Energy Economics 132 (2024) 107432

Contents lists available at ScienceDirect

Energy Economics

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

# Stock market bubbles and the realized volatility of oil price returns☆

Rangan Gupta [a,*], Joshua Nielsen [b], Christian Pierdzioch [c]

[a] Department of Economics, University of Pretoria, Private Bag X20, Hatfield 0028, South Africa
[b] Boulder Investment Technologies, LLC, 1942 Broadway Suite 314C, Boulder, CO, 80302, USA
[c] Department of Economics, Helmut Schmidt University, Holstenhofweg 85, P.O.B. 700822, 22008 Hamburg, Germany

ARTICLE INFO

*JEL classification:*
C22
C53
G15
Q02

*Keywords:*
Realized volatility
Oil price
Stock market bubbles
Forecasting
Shrinkage estimators

ABSTRACT

Using monthly data for the G7 countries from 1973 to 2020, we study whether stock market bubbles help to forecast out-of-sample the realized volatility of oil price returns. We use the Multi-Scale Log-Periodic Power Law Singularity Confidence Indicator (MS-LPPLS-CI) approach to identify both positive and negative bubbles in the short-, medium, and long-term. First, we successfully detect major crashes and rallies using the MS-LPPLS-CIs. Having established the relevance of the bubbles indicators, and given the large number of them, we use widely-studied shrinkage (Lasso, elastic net, ridge regression) approaches to estimate our forecasting models. We find that stock market bubbles have predictive value for realized volatility at a short to intermediate forecast horizon. The number of bubble predictors included in the penalized forecasting models tend to increase in the forecast horizon. We obtain our main finding for the various types of stock market bubbles, and for good and bad realized volatilities.

## 1. Introduction

We forecast the monthly realized volatility of returns of the West Texas Intermediate (WTI) oil price, over the period of 1973:01 to 2020:09, based on the information content of indicators of stock market bubbles of the G7 countries (i.e., Canada, France, Germany, Italy, Japan, the United Kingdom (UK), and the United States (US)). The predictive value of stock market bubbles for explaining oil market volatility can emanate through multiple channels.

First, it is now theoretically (Biswas et al., 2020) and empirically (Reinhart and Rogoff, 2009; Brunnermeier and Oehmke, 2013; Jordà et al., 2015) well-established that bursting of bubbles leads to severe recessions and major economic losses. At the same time, recoveries or rallies of stock markets tend to be associated with better performance of the economy in terms of growth (Caraiani et al., 2023). The resulting up- and downswings of economic activity, in turn, can have direct repercussions on oil-price movements.

Second, stock market bubbles and the ensuing economic booms and busts can increase macroeconomic uncertainty. As initially pointed out by Bernanke (1983), and more recently by Ludvigson et al. (2021), economic conditions are negatively associated with macroeconomic uncertainty. Rising (falling) macroeconomic uncertainty, in turn, can cause risk-averse commodity producers to increase (reduce) holding of physical inventory, as future cash flows are expected to be negatively (positively) impacted (Gupta and Pierdzioch, 2023). This can result in a rise (fall) in the convenience yield, which, in turn, will produce increased (decreased) oil market volatility (Gupta and Pierdzioch, 2021a; Gupta and Pierdzioch, 2022; Çepni et al., 2022). In other words, a stock market crash (recovery) can raise (reduce) uncertainty so as to cause oil market volatility to rise (fall) due to deteriorating (improved) economic conditions.

Third, the effect of stock market bubbles can feed into the volatility of the oil market via a "leverage effect" (as initially established by Geman and Shih (2009) and, more recently, stressed by Asai et al. (2019)), when one accounts for the high degree of correlation between stock and oil returns. The original idea of the leverage effect, proposed for the stock market by Black (1976), implies that negative (positive) returns are often associated with upward (downward) revisions of volatility. Aboura and Chevallier (2013), however, point out that it is also possible to obtain an increase in the volatility subsequent to a hike in the oil price as it may reflect that oil consumers fear a rising oil price. Irrespective of the underlying nature of the leverage effect, and the sign of the correlation, which seems to have changed from the historically negative one to a positive value since the turn of the century, and consistently so after the Global Financial Crisis due to

☆ We would like to thank two anonymous referees, and the Editor, Professor Richard S.J. Tol, for many helpful comments. However, any remaining inaccuracies are solely ours.

* Corresponding author.
*E-mail addresses:* rangan.gupta@up.ac.za (R. Gupta), josh@boulderinvestment.tech (J. Nielsen), macroeconomics@hsu-hh.de (C. Pierdzioch).

https://doi.org/10.1016/j.eneco.2024.107432
Received 23 August 2023; Received in revised form 4 February 2024; Accepted 26 February 2024
Available online 27 February 2024

the process of financialization (Antonakakis et al., 2017), this potential channel of stock market bubbles impacting oil returns volatility cannot be ignored. For example, declining (increasing) stock returns due to a crash (recovery) would translate into decreasing or increasing (increasing or decreasing) oil returns, depending on the part of sample period being analyzed, which would then result in a increase or a decrease (a decrease or an increase) of oil returns volatility, contingent on the traditional (non-traditional) definition of the leverage effect.

Fourth, stock market bubbles are known to be driven by investor sentiment (Pan, 2020; van Eyden et al., 2023), and with a more financialized oil market, implying an increased importance of the financial sector and a more important role of financial motives, speculation is likely to have become a common factor driving the higher interconnectedness of returns between stocks and crude oil. The interconnectedness of the two returns implies that stock market bubbles could impact volatility of the oil market either directly or indirectly when stock market bubbles work their way to the oil market via the leverage effect.

Finally, a vast inflow of institutional investors in the oil market, especially over the last two decades, has led to increasing integration of the oil and stock markets such that portfolio rebalancing of index investors can result in volatility spillovers, caused by stock market bubbles through the leverage channel (strong evidence of which can be found in Goswami et al. (2020)) to the oil market (Bampinas and Panagiotidis, 2017).

