2}$$

We obtain forecasting Model 3 when we replace EPU with its state-level components (SEPU):

$$RV_{t+h} = \beta_0 + \beta_1 RV_t + \sum_{i=1}^{n} \gamma_i SEPU_{i,t} + \eta_{t+h}. \tag{3}$$

where γ denote regression coefficients, and n denotes the number of states being used to estimate Equation (3). This number depends on the availability of state-level SEPU data and, thus, on the sample period that we study.

Finally, we also shall present results for a forecasting model that features both EPU and the SEPU data. This gives Model 4:

$$RV_{t+h} = \beta_0 + \beta_1 RV_t + \beta_2 EPU_t + \sum_{i=1}^{n} \gamma_i SEPU_{i,t} + \eta_{t+h}. \tag{4}$$

In summary, we consider four different forecasting models, where the number of predictors and, thus, the complexity of the forecasting models increases as we move from Model 1 to Model 4. As a result, we estimate the parsimonious Models 1 and 2 by the ordinary-least-squares (OLS) technique, and Models 3 and 4 by two alternative machine-learning techniques. For convenience, Table 1 provides summary information of the estimated models and the estimation techniques being used.

– Table 1 about here. –

The first machine-learning technique that we consider to estimate Models 3 to 4 is the least absolute shrinkage and selection operator (Lasso), as introduced into the literature by Tibshirani

(1996). The Lasso estimates of the coefficients of the forecasting models are selected so as to minimize, using Model 4 as an example, the following expression (for a detailed textbook exposition, see, e.g., Hastie et al., 2009):

$$\sum_{t=1}^{N}(RV_{t+h}-\beta_0-\beta_1 RV_t-\beta_2 EPU_t-\sum_{i=1}^{n}\gamma_i SEPU_{i,t})^2+\lambda\left(\sum_{i=0}^{2}|\beta_i+\sum_{i=1}^{n}|\gamma_i|\right), \tag{5}$$

where we let N denote the number of observations available for estimation of the model. Equation (5) witnesses that the Lasso estimator is a shrinkage estimator that uses the L1 norm of the coefficient vector to shrink the dimension of the estimated forecasting model, where the penalty parameter, λ, governs the strength of this shrinkage effect. Hence, when λ takes on a large value, the Lasso estimator can even set to zero some of the model coefficients.

The second machine-learning technique that we consider is the random forest technique (Breiman, 2001). The advantage of random forests are twofold. First, random forests capture in a purely data-driven way any interaction effects between the predictors, which may arise due to spatial correlations of state-level economic-policy uncertainty in the context of our empirical analysis or due to the common response of level economic-policy uncertainty to a common macroeconomic shock. Second, random forests automatically capture potential nonlinear links between the realized variance of oil-price returns and the predictors. Hence, random forests, unlike the Lasso estimator, do not invoke a linear structure on the forecasting model but rather serve to recover a forecasting model of the following general format (again using Model 4 as an example):

$$RV_{t+h}=f(RV_t,EPU_t,SEPU_{1,t},...,SEPU_{n,t}) \tag{6}$$

Random forests approximate the function, f, in an automatic and data-driven way. This approach involves a large number of individual random regression trees and, hence, is associated with the class of ensemble machine-learning techniques.[10] A regression tree has branches that recursively partition the space of predictors into non-overlapping regions in hierarchical and binary manners. The growth of a large regression tree, which translates into many such regions, allows for the computation of increasingly granular predictions of the RV. However, simultaneously, the

[10]The interested reader is referred to the textbook by Hastie et al., (2009).

complicated hierarchical structure of a regression tree eventually leads to over-fitting problems and sensitiveness to the data. A random forest fixes these problems based on a combination approach involving many random individual regression trees. In this regard, the researcher draws a large number of bootstrap samples from the data and then estimates a random regression tree associated with every bootstrap sample. Note that a random regression tree tends to differ from a standard regression tree since a random subset of the predictors is used to produce growth in the nodes and branches of the tree. This additional aspect of randomness ameliorates the potential effect of influential individual predictors (possibly due to outliers in the data) associated with the building of a tree. Furthermore, because a random forest involves a large number of random regression trees, it causes a reduction in the correlation of predictions, which, in turn, are computed utilizing the individual random regression trees. Ultimately, the computation of the average of the predictions across the individual regression trees results in the stabilization of the random-forest-based forecasts of the *RV* of oil returns.

For estimation of our forecasting models, we use the R language and environment for statistical computing (R Core Team, 2021). We use the R add-on package “glmnet” (Friedman et al., 2010) to implement the Lasso estimator and the add-on package “grf” (Tibshirani et al., 2021) to estimate random forests. We use 10-fold cross-validation to identify the numerical value of the shrinkage parameter that minimizes the mean cross-validated error. As for random forests, every random forest consists of 500 individual random regression trees. We use cross-validation to select the tree parameters (number of randomly sampled predictors selected for tree building, minimum number of data at a terminal tree node, and maximum imbalance of a split at a node of a tree).

We estimate our forecasting models by means of a recursive- and a rolling-estimation window. We use a rolling-estimation window of length five and ten years and, in case of the recursive-estimation window, a training period of five or ten years to initialize the estimates. Moreover, because the magnitude of fluctuations of oil-price returns varied widely during our sample period (see also Figure 1), we shall present, as a robustness check, results for both the realized variance of oil-price returns (*RV*) and the corresponding realized volatility (that is, $\sqrt{RV}$).

In order to evaluate forecast accuracy, we use variants of the out-of-sample R^2 statistic (Campbell and Thompson, 2008). For a squared-error-loss function, we compute the out-of-sample R^2 statistic as follows:

$$R^2 = 1 - \frac{\sum_{t=1}^{T}(x_{R,t} - \hat{x}_{R,t})^2}{\sum_{t=1}^{T}(x_{B,t} - \hat{x}_{B,t})^2}, \tag{7}$$

where T denotes the number of out-of-sample forecasts available, x_t denotes the realization of either RV_t or $\sqrt{RV_t}$, a hat denotes a forecast, and the indexes, R and B denote the rival and the benchmark model. When we observe $R^2 > 0$ ($R^2 < 0$) then the rival (benchmark) model outperforms the benchmark (rival) model.

As a robustness check, we also compute the out-of-sample R^2 statistic for an absolute-error-loss function. In this case, the out-of-sample R^2 statistic is given by

$$R^2 = 1 - \frac{\sum_{t=1}^{T}|x_{R,t} - \hat{x}_{R,t}|}{\sum_{t=1}^{T}|x_{B,t} - \hat{x}_{B,t}|\surd}. \tag{8}$$

Finally, we use a general loss function considered by Patton (2011) to evaluate forecasts. This loss function, L, is given by

$$L = \begin{cases} \frac{1}{(b+1)(b+2)}\left(RV^{b+2} - \hat{RV}\surd^{b+2}\right) - \frac{1}{b+1}\hat{RV}\surd^{b+1}\left(RV - \hat{RV}\surd\right) & \text{for} \quad b \neq -2, b \neq -1, \\ \hat{RV}\surd - RV + RV \log\frac{RV}{\hat{RV}\surd} & \text{for} \quad b = -1, \\ \frac{RV}{\hat{RV}\surd} - \log\frac{RV}{\hat{RV}\surd} - 1 & \text{for} \quad b = -2, \end{cases} \tag{9}$$

where b denotes the shape parameter of this loss function. For $b = -2$ the general loss function degenerates to the QLIKE loss function studied by Patton (2011), and for $b = 0$ the standard symmetric squared-error loss function emerges as a special case. Hence, this loss function covers the case of a symmetric loss function but also accounts for the possibility of an asymmetric loss function, which naturally arises when a forecast consumer, due to a specific trading strategy or behavioral dispositions, attaches different losses to over- and underestimations of the same size. Specifically, the loss function attaches a higher loss to an under-prediction when $b < 0$ and to an over-prediction when $b > 0$. The out-of-sample R^2 statistic corresponding to the loss function specified in Equation (9) is given by

$$R^2 = 1 - \frac{\sum_{t=1}^{T} L(b, RV_{R,t}, \hat{RV}_{R,t})}{\sum_{t=1}^{T} L(b, RV_{B,t}, \hat{RV}_{B,t})\surd}. \tag{10}$$

It should be noted that this out-of-sample R^2 statistic is a function of the shape parameter, b. Hence, we will analyze the performance of the respective benchmark and rival models b means of a graph that depicts the out-of-sample R^2 as a function of the shape parameter, b.