It is well-known from the extensive research on speculative bubbles in financial markets that bubbles are notoriously difficult to detect. For this reason, we not only use the advanced Log-Periodic Power Law Singularity (LPPLS) model, tracing back to the work by Johansen et al. (1999, 2000) and Sornette (2003), to recover positive (upward accelerating price followed by a crash) as well as negative (downward accelerating price followed by a rally) bubbles, but we also apply the multi-scale LPPLS confidence indicators recently studied by Demirer et al. (2019) to characterize positive and negative bubbles at short-, medium- and long-term time scales.[1] In this regard, it is important to emphasize that both positive and negative multi-scale bubbles cannot be detected based on other statistical tests (see, Balcilar et al. (2016), Zhang et al. (2016), and Skrobotov (2023) for detailed reviews). We consider the two characteristics (nature and time-scale) of bubbles important for two reasons. First, these characteristics render it possible to trace out the potential differential predictive effects for the realized volatility of oil price returns of positive and negative news, resulting from positive and negative bubbles, especially in the context of the leverage-effect-related channel outlined above. Second, the two characteristics capture the idea that crashes and recoveries at alternative horizons can convey differential information for different types of traders, as implied by the Heterogeneous Market Hypothesis (HMH; Müller et al., 1997). The HMH stipulates that different types of traders (i.e., investors, speculators, and traders) populate financial markets, and that the different types of traders differ with regard to their perception of information flows at different time horizons. For example, speculators may be more likely to pay attention particularly to short- and medium-term bubbles, whereas investors possibly mainly react to long-term bubbles. In other words, a disaggregation into sign and time-scale of stock market bubbles is likely to prevent loss of predictive signals due to aggregation.

We model in terms of the realized volatility as the square root of the sum of daily squared returns over a month (based on the research by Andersen and Bollerslev, 1998), which has the advantage that it is an observable and unconditional metric of volatility — an otherwise latent process. Traditionally, researchers have analyzed the time-varying volatility of oil-price returns using various models belonging to the generalized autoregressive conditional heteroskedasticity (GARCH)-family. The characteristic feature of models belonging to the GARCH-family is that the conditional variance is a deterministic function of model parameters and past data. In more recent research, some researchers have studied stochastic volatility (SV) models, which imply that volatility is a latent variable governed by some stochastic process (for further details, see Chan and Grant (2016) and Lux et al. (2016)). Irrespective of whether one uses a GARCH or a SV model, one obtains an estimate of volatility that is not unconditional (model-free), as it is in the case with realized volatility.

We also must stress that, while oil is a global commodity, we consider the WTI as our proxy for the world oil price, not only due to consistent availability of data to associate with major episodes of bubbles covering nearly the last 50 years, but also because the US is a major player on both the demand- and the supply-front of the oil market. In this regard, our decision to analyze the G7 stock market bubbles should also be understandable, because the G7 countries account for nearly two-third of global net wealth and nearly half of world output, and, as such, extreme bubble-induced movements in the stock markets of the G7 countries are likely to trigger effects that radiate on a global scale to both the real economy and financial and commodity markets (Das et al., 2019).

In terms of our empirical research strategy, we rely on a standard linear forecasting regression framework, with the benchmark being an autoregressive model of realized volatility of oil returns. In order to separate out the influence of stock market bubbles on oil market developments, required to validate our lines of economic intuition, we augment our benchmark model with structural oil shocks, known to drive the oil price, following the seminal contribution of Kilian (2009), as well as oil returns realized volatility (Demirer et al., 2020). In this regard, we consider oil supply shocks, economic activity shocks, oil-specific consumption demand shocks, and oil inventory demand shocks.

When we add the 42 (six each: positive and negative at short-, medium-, and long-horizon, for seven countries) bubbles indicators to our forecasting model and also the shocks, we can be absolutely sure to capture the additional effect of the booms and busts of the G7 stock markets on oil price volatility via its returns. In addition, the global demand and oil-specific consumption demand shocks, make it possible to filter out the economic activity and speculative channels via which bubbles can impact oil market volatility, as outlined above. In other words, including the four structural oil shocks in our model as control variables is strongly warranted in terms of appropriately isolating the predictive role of stock market bubbles on the realized volatility of oil returns. The oil shocks data, based on the work of Baumeister and Hamilton (2019), is available from 1975:02, but because we detect an important crash episode in 1973 (from when our bubbles indicators start), as part of the collapse of the Bretton Woods system, we also conduct a forecasting analysis starting in 1973:01 by using oil returns instead of the oil shocks. While it is understandable that oil returns is an imperfect substitute of the oil shocks, as it might be masking many of the channels involving the bubbles-oil volatility nexus, it does serve as a valuable control in the context of our forecasting experiment.

In terms of econometrics, while we rely on the standard Ordinary Least Squares (OLS) estimator to estimate our benchmark forecasting model, and the augmented one with the four shocks, we use a shrinkage estimator known as the least absolute shrinkage and selection operator (Lasso) estimator (Tibshirani, 1996) to analyze the predictive value of the 42 bubbles indicators. The idea underlying this shrinkage estimator (as well as its relatives) is to reduce the dimension of a regression model in a data-driven manner to improve the accuracy of predictions derived from the penalized model. The relatives of the Lasso estimator that we study in our empirical research are the elastic net estimator and the Ridge regression estimator, two other widely studied shrinkage estimators.

[1] The short-, medium-, and long-term corresponds to estimation windows associated with trading activities covering one- to three-months, three-months to a year, and one- to two-years, respectively.

Our analysis stretches beyond academic value, and our findings should be of importance to investors as well as policymakers, because the trend towards increasing financialization of the oil and other commodity markets (Tang and Xiong, 2012; Büyükşahin and Robe, 2014), has resulted in increased market participation of institutional investors like hedge funds, pension funds, and insurance companies. In fact, crude oil is now considered a profitable "alternative" investment in the portfolios of financial institutions (Degiannakis and Filis, 2017).