4 Empirical Results

Table 2 summarizes the out-of-sample R^2 statistic for the sample period 2000/01−2019/12, for which EPU data are available for most states.[11] Estimation of Models 3 and 4 is done by the Lasso estimator. We present results for a recursive and a rolling-estimation window, and for the realized variance (RV) as well as the realized volatility ($\sqrt{RV}$). We observe a strong tendency of the out-of-sample R^2 statistic to turn from being negative for the short forecast horizons to become positive for the intermediate and long forecast horizons.[12] Hence, Models 2 to 4 tend to outperform Model 1 when we let the length of the forecast horizon to increase. Similarly, Model 4, the model that includes the SEPU data in the list of predictors, outperforms Model 2, which uses EPU as a predictor, when the forecast horizon gets longer. This pattern consistently emerges irrespective of whether we study the realized variance or the realized volatility of oil-price returns, or whether we use a recursive- or a rolling-estimation window to compute forecasts. The results also witness that there is a tendency for the out-of-sample R^2 statistics to increase when the forecast horizon increases. Hence, the predictive value of EPU and its state-level component tends to strengthen when we forecast the longer-term realized variance and realized volatility of oil-price returns.

[11] We exclude Delaware and Wyoming, but retain states for which only few data are not available for 2000 (like Rhode Island) so as to maximize the number of states used for estimation of the forecasting models.

[12] As market agents care not only about the nature of volatility, but also of its level, with all traders making distinctions between “good" and “bad" volatilities (Giot et al., 2010), we also carried out our forecasting exercise by decomposing the overall RV into these two components. We continued to observe a similar pattern in terms of the results, i.e., the out-of-sample R^2 statistic, when we use the positive and negative realized semi-variances (i.e., ‘good and bad realized variances, respectively) of oil-price returns. Due to this finding, detailed results have not been formally reported for the sake of brevity, but are available upon request from the authors.

– Please include Tables 2 and 3 about here. –

In Table 3, we use the absolute-error-loss function rather than the conventional squared-error-loss function to compute the out-of-sample R^2 statistic. Results are again for the Lasso estimator and the sample period 2000/01−2019/12. The results for the absolute-error-loss function corroborate the results fore the squared-error-loss function. Again, we observe that Models 2 to 4 outperform Model 1, and Model 4 outperforms Model 2, for the intermediate and long forecast horizons, where the magnitude of the out-of-sample R^2 statistic shows a tendency to increase in the forecast horizon.

– Please include Table 4 about here. –

Table 4 depicts the results we obtain when we switch from the Lasso estimator to random forests. Random forests have the advantage that they automatically identify in a data-driven way potential nonlinear links in the data and also interaction effects among the predictors. The sample period is 2000/01−2019/12 and out-of-sample R^2 statistic is based on the squared-error-loss function. Again, we find that the out-of-sample R^2 statistic turns from negative to positive and tends to increase when we opt for a longer forecast horizon. We obtain this finding for both a recursive- and a rolling-estimation window, and for the realized variance as well as the realized volatility of oil-price returns.

– Figure 4 about here. –

In Figure 4, we use the more general loss function given in Equation (9) to study the the predictive value of economic-policy uncertainty and its state-level realizations. The results are based on the realized-variance forecasts implied by the estimated random forests. The out-of-sample R^2 statistic on the vertical axis is depicted as a function of the shape parameter, b, of the loss function. The dashed vertical lines show the special cases of the QLIKE loss function (for $b = -2$) and the standard symmetric squared-error loss function (for $b = 0$). We present the results for a recursive-estimation window. The figure visualizes how the performance of the rival models relative to the benchmark models improves as the length of the forecast horizon increases. While the relative performance of the rival models depends on the shape parameter, b, for intermediate forecast horizons (for example, for $h = 12$ when we use the short training period), the rival models outperform the benchmark models for all values of the shape parameter plotted in the figure for the long forecast horizons.

– Please include Table 5 about here. –

Table 5 summarizes results for two extended sample periods. The first extended sample period ranges from 1984/01 to 2019/12, but includes only SEPU data for Texas, as this is the only major oil-producing state for which data extend back that far in time. The second extended sample period is somewhat shorter and starts in 1990/01, but includes the SEPU data for three major oil-producing states: Texas, North Dakota, and Oklahoma. We present results for the Lasso estimator (Models 3 and 4), and we focus on a squared-error-loss function and the realized volatility of oil-price returns. The results are consistent with the findings of the shorter sample period that starts in 2000. The out-of-sample R^2 statistics, in general, are negative for the short forecast horizons, but turn positive (and increase in value) for the intermediate and long forecast horizon. Hence, economic-policy uncertainty becomes more important for forecasting the realized volatility of oil-price returns when we increase the length of the forecast horizon. Similarly, the comparison of Models 2 and 4 shows that the state-level components of economic-policy uncertainty tend to add predictive value beyond the predictive value already contributed by aggregate economic-policy uncertainty at intermediate and long forecast horizons.

– Please include Table 6 about here. –

Our finding that the out-of-sample R^2 statistic is increasing in the forecast horizon implies that forecast consumers who target the longer-term low-frequency movements of the realized variance/volatility benefit from using S(EPU) data as predictors. A natural question, therefore, how the results for the out-of-sample R^2 statistic look like when we target the future realization of the realized variance/volatility of oil-price returns rather than its mean value over the respective forecast horizon. We call this setting the "point forecast" scenario and report the results in Table 6. The results are in line with what we expected. First, the out-of-sample R^2 statistics tend to be smaller than in Table 5. Second, the out-of-sample R^2 statistics exhibit a U-shaped pattern. The statistics are negative for the short forecast horizons, as for those forecast horizons differentiating between mean forecasts and point forecasts should not make much difference. The statistics then increase in the forecast horizon, but eventually they decrease again for the very long forecast horizon(s) because at those forecast horizons it becomes increasingly more difficult to make accurate point forecasts than to target lower frequency movements of the realized variance/volatility of oil-price returns.

5 Concluding Remarks

Our findings, derived using two widely-studied machine-learning techniques, lend support to the hypotheses that for a sufficiently long forecast horizon: (i) including aggregate economic-policy uncertainty as a predictor improves the accuracy of forecasts of the realized variance/volatility of oil-price returns, and, (ii) that using disaggregate U.S. state-level economic-policy uncertainty rather than its aggregate counterpart improves the accuracy of forecasts of the realized variance/volatility of oil-price returns in a systematic way. We deem this to be an important result not only from the perspective of a forecasting practitioner, but also for academic research on the sources of fluctuations of oil prices. These implications are of course in addition to the importance of these results, in terms of the predictive role of aggregate and regional uncertainties for oil price variability which would assist in the portfolio allocation decisions of oil traders, and also for policymakers in designing macroeconomic policies under the context of oil volatility-recession nexus.

While we have focuses in this research on (state-level) U.S. economic-policy uncertainty, given that the U.S. is a major player in the international oil market, it is interesting in future research to ask whether U.S. or international sources of economic-policy uncertainty play a more prominent role as oil-market-volatility predictors (as in Gupta and Pierdzioch (2021)). This question is interesting given that results reported in recent empirical research (Gupta and Pierdzioch, forthcoming) suggest that accounting for disaggregate international country-specific measures of economic-policy uncertainty helps forecast consumers to improve their forecasts of oil-market volatility. Hence, a natural extension of our research is to run a forecasting competition between state-level U.S. economic-policy uncertainty and international economic-policy uncertainty as predictors of oil-market volatility. Given that such a competition requires to account for a large number of predictors, the machine-learning techniques that we have used in our research lend themselves to run such a horse-race.