In light of this, in 2022, the world produced an average of 80.75 million barrels of oil per day (including condensates), such that annual crude oil production reached approximately 29.5 billion barrels, with the market size exceeding $2 trillion at current prices (US Energy Information Administration (EIA); BP Statistical Review of World Energy), making oil the by far most actively traded commodity.[2] Because standard finance theory implies that the volatility of price movements is an important input to investment decisions and portfolio choices (Poon and Granger, 2003), it is clear that accurate forecasts of the realized volatility of oil-price returns are of key interest to oil traders. At the same time, the variability of oil-price returns has been shown to be associated with slowdowns in worldwide economic growth (van Eyden et al., 2019; Salisu et al., 2021). It follows that precise estimates of future movements of the realized volatility of oil price returns should also be important for policymakers, who need to design macroeconomic policies ahead of time to prevent the possibility of deep economic recessions.

To the best of our knowledge, this is the first paper to analyze the role of stock market bubble indicators of the G7 for forecasting the realized volatility of WTI oil-price returns, using statistical learning techniques (Lasso, ridge regression and elastic net), besides the standard OLS estimator applied to the smaller models involving only a lag of the dependent variable, and structural oil shocks. In this way, we contribute to the already existing significant literature on forecasting of oil-returns volatility by studying the role of crashes and recoveries of major international equity markets, with existing studies having considered the predictive value of a large number of macroeconomic, financial, behavioral, and climate patterns-related variables based on a plethora of modeling techniques (see, for example, Asai et al. (2020), Bonato et al. (2020), Bouri et al. (2020, 2021), Gkillas et al. (2020), Gupta and Pierdzioch (2021b), Demirer et al. (2022), Salisu et al. (2022), and Luo et al. (2022), and the references cited within these papers).

We structure the remaining sections of this research as follows: In Section 2, we provide a description of the data we used in our study, along with the detected bubbles based on the multi-scale LPPLS confidence indicators (MS-LPPLS-CI) approach. We outline in Section 3 our forecasting models and discuss our empirical results in Section 4. Finally, we conclude in Section 5.

## 2. Data

Given the restriction imposed by the availability of data, our sample period cover the monthly period of 1973:01 to 2020:09. We derive our metric of the realized volatility of oil-price returns in several steps. First, we compute the daily log-returns, that is, the first-difference of the natural logarithm of the WTI oil price. The daily WTI crude oil nominal price data is derived from Global Financial Data (GFD).[3] Second, we take the sum of the daily squared log-returns over a specific month. Third, we take the square root of this sum. Fourth and finally, due to a large peak of realized volatility at the end of the sample period, which is associated with the outbreak of the COVID-19 pandemic, we work with the natural logarithm, $\ln(1 + RV)$, as this also avoids zero entries in the early part of the sample. We plot the resulting metric of realized volatility in Fig. 1.

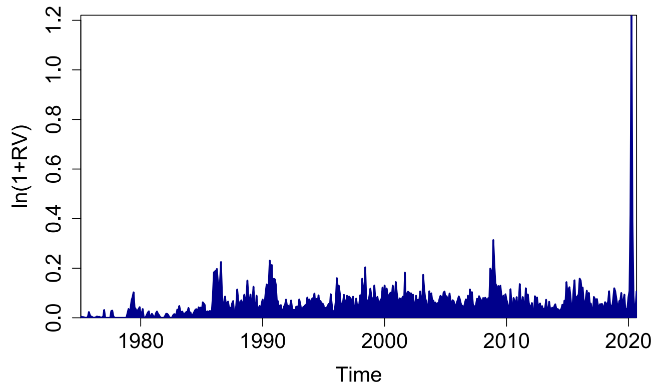


**Fig. 1.** Realized volatility.

We control for monthly log-returns of the WTI price, which starts in 1973:01, also obtained from GFD. Because it is now well-established, however, that oil-price movements are driven by structural shocks, as discussed in Section 1, our main analysis primarily focuses on four monthly shocks to separate out the effect of the stock market bubbles. We account for the following four shocks: an oil supply shock, an economic activity shock, an oil-specific consumption demand shock, and an oil inventory demand shock, with the data on shocks obtained from the structural vector autoregressive (SVAR) model of Baumeister and Hamilton (2019),[4] whose framework is less restrictive than what has been predominantly used in the recent literature following the research by Kilian (2009). The modeling framework proposed by Baumeister and Hamilton (2019) incorporates uncertainty about the identifying assumptions of the SVAR model. Their oil shocks, thus, are relatively more accurately estimated, with each of the shocks accounting for a distinct aspect of the demand for and supply of oil, that is, the information conveyed by the shocks do not overlap. The data on the structural shocks start from 1975:02, and are depicted in Fig. 2.

Turning to our main predictors, we compute first weekly bubble indicators, which we derive by studying the natural logarithmic values of the daily dividend–price ratio of the G7 countries, using the (local currency) dividend and the stock price index series, provided by Refinitiv Datastream. The estimated bubbles indicators cover the weekly sample period from the 1st week of (7th) January, 1973 to 2nd week of (13th) September, 2020. Because our dependent variable is measured at the monthly frequency, in order to obtain a monthly estimate for every multi-scale confidence indicator, we compute the mean of the weekly values for each of the scales for a given month. For the details underlying how we estimate the LPPLS model, and obtain the MS-LPPLS-CIs, the technically-minded reader is referred to the end of this paper (Appendix), where we describe further details of the model.

The evolution of the MS-LPPLS-CIs can be used to detect crashes and rallies in real-time. To this end, we plot the short-, medium-, and long-term indicators (blue, red, and yellow lines), while we show the log price-to-dividend ratio as a black line in Fig. 3. A larger LPPLS-CI value for a particular scale shows that the LPPLS signature is present for many of the fitting windows to which we calibrated the model, making it a more reliable bubble indicator. The key message conveyed by Fig. 3 is

[2] Iron Ore comes a distant second with 2.6 billion tonnes of annual production in 2022, and an associated market size of $283.4 billion. In fact, the combined market size of the top 10 metal markets amounts to $967 billion, which is still less than half that of the oil market.

[3] See the following internet page: https://globalfinancialdata.com/. Since daily data starts from January, 1977, for the period 1973 to 1976, we use the square root of the squared log-returns.

[4] The data is downloadable from the internet page of Professor Christiane Baumeister at: https://sites.google.com/site/cjsbaumeister/research.

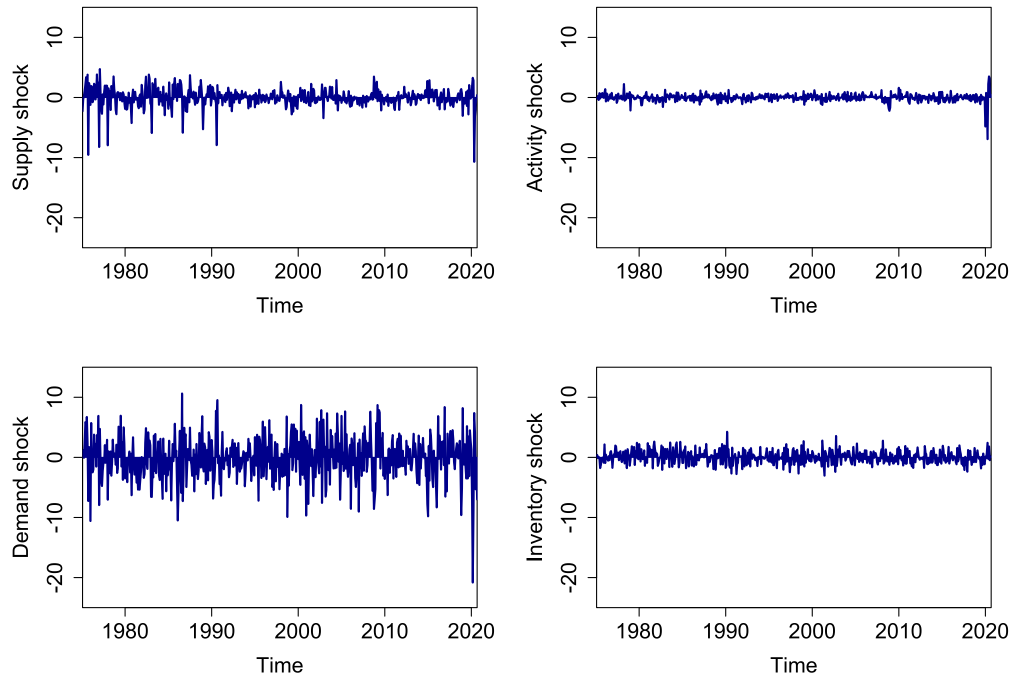


**Fig. 2.** Shocks.

that there are many peaks in the LPPLS-CIs preceding substantial shifts in the log price-to-dividend ratio. We note that the bubble indicators across the G7 countries, in general, display peaks in periods of time corresponding to crashes and recoveries before and around the collapse of the Bretton Woods system in 1973, the "Black Monday" episode in 1987, the Asian Financial Crisis of 1997, the Dot-com bubble burst from 2000 to 2002, the Global Financial Crisis of 2007 to 2008, the European sovereign debt crisis from 2009 to 2012, the "Brexit" in 2016, and to some extent during the COVID-19 episode as well. In other words, the MS-LPPLS-CIs are capable of providing leading information on all the major episodes of booms and busts witnessed globally over 1973 to 2020.

In general, smaller crashes or rallies can best be recovered using shorter time-scales, while longer time-scales help to detect larger crashes or rallies, with the short-term LPPLS-CIs preceding the medium-term ones, with the latter leading the long-run indicators, i.e., maturation of the bubble heading towards instability is present across several distinct time-scales. More importantly, we observe a similar timing of the strong (positive as well as negative) MS-LPPLS-CI values in the cross-section of G7 countries, in line with the intuition that boom and bust cycles of the seven developed equity markets often occur in tandem, motivating the need to combine the information content of the bubbles to capture fully the potential effect of financial market collapses and recoveries on the accuracy of forecasts of the realized volatility of oil price returns.

## 3. Forecasting models

We consider two benchmark models: the AR-CORE model and the AR-SHOCKS model, given by the following regression equations:

$$RV_{t+h} = \beta_0 + \beta_1 RV_t + u_{t+h}, \tag{1}$$

$$RV_{t+h} = \beta_0 + \beta_1 RV_t + \beta_2 S_t + u_{t+h}, \tag{2}$$

where $h$ denotes the forecast horizon (in months), $RV_t$ denotes realized volatility in period of time $t$, $RV_{t+h}$ denotes the average of realized volatility between period of time $t$ and period of time $t+h$, $u_{t+h}$ denotes the usual disturbance term, $S_t$ denotes a vector of (oil supply, economic activity, oil-specific consumption demand, and oil inventory demand) shocks, and $\beta_j, j = 0, 1$ ($\beta_2$) denote coefficients (a vector of coefficients) to be estimated. The AR-CORE model and the AR-SHOCKS model feature a small number of predictor variables and can easily be estimated by the OLS estimator.

The extended rival forecasting models are given by the following AR-CORE-BUBBLES and the AR-SHOCKS-BUBBLES models:

$$RV_{t+h} = \beta_0 + \beta_1 RV_t + \beta_3 B_t + u_{t+h}, \tag{3}$$

$$RV_{t+h} = \beta_0 + \beta_1 RV_t + \beta_2 S_t + \beta_3 B_t + u_{t+h}. \tag{4}$$

While the two rival models can in principle also be estimated by the OLS estimator, such an estimation approach does not take into account that the vector, $B_t$, of stock market bubbles has many components. With seven (G7) countries in our sample, and short-term, medium-term, and long-term bubble predictors, with all three categories further split into positive and negative bubbles, we have a total of $7 \times 3 \times 2 = 42$ bubble predictors. Adding the autoregressive predictor and the four shocks gives a total of 47 predictors (excluding the intercept term).