References

Andersen T.G., and Bollerslev T. (1998). Answering the skeptics: yes, standard volatility models do provide accurate forecasts. International Economic Review, 39(4), 885−905.

Asai, M. Gupta, R., and McAleer, M. (2020). Forecasting volatility and co-volatility of crude oil and gold futures: Effects of leverage, jumps, spillovers, and geopolitical risks. International Journal of Forecasting, 36(3), 933−948.

Asai, M. Gupta, R., and McAleer, M. (2019). The impact of jumps and leverage in forecasting the co-volatility of oil and gold futures. Energies, 12, 3379.

Bakas, D., and Triantafyllou, A. (2018). The impact of uncertainty shocks on the volatility of commodity prices. Journal of International Money and Finance, 87, 96−111.

Bakas, D., and Triantafyllou, A. (2020). Commodity price volatility and the economic uncertainty of pandemics. Economics Letters, 193, 109283.

Baker, S.R., Bloom, N.A., and Davis, S.J. (2016). Measuring economic policy uncertainty. Quarterly Journal of Economics, 131(4), 1593−1636.

Bahloul, W., Balcilar, M., Cunado, J., and Gupta, R. (2018). The role of economic and financial uncertainties in predicting commodity futures returns and volatility: Evidence from a nonparametric causality-in-quantiles test. Journal of Multinational Financial Management, 45, 52−71.

Balcilar, M., Demirer, R., and Gupta, R. (2017). Do sustainable stocks offer diversification benefits for conventional portfolios? An empirical analysis of risk spillovers and dynamic correlations. Sustainability, 9, 1799.

Balcilar, M., Demirer, R., Gupta, R., and Wohar, M.E. (2020). The effect of global and regional stock market shocks on safe haven assets. Structural Change and Economic Dynamics, 54, 297−308.

Bampinas, G., and Panagiotidis, T. (2015). On the relationship between oil and gold before and after financial crisis: linear, nonlinear and time-varying causality testing. Studies in Nonlinear Dynamics and Econometrics, 19(5), 657−668.

Bampinas, G., and Panagiotidis, T. (2017). Oil and stock markets before and after financial crises: A local Gaussian correlation approach. Journal of Futures Markets, 37(12), 1179−1204.

Bonaccolto, M., Caporin, M., and Gupta, R. (2018). The dynamic impact of uncertainty in causing and forecasting the distribution of oil returns and risk? Physica A: Statistical Mechanics and its Applications, 507, 446−469.

Bonato, M. (2019). Realized correlations, betas and volatility spillover in the agricultural commodity market: What has changed? Journal of International Financial Markets, Institutions and Money, 62, 184−202.

Bonato, M., Gkillas, K., Gupta, R., and Pierdzioch, C. (2020c). Investor happiness and predictability of the realized volatility of oil price. Sustainability, 12, 4309.

Bouri, E., Gkillas, K., Gupta, R., and Pierdzioch, C. (2020). Infectious diseases, market uncertainty and realized volatility of oil. Energies, 13(16), 4090.

Bouri, E., Gupta, R., Pierdzioch, C., and Salisu, A.A. (2021). El Niño and forecastability of oil-price realized volatility. Theoretical and Applied Climatology, 144, 1173−1180.

Breiman, L. (2001). Random forests. Machine Learning, 45, 5−32.

Campbell, J.Y., and Thompson, S.B. (2008). Predicting excess stock returns out of sample: Can anything beat the historical average? Review of Financial Studies 21(4), 1509−1531.

Chan, J.C.C, and Grant, A. (2016). Modeling energy price dynamics: GARCH versus stochastic volatility. Energy Economics, 54, 182−189.

De Michelis, A, Ferreira, T, Iacoviello, M. (2020). Oil prices and consumption across countries and U.S. States. International Journal of Central Banking, 16(2), 3−43.

Degiannakis, S., and Filis, G. (2017). Forecasting oil price realized volatility using information channels from other asset classes. Journal of International Money and Finance, 76, 28−49.

Degiannakis, S., Filis, G., and Panagiotakopoulou, S. (2018). Oil price shocks and uncertainty: How stable is their relationship over time? Economic Modelling, 72(C), 42−53.

Demirer, R., Gkillas, K., Gupta, R., and Pierdzioch, C. (2021). Risk aversion and the predictability of crude oil market volatility: A forecasting experiment with random forests. Journal of the Operational Research Society. DOI: https://doi.org/10.1080/01605682.2021.1936668.

Demirer, R., Gupta, R., Pierdzioch, C., and Shahzad, S.J.H. (2020). The predictive power of oil price shocks on realized volatility of oil: A note. Resources Policy, 69(C), 101856.

Dutta, A., Bouri, E., and Saeed, T. (2021). News-based equity market uncertainty and crude oil volatility. Energy, 222, 119930.

Elkamhi, R., Jo, C., and Salerno, M. (2020). Measuring state-level economic policy uncertainty. Available for download from SSRN at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3695365.

Friedman, J. Hastie, T., and Tibshirani, R. (2010). Regularization paths for Generalized Linear Models via coordinate descent. Journal of Statistical Software, 33(1), 1−22.

Giot, P., Laurent, S., and Petitjean, M. (2010). Trading activity, realized volatility and jumps. Journal of Empirical Finance, 17(1), 168−175.

Gkillas, K., Gupta, R., and Pierdzioch, C. (2020). Forecasting realized oil-price volatility: The role of financial stress and asymmetric loss. Journal of International Money and Finance, 104(C), 102137.

Gupta, R., Lau, C.K.M., and Sheng, X. (2020). Graph theory-based network analysis of regional uncertainties of the US Economy. Physica A: Statistical Mechanics and its Applications, 540(1), 123064.

Gupta, R., and Pierdzioch, C. (Forthcoming). Forecasting the realized variance of oil-price returns: A disaggregated analysis of the role of uncertainty and geopolitical risk. Environmental Science and Pollution.

Gupta, R., and Pierdzioch, C. (2021a). Forecasting the volatility of crude oil: The role of uncertainty and spillovers. Energies, 14(14), 4173.

Gupta, R., and Pierdzioch, C. (2021b). Climate risks and the realized volatility oil and gas prices: Results of an out-of-sample forecasting experiment. Energies, 14(23), 8085.

Gupta, R., and Sheng, X. (2021). The effects of oil shocks on macroeconomic uncertainty: Evidence from a large panel dataset of US States. Computational Management: Applications of Computational Intelligence in Business Management, Edited by Srikanta Patnaik, Vol. 18, Chapter 7, 159−175.

Hastie, T., Tibshirani, R., and Friedman, J. (2009). The elements of statistical learning: Data mining, inference, and prediction. 2nd ed., Springer: New York, NY, USA.

Li, X., Wei, Y., Chen, X., Ma, F., Liang, C., and Chen, W. (2020). Which uncertainty is powerful to forecast crude oil market volatility? New evidence. International Journal of Finance and Economics. DOI: `https://doi.org/10.1002/ijfe.2371`.

Liang, C., Wei, Y., Li, X., Zhang, X., and Zhang, Y. (2020). Uncertainty and crude oil market volatility: new evidence. Applied Economics, 52(27), 2945−2959.

Luo, J., Demirer, R., Gupta, R., and Ji, Q. (2022). Forecasting oil and gold volatilities with sentiment indicators under structural breaks. Energy Economics, 105, 105751.

item Lux, T., Segnon, M., and Gupta, R. (2016). Forecasting crude oil price volatility and value-at-risk: Evidence from historical and recent data. Energy Economics, 56, 117−133.

Mumtaz, H. (2018). Does uncertainty affect real activity. Evidence from state-level data. Economics Letters, 167, 127−130.

Mumtaz, H. , Sunder Plassmann, L., and Theophilopoulou, A. (2018), The state-level impact of uncertainty shocks. Journal of Money, Credit and Banking, 50(8), 1879−1899.

Muteba Mwamba, J. W., Hammoudeh, S., and Gupta, R. (2017). Financial tail risks in conventional and Islamic stock markets: a comparative analysis. Pacific-Basin Finance Journal, 42, 60−82.

Pan, Z., Wang, Y., Wu, C., and Yin, L. (2017). Oil price volatility and macroeconomic fundamentals: A regime switching GARCH-MIDAS model. Journal of Empirical Finance, 43(C), 130−142.

Patton, A.J. (2011). Volatility forecast comparison using imperfect volatility proxies. Journal of Econometrics, 160, 246−256.

Poon, S-H, and Granger, C.W.J. (2003). Forecasting volatility in financial markets: A review. Journal of Economic Literature, 41(2), 478−539.

R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL: `https://www.R-project.org/`.

Salisu, A.A., Gupta, R., Bouri, E., and Ji, Q. (2022). Mixed-frequency forecasting of crude oil volatility based on the information content of global economic conditions. Journal of Forecasting, 41(1), 134−157.

Salisu, A.A., Gupta, R., and Olaniran, A. (2021). The effect of oil uncertainty shock on real GDP of 33 countries: a global VAR approach. Applied Economics Letters. DOI: `https://doi.org/10.1080/13504851.2021.1983134`.

Shahzad, S.J.H., Gupta, R., Demirer, R., and Pierdzioch, C. (2021). Oil shocks and directional predictability of macroeconomic uncertainties of developed economies: Evidence from high-frequency data. Scottish Journal of Political Economy. DOI: `https://doi.org/10.1111/sjpe.12280`.

Tibshirani, J., Athey, S. E. Sverdrup, and Wager, S. (2021). grf: Generalized Random Forests. R package version 2.0.2. `https://CRAN.R-project.org/package=grf`.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society, Series B*, 58: 267−288.

Van Robays, I. (2016). Macroeconomic uncertainty and oil price volatility. Oxford Bulletin of Economics and Statistics, 78(5), 671−693.

van Eyden R., Difeto, M., Gupta, R., and Wohar, M.E. (2019). Oil price volatility and economic growth: Evidence from advanced economies using more than a century of data. Applied Energy, 233, 612−621.

Figure 1: Realized Variance of OIl-Price Returns

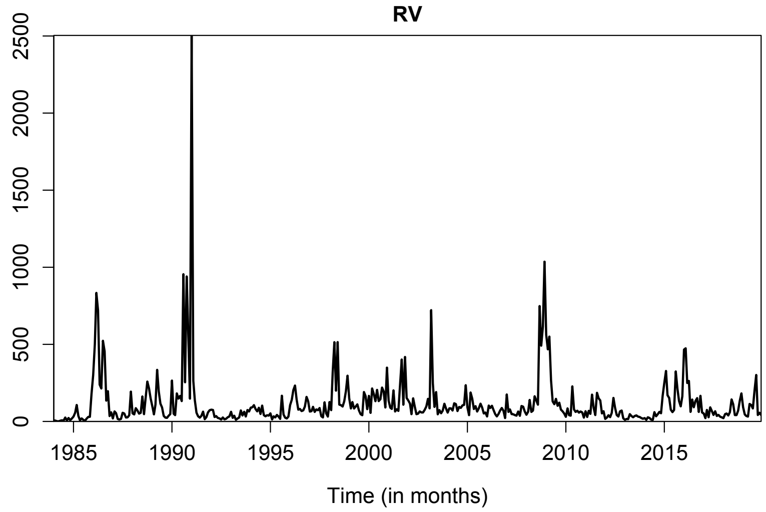


Figure 2: Aggregate EPU Data

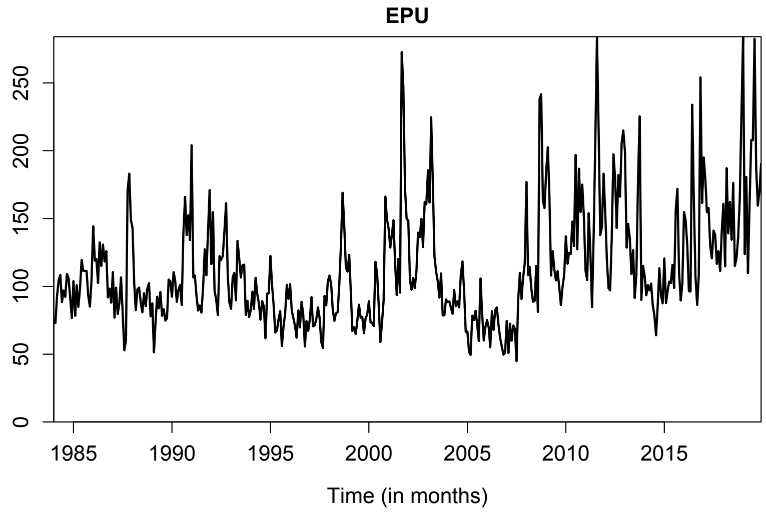


Figure 3: Examples of State-Level EPU Data

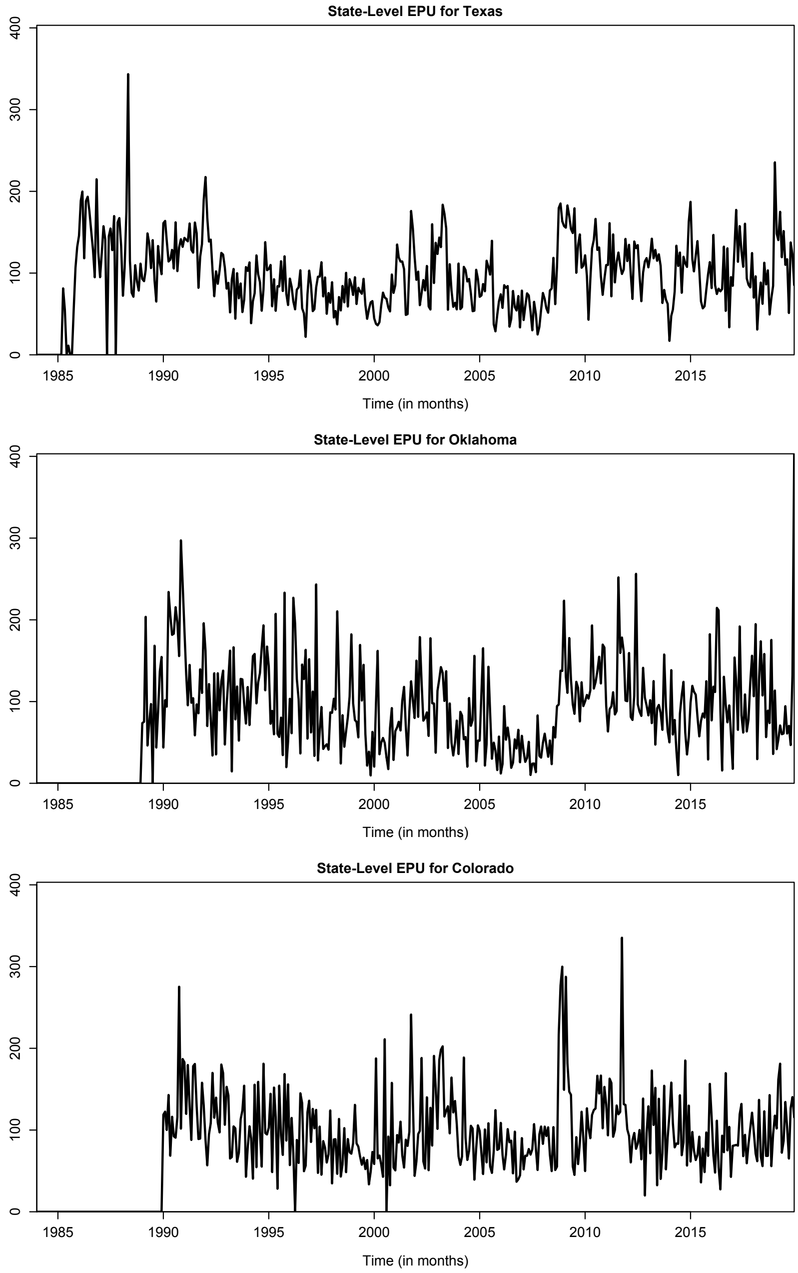
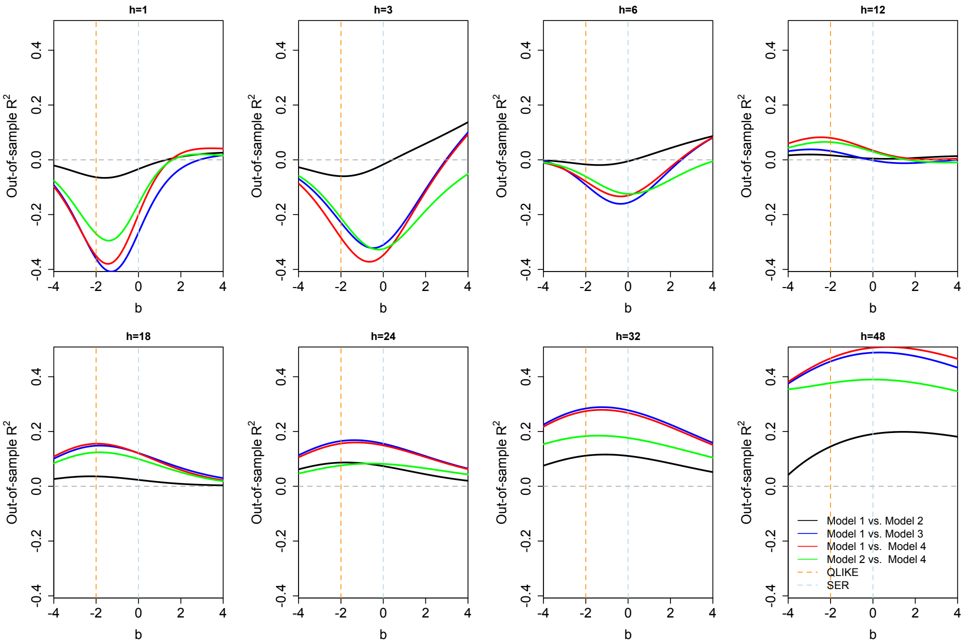