Given the large number of predictors, we use a shrinkage estimator to estimate the forecasting models given in Eqs. (3) and (4). One popular shrinkage estimator is the least absolute shrinkage and selection operator (Lasso) developed by Tibshirani (1996). In the case of the AR-SHOCKS-BUBBLES model, the Lasso estimator leads to a parsimonious (penalized) forecasting model by minimizing the following expression (for a textbook exposition, see Hastie et al. (2009)):

$$\sum_{t=1}^{T} \left( RV_{t+h} - \beta_0 - \beta_1 RV_t - \beta_2 S_t - \beta_3 B_t \right)^2 + \lambda \left( \sum_{i=1}^{n} |\beta_i| \right), \tag{5}$$

where $T$ denotes the number of observations used to estimate the forecasting model, $\lambda$ denotes a shrinkage parameter, and $n$ denotes the number of coefficients that are subject to the shrinkage process. Hence, the Lasso estimator adds to the standard quadratic loss function that forms the foundation of the OLS estimator a penalty term that increases in the absolute value of the coefficients. The Lasso estimator, thereby, implies that it is preferable to select coefficients that are small in absolute value or even zero, where the effect of this shrinking of coefficients towards zero must be balanced against the resulting effect on the quadratic loss function.

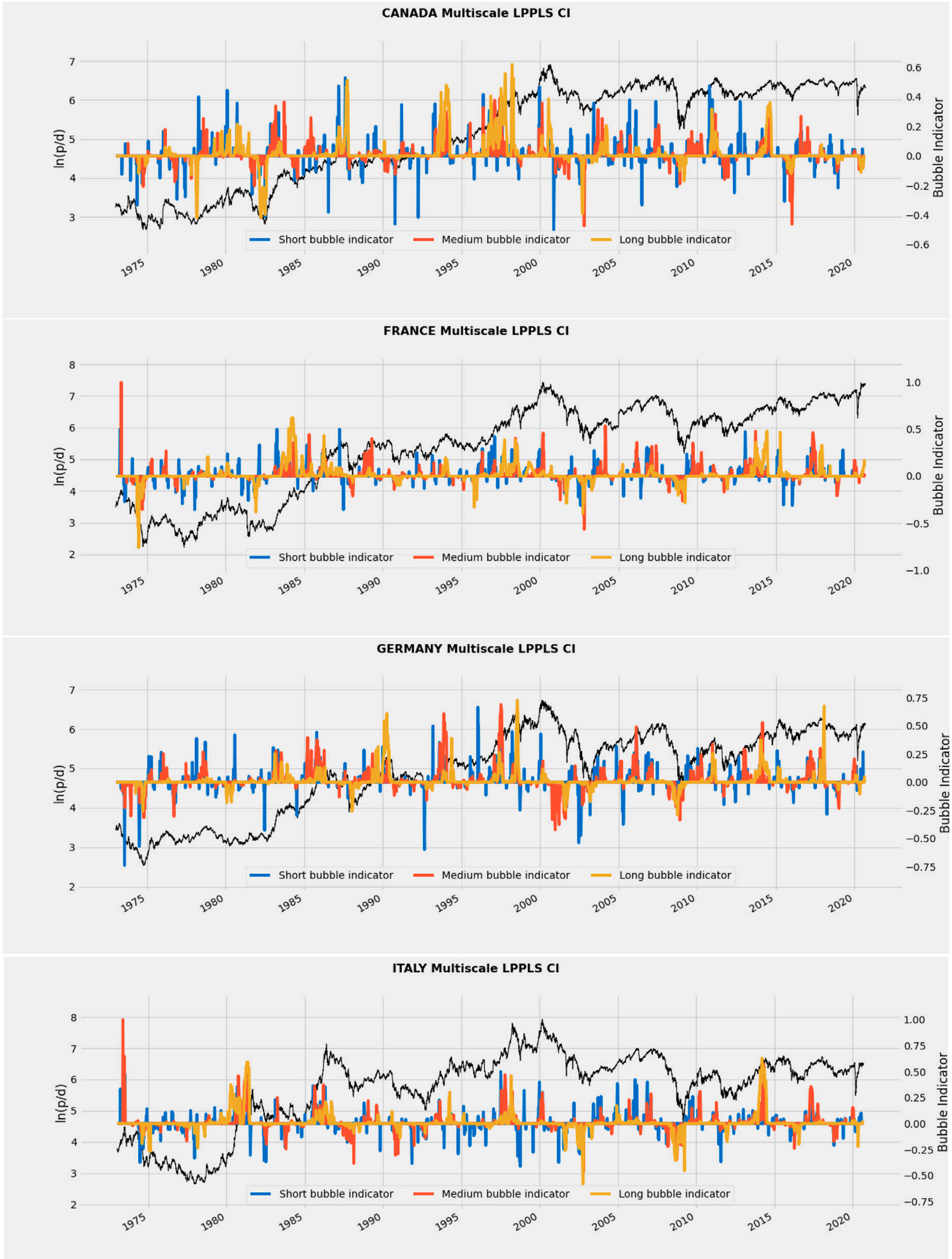


**Fig. 3.** G7 Monthly Mulit-Scale LPPLS-CI. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Eq. (6) can be easily modified to obtain two other popular shrinkage estimators. The first alternative shrinkage estimator is the so-called ridge regression estimator. This estimator uses the squared coefficients rather than the absolute coefficients in the penalty term of Eq. (6). The second alternative shrinkage estimator is the elastic net estimator, which uses a weighted average of the penalty terms that form the