Table 1: Estimated Models

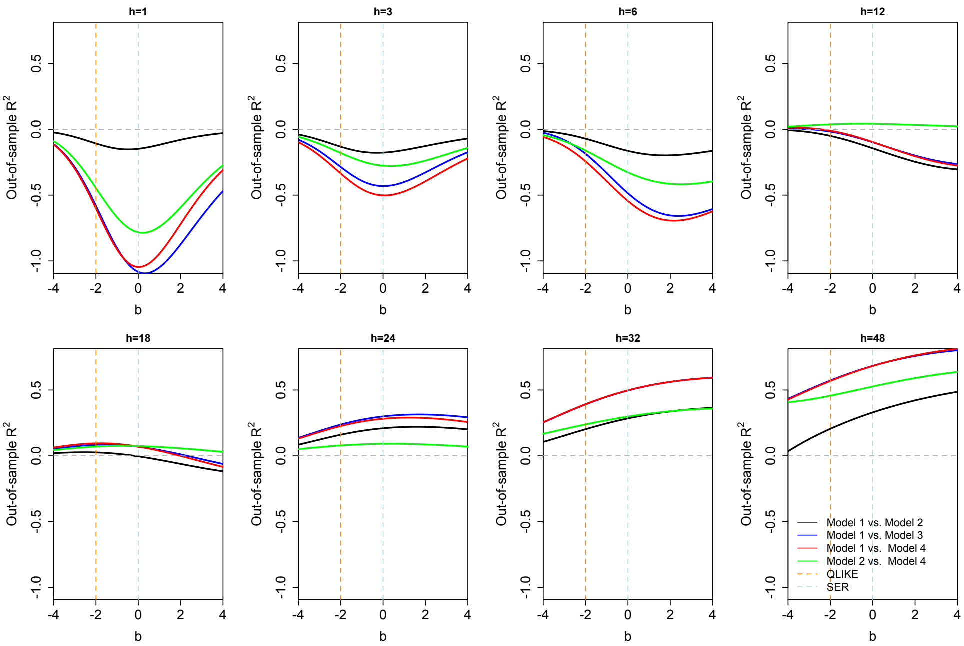
Model	Predictors	Estimation technique
Model 1	AR model	Ordinary-least squares
Model 2	AR term plus aggregate EPU	Ordinary-least squares
Model 3	AR term plus state-level EPUs	Lasso / random forests
Model 4	AR term plus aggregate and state-level EPUs	Lasso / random forests

Figure 4: Results for a General Loss Function

Panel A: Training period: five years



Panel B: Training period: ten years



The out-of-sample R^2 statistic is computed based on a general loss function and for a recursive-estimation window. The benchmark (rival) model is the first (second) model given in the legend. A positive out-of-sample R^2 statistic shows that the rival model outperforms the benchmark model. Results are for *RV* based on random forests and the sample period 2000/01−2019/12. The parameter h denotes the forecast horizon (in months). SER = Squared-error-loss function.

Table 2: Configuration: Lasso, Squared-Error-Loss Function, 2000/01−2019/12

Benchmark vs. rival model	Window	h=1	h=3	h=6	h=12	h=18	h=24	h=32	h=48
					Recursive-estimation window (RV)				
Model 1 vs. Model 2	60	-0.0337	-0.0169	-0.0052	0.0045	0.0234	0.0736	0.1108	0.1907
Model 1 vs. Model 2	120	-0.1479	-0.1770	-0.1639	-0.1443	-0.0042	0.2094	0.2841	0.3298
Model 1 vs. Model 3	60	-0.2523	-0.2764	-0.1136	-0.0439	0.1143	0.1761	0.2626	0.4993
Model 1 vs. Model 3	120	-0.6748	-0.2620	-0.2089	-0.0666	0.1245	0.2816	0.3977	0.6420
Model 1 vs. Model 4	60	-0.2972	-0.2951	-0.1483	-0.0106	0.1124	0.1882	0.2522	0.4907
Model 1 vs. Model 4	120	-0.6482	-0.2792	-0.2524	-0.1653	0.0641	0.3035	0.3936	0.6495
Model 2 vs. Model 4	60	-0.2549	-0.2736	-0.1423	-0.0151	0.0912	0.1237	0.1590	0.3707
Model 2 vs. Model 4	120	-0.4358	-0.0868	-0.0761	-0.0183	0.0681	0.1191	0.1529	0.4771
					Recursive-estimation window ($\sqrt{RV}$)				
Model 1 vs. Model 2	60	-0.0292	-0.0137	0.0075	0.0288	0.0636	0.1151	0.1430	0.2002
Model 1 vs. Model 2	120	-0.1125	-0.1052	-0.0607	-0.0096	0.0859	0.2007	0.2347	0.2954
Model 1 vs. Model 3	60	-0.2747	-0.1433	-0.1537	0.0177	0.1505	0.2335	0.3257	0.5550
Model 1 vs. Model 3	120	-0.3508	-0.1871	-0.2026	0.0173	0.2387	0.3336	0.4390	0.6625
Model 1 vs. Model 4	60	-0.2752	-0.1180	-0.1424	-0.0050	0.1660	0.2410	0.3316	0.5550
Model 1 vs. Model 4	120	-0.3623	-0.1742	-0.1465	-0.0078	0.2049	0.3502	0.4158	0.6478
Model 2 vs. Model 4	60	-0.2390	-0.1029	-0.1510	-0.0348	0.1094	0.1423	0.2201	0.4436
Model 2 vs. Model 4	120	-0.2245	-0.0624	-0.0809	0.0018	0.1302	0.1870	0.2367	0.5002
					Rolling-estimation window (RV)				
Model 1 vs. Model 2	60	-0.0307	0.0632	0.0476	-0.0925	0.0394	0.1647	0.1663	0.1811
Model 1 vs. Model 2	120	-0.1260	-0.1958	-0.1944	-0.1789	-0.0083	0.2205	0.2617	0.2211
Model 1 vs. Model 3	60	-0.3640	-0.2058	-0.2206	-0.1029	0.1110	0.2372	0.4206	0.4731
Model 1 vs. Model 3	120	-0.6834	-0.1962	-0.2125	-0.2767	-0.0469	0.1695	0.4161	0.6246
Model 1 vs. Model 4	60	-0.3374	-0.1198	-0.2439	-0.1979	0.0914	0.2469	0.3973	0.4977
Model 1 vs. Model 4	120	-0.6418	-0.2424	-0.2008	-0.3318	-0.1434	0.1765	0.4099	0.6075
Model 2 vs. Model 4	60	-0.2975	-0.1953	-0.3060	-0.0964	0.0541	0.0984	0.2771	0.3866
Model 2 vs. Model 4	120	-0.4581	-0.0389	-0.0054	-0.1297	-0.1339	-0.0564	0.2007	0.4961
					Rolling-estimation window ($\sqrt{RV}$)				
Model 1 vs. Model 2	60	0.0049	0.0870	0.1096	0.0219	0.1085	0.2159	0.2251	0.2076
Model 1 vs. Model 2	120	-0.0933	-0.1175	-0.0818	-0.0306	0.0766	0.1968	0.2077	0.2068
Model 1 vs. Model 3	60	-0.2698	-0.1958	-0.2894	-0.1539	0.0875	0.2405	0.4270	0.5281
Model 1 vs. Model 3	120	-0.3280	-0.1534	-0.0661	-0.0296	0.1610	0.2711	0.4333	0.6114
Model 1 vs. Model 4	60	-0.3724	-0.2125	-0.1968	-0.0403	0.1484	0.2665	0.4487	0.5370
Model 1 vs. Model 4	120	-0.3133	-0.0960	-0.1076	-0.0616	0.1343	0.3023	0.4296	0.6294
Model 2 vs. Model 4	60	-0.3792	-0.3281	-0.3442	-0.0636	0.0449	0.0646	0.2885	0.4157
Model 2 vs. Model 4	120	-0.2012	0.0193	-0.0239	-0.0301	0.0625	0.1313	0.2801	0.5327