**Fig. 3.** (*continued*).

foundation of the Lasso and Ridge regression estimators, with some mixing parameter $0 \leq m \leq 1$ (such that the Lasso estimator obtains as a special case for $m = 1$). In our empirical analysis, we shall mainly set $m = 0.5$ when we study the elastic net estimator, but we shall also present results for an alternative numerical value of the mixing parameter. Formally, the three shrinkage estimators (that is, the Lasso estimator, the Ridge regression estimator, and the elastic net estimator) obtain upon rewriting Eq. (6) in a generalized format as follows:

$$\sum_{t=1}^{T} \left( RV_{t+h} - \beta_0 - \beta_1 RV_t - \beta_2 S_t - \beta_3 B_t \right)^2 + \lambda \left( m\|\beta\|_1 + (1-m)\|\beta\|_2^2/2 \right), \tag{6}$$

where $\beta$ (without the index, $i$) denotes the respective vector of coefficients to be estimated (excluding the constant), and $\|.\|$ is the usual norm notation. We estimate our forecasting models on four different estimation training-data windows covering 50% up to 75% of the data, and use the remaining test data to compute forecasts of realized volatility. For example, a training-data window covering 50% of the data means that we use, starting at the beginning of the sample period, the first half of the data to estimate our forecasting models, and the remaining second half of the data to obtain forecasts. Similarly, a training-data window of 75% means that we use the first 75% of the data for estimation, and the remaining 25% for testing. Studying different training-data windows shows that our results are robust to the specific choice of the training and test sample. We subtract the actual realizations of realized volatility from the forecasts to obtain the forecast error, which we then scale by the actual realizations of realized volatility to obtain the relative forecast error. We use the relative forecast error because, as we show in Fig. 1, realized volatility exhibits a very large outburst at the end of the sample period. Using the relative forecast error mitigates the effect of this clearly atypical period of very large realized volatility on our results. Equipped with the relative forecast error, we compute the root-mean-squared-relative-forecast error (RMSRFE) for the benchmark and the rival forecasting models, which we then combine to form a RMSRFE ratio. A RMSRFE ratio that exceeds unity shows that the rival model performs better than the benchmark model. In addition, we use the modified (Diebold and Mariano, 1995) test (Harvey et al., 1997) to shed light on the statistical significance of our results.[5] In the context of our empirical analysis, the Diebold–Mariano test has the advantage that we can easily apply the test to compare forecasts under different loss functions. In our empirical analysis, we study the case of a quadratic and an absolute-error loss function. The latter is less sensitive to large forecast errors, which clearly could be an issue give that realized volatility (even though we study the natural logarithm) displayed large fluctuations especially at the end of our sample period (see Fig. 1). In this regard, we should also mention that we apply the test to study relative forecast errors (that is, forecast errors scaled by actual realizations of realized volatility).[6]

We use the R language and environment for statistical computing (R Core Team, 2023; version 4.3.0) the R add-on package "glmnet" (Friedman et al., 2010; version 4.1–7) to estimate our forecasting models. We use 10-fold cross-validation to determine the optimal shrinkage parameter (the one that minimizes the mean cross-validated error).

## 4. Empirical results

We set the stage for our empirical analysis by summarizing in Table 1 the number of bubble predictors included in the penalized forecasting models, where we use the Lasso estimator, the elastic net estimator, and the Ridge regression estimator to select among all predictors and, thus, to obtain a parsimonious penalized version of the AR-CORE-BUBBLES model. We present results for five different forecast horizons and four different training windows. The results show that the number of bubble predictors included in the penalized forecasting model under the Lasso and elastic net estimators tends to increase in the forecast horizon, whereas the penalized forecasting models selected by means of the Ridge regression estimator always include all bubble predictors. Hence, as expected, the elastic net strikes a balance between the Lasso estimator and the Ridge regression estimator in terms of the parsimony of the penalized forecasting models. Moreover, the penalized forecasting models estimated by means of the Lasso estimator, in the majority of model configurations, do not include bubble predictors at all for the one-month forecast horizon, while the forecasting models selected by means of the elastic net estimator typically feature a few bubble predictors also at the short forecast horizon. It follows that for the elastic net estimator (and the Ridge regression estimator) the selected bubble predictors contribute to forecast accuracy also at the one-month-ahead horizon.[7]

In Table 2, we compare the three estimators in terms of the RMSRFE ratio. The RMSRFE ratio is computed by dividing the RMSRFE of the first model reported in the first column of the table model by the RMSRFE (of the second model reported in the first column. A RMSRFE ratio larger than unity, thereby, indicates that the second model is superior relative to the first model. The results of the model comparison show that the elastic net estimator tends to dominate the Lasso estimator for

**Table 1**
Inclusion of bubble predictors in the forecasting model.

| Panel A: AR core model is the benchmark model | | | | | |
|---|---|---|---|---|---|
| Models | h = 1 | h = 2 | h = 3 | h = 6 | h = 12 |
| | Lasso estimator | | | | |
| w = 0.5 | 1 | 2 | 9 | 28 | 27 |
| w = 0.58 | 0 | 0 | 10 | 22 | 25 |
| w = 0.67 | 0 | 8 | 8 | 16 | 21 |
| w = 0.75 | 2 | 5 | 11 | 25 | 28 |
| | Elastic net estimator | | | | |
| w = 0.5 | 2 | 13 | 10 | 26 | 29 |
| w = 0.58 | 3 | 9 | 12 | 18 | 25 |
| w = 0.67 | 3 | 9 | 5 | 17 | 22 |
| w = 0.75 | 4 | 11 | 15 | 26 | 25 |
| | Ridge regression estimator | | | | |
| w = 0.5 | 42 | 42 | 42 | 42 | 42 |
| w = 0.58 | 42 | 42 | 42 | 42 | 42 |
| w = 0.67 | 42 | 42 | 42 | 42 | 42 |
| w = 0.75 | 42 | 42 | 42 | 42 | 42 |
| Panel B: AR plus shocks model is the benchmark model | | | | | |
| Models | h = 1 | h = 2 | h = 3 | h = 6 | h = 12 |
| | Lasso estimator | | | | |
| w = 0.5 | 2 | 2 | 8 | 21 | 29 |
| w = 0.58 | 0 | 0 | 1 | 20 | 24 |
| w = 0.67 | 0 | 0 | 8 | 16 | 25 |
| w = 0.75 | 2 | 2 | 13 | 25 | 30 |
| | Elastic net estimator | | | | |
| w = 0.5 | 2 | 6 | 15 | 27 | 22 |
| w = 0.58 | 0 | 8 | 11 | 18 | 25 |
| w = 0.67 | 1 | 8 | 9 | 19 | 22 |
| w = 0.75 | 4 | 8 | 13 | 26 | 27 |
| | Ridge regression estimator | | | | |
| w = 0.5 | 42 | 42 | 42 | 42 | 42 |
| w = 0.58 | 42 | 42 | 42 | 42 | 42 |
| w = 0.67 | 42 | 42 | 42 | 42 | 42 |
| w = 0.75 | 42 | 42 | 42 | 42 | 42 |