The out-of-sample R^2 statistic is computed as $R^2 = 1 - \sum e^2_{rival} / \sum e^2_{benchmark}$. The benchmark (rival) model is the first (second) model given in the first column of the table. A positive out-of-sample R^2 statistic shows that the rival model outperforms the benchmark model. The parameter h denotes the forecast horizon (in months).

Table 3: Configuration: Lasso, Absolute-Error-Loss Function, 2000/01−2019/12

Benchmark vs. rival model	Window	h=1	h=3	h=6	h=12	h=18	h=24	h=32	h=48
				Recursive-estimation window (RV)					
Model 1 vs. Model 2	60	-0.0239	-0.0148	-0.0029	0.0033	0.0208	0.0529	0.0861	0.0792
Model 1 vs. Model 2	120	-0.1489	-0.1138	-0.1022	-0.0755	-0.0002	0.1043	0.1830	0.1161
Model 1 vs. Model 3	60	-0.2487	-0.1768	-0.0959	0.0069	0.1130	0.1418	0.1756	0.2998
Model 1 vs. Model 3	120	-0.4122	-0.1951	-0.1559	0.0030	0.1333	0.1889	0.2394	0.3384
Model 1 vs. Model 4	60	-0.2694	-0.1796	-0.1010	0.0321	0.1015	0.1480	0.1691	0.2969
Model 1 vs. Model 4	120	-0.4041	-0.2039	-0.1691	-0.0403	0.0931	0.1957	0.2355	0.3632
Model 2 vs. Model 4	60	-0.2398	-0.1624	-0.0978	0.0289	0.0824	0.1005	0.0909	0.2364
Model 2 vs. Model 4	120	-0.2221	-0.0809	-0.0607	0.0328	0.0933	0.1021	0.0642	0.2796
				Recursive-estimation window ($\sqrt{RV}$)					
Model 1 vs. Model 2	60	-0.0234	-0.0136	0.0122	0.0150	0.0343	0.0566	0.0820	0.0872
Model 1 vs. Model 2	120	-0.0911	-0.0664	-0.0368	-0.0193	0.0351	0.0944	0.1445	0.1367
Model 1 vs. Model 3	60	-0.1823	-0.0987	-0.0793	0.0360	0.1174	0.1500	0.1771	0.3044
Model 1 vs. Model 3	120	-0.2220	-0.1066	-0.1268	0.0211	0.1585	0.1976	0.2427	0.3708
Model 1 vs. Model 4	60	-0.1939	-0.0925	-0.0834	0.0261	0.1182	0.1546	0.1857	0.3108
Model 1 vs. Model 4	120	-0.2240	-0.1141	-0.1066	0.0055	0.1333	0.2088	0.2335	0.3518
Model 2 vs. Model 4	60	-0.1665	-0.0779	-0.0967	0.0112	0.0869	0.1039	0.1130	0.2450
Model 2 vs. Model 4	120	-0.1219	-0.0447	-0.0674	0.0243	0.1017	0.1263	0.1040	0.2492
				Rolling-estimation window (RV)					
Model 1 vs. Model 2	60	-0.0324	-0.0414	0.0015	-0.0248	0.0327	0.1095	0.1109	0.1004
Model 1 vs. Model 2	120	-0.1473	-0.1358	-0.1180	-0.0893	0.0001	0.1087	0.1738	-0.0210
Model 1 vs. Model 3	60	-0.2490	-0.1817	-0.1827	-0.0412	0.1260	0.2032	0.2666	0.3457
Model 1 vs. Model 3	120	-0.4703	-0.1708	-0.1474	-0.0592	0.0931	0.1519	0.2733	0.3032
Model 1 vs. Model 4	60	-0.2156	-0.1707	-0.1812	-0.0926	0.1208	0.2013	0.2636	0.3527
Model 1 vs. Model 4	120	-0.4456	-0.1930	-0.1529	-0.0868	0.0342	0.1633	0.2710	0.2867
Model 2 vs. Model 4	60	-0.1774	-0.1241	-0.1829	-0.0662	0.0911	0.1032	0.1718	0.2805
Model 2 vs. Model 4	120	-0.2600	-0.0503	-0.0312	0.0023	0.0341	0.0613	0.1177	0.3013
				Rolling-estimation window ($\sqrt{RV}$)					
Model 1 vs. Model 2	60	-0.0165	0.0097	0.0781	0.0117	0.0530	0.1240	0.1294	0.1035
Model 1 vs. Model 2	120	-0.0859	-0.0780	-0.0414	-0.0312	0.0256	0.0957	0.1338	0.0682
Model 1 vs. Model 3	60	-0.1105	-0.1240	-0.1207	-0.0475	0.1109	0.1933	0.2834	0.3510
Model 1 vs. Model 3	120	-0.1840	-0.1010	-0.0570	0.0088	0.1406	0.1818	0.2503	0.3285
Model 1 vs. Model 4	60	-0.1450	-0.1267	-0.0809	-0.0047	0.1425	0.1988	0.2871	0.3604
Model 1 vs. Model 4	120	-0.1835	-0.0902	-0.0879	-0.0032	0.1143	0.2037	0.2657	0.3482
Model 2 vs. Model 4	60	-0.1264	-0.1377	-0.1725	-0.0166	0.0946	0.0854	0.1812	0.2866
Model 2 vs. Model 4	120	-0.0899	-0.0114	-0.0446	0.0271	0.0910	0.1194	0.1523	0.3004

The out-of-sample R^2 statistic is computed as $R^2 = 1 - \sum |e_{rival}| / \sum |e_{benchmark}|$. The benchmark (rival) model is the first (second) model given in the first column of the table. A positive out-of-sample R^2 statistic shows that the rival model outperforms the benchmark model. The parameter h denotes the forecast horizon (in months).