The AR-CORE and AR-SHOCKS forecasting models are estimated by the OLS technique. The AR-SHOCKS-BUBBLES model, which includes all bubbles as potential predictors, is estimated by the Lasso estimator, elastic net estimator, and the Ridge regression estimator. The estimations are done for four different windows of training data covering 50% up to 75% of the data, as indicated by the parameter $w$. The forecasting models are used to compute forecasts of the remaining test data. The numbers in the table summarize how many bubble predictors the penalized forecasting models include. $h$ denotes the forecast horizon (in months).

forecast horizons up to and including three months, and for some model configurations also for the two long forecast horizons. The results for the two long forecast horizons, however, are more balanced than those for the shorter forecast horizons, which is not surprising given that the number of bubble predictors selected by the two estimators is similar for the two long forecast horizons (see Table 1). Moreover, the Lasso estimator tends to be superior (inferior) to the Ridge regression estimator for a combination of a short training-data window and a short forecast horizon, while the Ridge regression estimator gains ground for the long forecast horizons. Finally, the elastic net estimator dominates the Ridge regression estimator mainly at the two short training-data windows, while the Ridge regression estimator appears to perform slightly better for the two long training-data windows.

In sum, the elastic net estimator performs reasonably well relative to the other two estimators, which reflect that this estimator yields a reasonable compromise between the aim (i) to obtain a parsimonious forecasting model, and, (ii) to capture the informational content of the bubble predictors for subsequent the realized volatility. For this reason, we shall focus in the remainder of this section on the results for the elastic net estimator.

We summarize the results for the RMSRFE and MARFE ratios in Table 3. We compare the AR-CORE with the AR-CORE-BUBBLES model

[5] The modified test proposed by Harvey et al. (1997) is given by DM-modified $= ((n+1-2h+n^{-1}h(h-1))/n)^{1/2} \times$ DM, where $n$ denotes the number of forecast errors. For the implementation in R, see the add-on "forecast" package (Hyndman et al., 2023; Hyndman and Khandakar, 2008).

[6] In our implementation of the shrinkage estimators, we allow for a shrinkage of the coefficients of the benchmark predictors (that is, for example, the Lasso can set the coefficients of the benchmark predictors to zero), so that, from a purely conceptual point of view, setting the coefficients of the bubble predictors to zero does not necessarily give the benchmark model. While the AR term is, in our empirical analysis, always included in the model, this is not necessarily true for the shock predictors.

[7] As for the shocks, the results (not reported for reasons of space, but available from the authors upon request) show that mainly supply shocks are included in the penalized model (Lasso estimator and elastic net estimator), while demand shocks enter the penalized model for a combination of a long training-data window and a long forecast horizon. Inventory shocks, in turn, enter the penalized model for a combination of a short training-data window and a long forecast horizon.

**Table 2**
Comparison of estimator performance.

| Panel A: AR-CORE model is the benchmark model | | | | | |
|---|---|---|---|---|---|
| Models | h = 1 | h = 2 | h = 3 | h = 6 | h = 12 |
| Lasso estimator vs. elastic net estimator/w = 0.5 | 1.0082 | 1.0239 | 1.0118 | 0.9854 | 1.0111 |
| Lasso estimator vs. elastic net estimator/w = 0.58 | 1.0152 | 1.0036 | 1.0431 | 0.9854 | 0.9100 |
| Lasso estimator vs. elastic net estimator/w = 0.67 | 1.0171 | 1.0320 | 1.0207 | 1.0063 | 0.9960 |
| Lasso estimator vs. elastic net estimator/w = 0.75 | 1.0073 | 1.0237 | 1.0379 | 1.0143 | 1.0078 |
| Lasso estimator vs. Ridge regression estimator/w = 0.5 | 0.6014 | 1.0041 | 1.0313 | 0.8299 | 1.0268 |
| Lasso estimator vs. Ridge regression estimator/w = 0.58 | 0.5709 | 1.0042 | 0.9495 | 1.2227 | 0.8843 |
| Lasso estimator vs. Ridge regression estimator/w = 0.67 | 1.0296 | 1.0716 | 1.0365 | 1.0076 | 1.0086 |
| Lasso estimator vs. Ridge regression estimator/w = 0.75 | 1.0474 | 1.1052 | 1.0594 | 0.9965 | 0.9633 |
| Ridge regression estimator vs. elastic net estimator/w = 0.5 | 1.6764 | 1.0197 | 0.9811 | 1.1873 | 0.9847 |
| Ridge regression estimator vs. elastic net estimator/w = 0.58 | 1.7783 | 0.9994 | 1.0986 | 0.8060 | 1.0291 |
| Ridge regression estimator vs. elastic net estimator/w = 0.67 | 0.9879 | 0.9631 | 0.9848 | 0.9988 | 0.9875 |
| Ridge regression estimator vs. elastic net estimator/w = 0.75 | 0.9617 | 0.9262 | 0.9797 | 1.0178 | 1.0462 |
| Panel B: AR-SHOCKS-BUBBLE model is the benchmark model | | | | | |
| Estimators | h = 1 | h = 2 | h = 3 | h = 6 | h = 12 |
| Lasso estimator vs. elastic net estimator/w = 0.5 | 1.0242 | 1.0382 | 1.0298 | 1.0370 | 1.5074 |
| Lasso estimator vs. elastic net estimator/w = 0.58 | 1.0208 | 1.0583 | 1.0301 | 1.1239 | 1.0772 |
| Lasso estimator vs. elastic net estimator/w = 0.67 | 1.0088 | 1.0177 | 1.0514 | 1.0161 | 1.0094 |
| Lasso estimator vs. elastic net estimator/w = 0.75 | 1.0087 | 1.0213 | 1.0209 | 0.9970 | 0.9898 |
| Lasso estimator vs. Ridge regression estimator/w = 0.5 | 0.6233 | 0.9807 | 0.9745 | 1.1373 | 1.1310 |
| Lasso estimator vs. Ridge regression estimator/w = 0.58 | 0.6019 | 1.0047 | 1.0422 | 1.2106 | 1.1679 |
| Lasso estimator vs. Ridge regression estimator/w = 0.67 | 1.0349 | 1.0768 | 1.0546 | 1.0156 | 1.0302 |
| Lasso estimator vs. Ridge regression estimator/w = 0.75 | 1.0559 | 1.0990 | 1.1250 | 0.9911 | 0.9528 |
| Ridge regression estimator vs. elastic net estimator/w = 0.5 | 1.6431 | 1.0586 | 1.0567 | 0.9118 | 1.3328 |
| Ridge regression estimator vs. elastic net estimator/w = 0.58 | 1.6962 | 1.0533 | 0.9884 | 0.9283 | 0.9224 |
| Ridge regression estimator vs. elastic net estimator/w = 0.67 | 0.9747 | 0.9451 | 0.9969 | 1.0005 | 0.9798 |
| Ridge regression estimator vs. elastic net estimator/w = 0.75 | 0.9553 | 0.9293 | 0.9075 | 1.0060 | 1.0388 |