Table 4: Configuration: Random Forests, Squared-Error-Loss Function, 2000/01−2019/12

Benchmark vs. rival model	Window	h=1	h=3	h=6	h=12	h=18	h=24	h=32	h=48
					Recursive-estimation window (RV)				
Model 1 vs. Model 2	60	-0.0337	-0.0169	-0.0052	0.0045	0.0234	0.0736	0.1108	0.1907
Model 1 vs. Model 2	120	-0.1479	-0.1770	-0.1639	-0.1443	-0.0042	0.2094	0.2841	0.3298
Model 1 vs. Model 3	60	-0.2658	-0.3104	-0.1565	-0.0020	0.1214	0.1555	0.2776	0.4878
Model 1 vs. Model 3	120	-1.0832	-0.4312	-0.4856	-0.0984	0.0688	0.2995	0.4965	0.6833
Model 1 vs. Model 4	60	-0.1999	-0.3477	-0.1303	0.0340	0.1205	0.1492	0.2677	0.5060
Model 1 vs. Model 4	120	-1.0459	-0.5019	-0.5443	-0.0965	0.0701	0.2816	0.4977	0.6828
Model 2 vs. Model 4	60	-0.1608	-0.3253	-0.1245	0.0297	0.0994	0.0816	0.1765	0.3896
Model 2 vs. Model 4	120	-0.7823	-0.2761	-0.3268	0.0418	0.0740	0.0913	0.2983	0.5267
					Recursive-estimation window ($\sqrt{RV}$)				
Model 1 vs. Model 2	60	-0.0292	-0.0137	0.0075	0.0288	0.0636	0.1151	0.1430	0.2002
Model 1 vs. Model 2	120	-0.1125	-0.1052	-0.0607	-0.0096	0.0859	0.2007	0.2347	0.2954
Model 1 vs. Model 3	60	-0.3279	-0.3330	-0.2071	0.0262	0.1117	0.1845	0.2964	0.5103
Model 1 vs. Model 3	120	-0.7233	-0.3575	-0.3184	0.0161	0.0905	0.2402	0.4131	0.6200
Model 1 vs. Model 4	60	-0.3249	-0.3598	-0.1956	0.0610	0.1471	0.1846	0.2977	0.5157
Model 1 vs. Model 4	120	-0.7531	-0.3404	-0.3227	-0.0041	0.0959	0.2588	0.4157	0.6294
Model 2 vs. Model 4	60	-0.2873	-0.3414	-0.2046	0.0331	0.0892	0.0786	0.1805	0.3945
Model 2 vs. Model 4	120	-0.5758	-0.2128	-0.2470	0.0054	0.0109	0.0726	0.2366	0.4741
					Rolling-estimation window (RV)				
Model 1 vs. Model 2	60	-0.0307	0.0632	0.0476	-0.0925	0.0394	0.1647	0.1663	0.1811
Model 1 vs. Model 2	120	-0.1260	-0.1958	-0.1944	-0.1789	-0.0083	0.2205	0.2617	0.2211
Model 1 vs. Model 3	60	-0.5500	-0.4833	-0.2898	0.0139	0.2396	0.3142	0.4674	0.5581
Model 1 vs. Model 3	120	-1.1559	-0.5465	-0.5796	-0.1612	0.0034	0.2654	0.4849	0.6605
Model 1 vs. Model 4	60	-0.4780	-0.4488	-0.2599	0.0255	0.2614	0.3350	0.4614	0.5597
Model 1 vs. Model 4	120	-1.2483	-0.6287	-0.6134	-0.1412	0.0191	0.3056	0.4717	0.6625
Model 2 vs. Model 4	60	-0.4339	-0.5466	-0.3228	0.1081	0.2311	0.2039	0.3539	0.4624
Model 2 vs. Model 4	120	-0.9967	-0.3620	-0.3508	0.0319	0.0271	0.1092	0.2844	0.5667
					Rolling-estimation window ($\sqrt{RV}$)				
Model 1 vs. Model 2	60	0.0049	0.0870	0.1096	0.0219	0.1085	0.2159	0.2251	0.2076
Model 1 vs. Model 2	120	-0.0933	-0.1175	-0.0818	-0.0306	0.0766	0.1968	0.2077	0.2068
Model 1 vs. Model 3	60	-0.5873	-0.4928	-0.3860	-0.0410	0.1773	0.3061	0.4668	0.5873
Model 1 vs. Model 3	120	-0.7612	-0.4617	-0.4512	-0.1708	0.0344	0.2431	0.4209	0.5854
Model 1 vs. Model 4	60	-0.6306	-0.5702	-0.3826	-0.0098	0.1969	0.3314	0.4502	0.5886
Model 1 vs. Model 4	120	-0.7828	-0.4887	-0.5299	-0.0976	0.0973	0.2423	0.4335	0.5877
Model 2 vs. Model 4	60	-0.6387	-0.7198	-0.5528	-0.0324	0.0992	0.1472	0.2904	0.4808
Model 2 vs. Model 4	120	-0.6307	-0.3321	-0.4143	-0.0649	0.0224	0.0567	0.2850	0.4803

The out-of-sample R^2 statistic is computed as $R^2 = 1 - \sum e^2_{rival} / \sum e^2_{benchmark}$. The benchmark (rival) model is the first (second) model given in the first column of the table. A positive out-of-sample R^2 statistic shows that the rival model outperforms the benchmark model. The parameter h denotes the forecast horizon (in months).

Table 5: Configuration: Lasso, Squared-Error-Loss Function, $\sqrt{RV}$, Various Sample Periods

Benchmark vs. rival model	Window	h=1	h=3	h=6	h=12	h=18	h=24	h=32	h=48
		Recursive-estimation window (1984/01−2019/12))							
Model 1 vs. Model 2	60	-0.0316	-0.0297	-0.0194	-0.0039	0.0292	0.0849	0.1235	0.1402
Model 1 vs. Model 2	120	-0.0161	-0.0093	-0.0073	0.0096	0.0313	0.0666	0.1081	0.1827
Model 1 vs. Model 3	60	0.0011	-0.0023	-0.0075	-0.0100	0.0147	0.0623	0.0709	0.0604
Model 1 vs. Model 3	120	-0.0116	-0.0108	-0.0185	-0.0062	0.0420	0.0936	0.0921	0.1009
Model 1 vs. Model 4	60	-0.0058	-0.0140	-0.0225	-0.0138	0.0355	0.1051	0.1502	0.1520
Model 1 vs. Model 4	120	-0.0220	-0.0109	-0.0107	-0.0144	0.0587	0.1232	0.1743	0.2331
Model 2 vs. Model 4	60	0.0251	0.0153	-0.0031	-0.0099	0.0065	0.0221	0.0305	0.0137
Model 2 vs. Model 4	120	-0.0058	-0.0015	-0.0034	-0.0242	0.0282	0.0606	0.0742	0.0617
		Recursive-estimation window (1990/01−2019/12)							
Model 1 vs. Model 2	60	-0.0166	-0.0119	-0.0102	0.0106	0.0452	0.0872	0.1119	0.1522
Model 1 vs. Model 2	120	-0.0170	-0.0188	-0.0034	0.0001	0.0135	0.0441	0.0412	0.0462
Model 1 vs. Model 3	60	-0.1039	-0.1029	-0.0694	0.0437	0.1133	0.1855	0.2605	0.3478
Model 1 vs. Model 3	120	-0.0546	-0.0667	-0.0389	0.0411	0.1131	0.1767	0.2409	0.3003
Model 1 vs. Model 4	60	-0.0988	-0.0789	-0.0627	0.0320	0.1123	0.1993	0.2881	0.3834
Model 1 vs. Model 4	120	-0.0569	-0.0654	-0.0529	0.0337	0.0986	0.1636	0.2158	0.2685
Model 2 vs. Model 4	60	-0.0808	-0.0662	-0.0520	0.0216	0.0702	0.1229	0.1984	0.2727
Model 2 vs. Model 4	120	-0.0393	-0.0457	-0.0493	0.0336	0.0863	0.1250	0.1821	0.2331
		Rolling-estimation window (1984/01−2019/12)							
Model 1 vs. Model 2	60	-0.0250	0.0107	0.0275	0.0033	0.0636	0.1629	0.1904	0.1529
Model 1 vs. Model 2	120	-0.0234	-0.0026	-0.0044	-0.0093	0.0094	0.0482	0.0688	0.0976
Model 1 vs. Model 3	60	-0.0264	-0.0237	-0.0398	-0.0031	0.0460	0.0846	0.1192	0.1834
Model 1 vs. Model 3	120	-0.0213	-0.0311	-0.0460	0.0139	0.0573	0.0905	0.1309	0.1977
Model 1 vs. Model 4	60	-0.0498	-0.0352	-0.0070	0.0220	0.1073	0.2015	0.2148	0.2293
Model 1 vs. Model 4	120	-0.0289	-0.0418	-0.0401	0.0066	0.0536	0.1031	0.1586	0.2209
Model 2 vs. Model 4	60	-0.0242	-0.0464	-0.0355	0.0188	0.0466	0.0461	0.0301	0.0901
Model 2 vs. Model 4	120	-0.0054	-0.0391	-0.0356	0.0157	0.0446	0.0577	0.0965	0.1365
		Rolling-estimation window (1990/01−2019/12)							
Model 1 vs. Model 2	60	-0.0079	0.0502	0.0669	0.0260	0.0973	0.1866	0.1966	0.1905
Model 1 vs. Model 2	120	-0.0298	-0.0096	0.0017	-0.0081	0.0191	0.0727	0.0765	0.0732
Model 1 vs. Model 3	60	-0.1047	-0.0233	0.0216	0.0930	0.1302	0.2211	0.2692	0.3295
Model 1 vs. Model 3	120	-0.0567	-0.0721	-0.0413	0.0377	0.1386	0.2010	0.2462	0.3118
Model 1 vs. Model 4	60	-0.1290	-0.0192	0.0643	0.1129	0.2140	0.2814	0.3101	0.3547
Model 1 vs. Model 4	120	-0.0837	-0.0582	-0.0296	0.0302	0.1013	0.1766	0.2299	0.2555
Model 2 vs. Model 4	60	-0.1201	-0.0731	-0.0028	0.0892	0.1293	0.1166	0.1413	0.2028
Model 2 vs. Model 4	120	-0.0524	-0.0482	-0.0314	0.0380	0.0838	0.1121	0.1660	0.1967

The out-of-sample R^2 statistic is computed as $R^2 = 1 - \sum e^2_{rival} / \sum e^2_{benchmark}$. The benchmark (rival) model is the first (second) model given in the first column of the table. A positive out-of-sample R^2 statistic shows that the rival model outperforms the benchmark model. The parameter h denotes the forecast horizon (in months). The models estimated on data for the sample period 1984/01−2019/12 feature data for Texas as the only EPU variable. The models estimated on data for the sample period 1990/01−2019/12 feature data for Texas, Oklahoma, and Colorado as the EPU variables.

Table 6: Configuration: Lasso, Squared-Error-Loss Function, $\sqrt{RV}$, Point Forecasts

Benchmark vs. rival model	Window	h=1	h=3	h=6	h=12	h=18	h=24	h=32	h=48
		Recursive-estimation window (1984/01−2019/12))							
Model 1 vs. Model 2	60	-0.0316	-0.0070	-0.0038	0.0114	0.0204	0.0266	0.0037	-0.0054
Model 1 vs. Model 2	120	-0.0161	-0.0162	0.0006	0.0175	0.0185	0.0383	0.0139	0.0006
Model 1 vs. Model 3	60	0.0072	-0.0089	-0.0044	0.0235	0.0272	0.0466	-0.0199	0.0068
Model 1 vs. Model 3	120	-0.0144	-0.0044	-0.0099	0.0304	0.0482	0.0578	-0.0022	0.0008
Model 1 vs. Model 4	60	-0.0090	-0.0050	-0.0083	0.0324	0.0361	0.0673	-0.0183	0.0008
Model 1 vs. Model 4	120	-0.0067	-0.0090	-0.0043	0.0313	0.0445	0.0635	0.0020	-0.0031
Model 2 vs. Model 4	60	0.0220	0.0019	-0.0045	0.0213	0.0161	0.0417	-0.0221	0.0062
Model 2 vs. Model 4	120	0.0092	0.0071	-0.0049	0.0140	0.0264	0.0262	-0.0121	-0.0037
		Recursive-estimation window (1990/01−2019/12)							
Model 1 vs. Model 2	60	-0.0166	-0.0240	-0.0186	0.0245	0.0181	0.0350	0.0146	-0.0005
Model 1 vs. Model 2	120	-0.0170	-0.0023	-0.0148	0.0005	-0.0175	0.0209	-0.0214	-0.0291
Model 1 vs. Model 3	60	-0.0816	-0.0370	0.0052	0.0334	0.0683	0.0926	0.0256	0.0036
Model 1 vs. Model 3	120	-0.0485	-0.0313	0.0032	0.0741	0.0524	0.0372	0.0017	0.0062
Model 1 vs. Model 4	60	-0.1066	-0.0391	-0.0217	0.0448	0.0637	0.0749	0.0325	0.0092
Model 1 vs. Model 4	120	-0.0776	-0.0371	-0.0149	0.0609	0.0274	0.0380	-0.0018	-0.0217
Model 2 vs. Model 4	60	-0.0885	-0.0148	-0.0030	0.0208	0.0465	0.0414	0.0182	0.0098
Model 2 vs. Model 4	120	-0.0596	-0.0347	-0.0002	0.0604	0.0441	0.0175	0.0191	0.0072
		Rolling-estimation window (1984/01−2019/12)							
Model 1 vs. Model 2	60	-0.0250	0.0044	-0.0069	-0.0618	0.0396	-0.0364	-0.0054	0.0280
Model 1 vs. Model 2	120	-0.0234	-0.0150	-0.0062	-0.0011	-0.0121	0.0055	-0.0043	-0.0156
Model 1 vs. Model 3	60	-0.0244	-0.0900	0.0276	0.0391	0.0309	0.0288	0.0470	0.0216
Model 1 vs. Model 3	120	-0.0364	-0.0580	0.0209	0.0669	0.0231	0.0331	0.0211	0.0239
Model 1 vs. Model 4	60	-0.0519	-0.0573	0.0201	0.0084	0.0723	-0.0241	0.0449	0.0339
Model 1 vs. Model 4	120	-0.0614	-0.0423	-0.0050	0.0398	0.0310	0.0373	0.0103	0.0128
Model 2 vs. Model 4	60	-0.0262	-0.0619	0.0268	0.0661	0.0341	0.0118	0.0500	0.0061
Model 2 vs. Model 4	120	-0.0371	-0.0269	0.0012	0.0408	0.0426	0.0320	0.0146	0.0280
		Rolling-estimation window (1990/01−2019/12)							
Model 1 vs. Model 2	60	-0.0079	0.0117	-0.0029	-0.0779	0.0606	-0.0598	0.0193	0.0363
Model 1 vs. Model 2	120	-0.0298	0.0076	-0.0041	-0.0073	-0.0119	-0.0033	-0.0036	-0.0352
Model 1 vs. Model 3	60	-0.0794	-0.0197	0.0237	0.0357	0.0790	0.0493	0.0430	0.0166
Model 1 vs. Model 3	120	-0.0402	-0.0275	0.0210	0.0772	0.0461	0.0180	0.0011	0.0191
Model 1 vs. Model 4	60	-0.1057	0.0035	0.0132	0.0043	0.1072	0.0310	0.0367	0.0231
Model 1 vs. Model 4	120	-0.1080	-0.0407	-0.0234	0.0403	0.0525	0.0080	-0.0006	-0.0062
Model 2 vs. Model 4	60	-0.0971	-0.0083	0.0161	0.0763	0.0496	0.0857	0.0177	-0.0137
Model 2 vs. Model 4	120	-0.0759	-0.0488	-0.0192	0.0472	0.0636	0.0113	0.0030	0.0280

The out-of-sample R^2 statistic is computed as $R^2 = 1 - \sum e^2_{rival} / \sum e^2_{benchmark}$. The benchmark (rival) model is the first (second) model given in the first column of the table. A positive out-of-sample R^2 statistic shows that the rival model outperforms the benchmark model. The parameter h denotes the forecast horizon (in months). The models estimated on data for the sample period 1984/01−2019/12 feature data for Texas as the only EPU variable. The models estimated on data for the sample period 1990/01−2019/12 feature data for Texas, Oklahoma, and Colorado as the EPU variables.