The AR-CORE-BUBBLES (AR-SHOCKS-BUBBLES) model, which includes all bubbles as potential predictors, is estimated by the Lasso estimator, elastic net estimator, and the Ridge regression estimator. The estimations are done for four different windows of training data covering 50% up to 75% of the data, as indicated by the parameter $w$. The forecasting models are used to compute forecasts of the remaining test data. The RMSRFE ratio is defined by dividing the RMSRFE of the benchmark (first model reported in the first column) model by the RMSRFE of the rival (second model reported in the first column) model. A RMSRFE ratio larger than unity indicates the rival model produces more accurate forecasts than the benchmark model. The parameter. $h$ denotes the forecast horizon (in months).

**Table 3**
RMSRFE and MARFE ratios.

| Panel A: AR core model is the benchmark model | | | | | |
|---|---|---|---|---|---|
| Models | h = 1 | h = 2 | h = 3 | h = 6 | h = 12 |
| | RMSRFE ratios | | | | |
| AR-CORE vs. AR-CORE-BUBBLES/w = 0.5 | 1.0967 | 1.0605 | 1.0406 | 0.6001 | 0.4351 |
| AR-CORE vs. AR-CORE-BUBBLES/w = 0.58 | 1.0805 | 1.0823 | 1.0990 | 0.4256 | 0.2985 |
| AR-CORE vs. AR-CORE-BUBBLES/w = 0.67 | 1.0634 | 1.0980 | 1.0779 | 0.9371 | 0.9831 |
| AR-CORE vs. AR-CORE-BUBBLES/w = 0.75 | 1.0625 | 1.0903 | 1.0898 | 0.9536 | 1.0037 |
| | MARFE ratios | | | | |
| AR-CORE vs. AR-CORE-BUBBLES/w = 0.5 | 1.0780 | 1.0246 | 0.9893 | 0.7189 | 0.6021 |
| AR-CORE vs. AR-CORE-BUBBLES/w = 0.58 | 1.0622 | 1.0555 | 1.0470 | 0.6116 | 0.5139 |
| AR-CORE vs. AR-CORE-BUBBLES/w = 0.67 | 1.0445 | 1.0693 | 1.0446 | 0.9588 | 0.9821 |
| AR-CORE vs. AR-CORE-BUBBLES/w = 0.75 | 1.0504 | 1.0664 | 1.0438 | 0.9807 | 1.0182 |
| Panel B: AR-SHOCKS model is the benchmark model | | | | | |
| Models | h = 1 | h = 2 | h = 3 | h = 6 | h = 12 |
| | RMSRFE ratios | | | | |
| AR-SHOCKS vs. AR-SHOCK-SBUBBLES/w = 0.5 | 1.1068 | 1.1103 | 1.0738 | 0.5073 | 0.6162 |
| AR-SHOCKS vs. AR-SHOCK-BUBBLES/w = 0.58 | 1.0814 | 1.1251 | 1.1157 | 0.5018 | 0.2832 |
| AR-SHOCKS vs. AR-SHOCK-BUBBLES/w = 0.67 | 1.0620 | 1.0922 | 1.1137 | 0.9472 | 0.9978 |
| AR-SHOCKS vs. AR-SHOCK-BUBBLES/w = 0.75 | 1.0680 | 1.0902 | 1.0719 | 0.9501 | 1.0016 |
| | MARFE ratios | | | | |
| AR-SHOCKS vs. AR-SHOCK-BUBBLES/w = 0.5 | 1.0846 | 1.0656 | 1.0098 | 0.6586 | 0.7189 |
| AR-SHOCKS vs. AR-SHOCK-BUBBLES/w = 0.58 | 1.0629 | 1.0917 | 1.0552 | 0.6655 | 0.4981 |
| AR-SHOCKS vs. AR-SHOCK-BUBBLES/w = 0.67 | 1.0441 | 1.0562 | 1.0623 | 0.9653 | 0.9975 |
| AR-SHOCKS vs. AR-SHOCK-BUBBLES/w = 0.75 | 1.0537 | 1.0564 | 1.0273 | 0.9761 | 1.0151 |

The AR-CORE (AR-SHOCKS) forecasting model is estimated by the OLS technique. The AR-CORE-BUBBLES (AR-SHOCKS-BUBBLES) model, which includes all bubbles as potential predictors, is estimated by the elastic net estimator. The estimations are done for four different windows of training data covering 50% up to 75% of the data, as indicated by the parameter $w$. The forecasting models are used to compute forecasts of the remaining test data. The RMSRFE (MARFE) ratio is defined by dividing the RMSRFE (MARFE) of the benchmark (AR-CORE or AR-SHOCKS) model by the RMSRFE (MARFE) of the rival (AR-CORE-BUBBLES or AR-SHOCKS-BUBBLES) model. A RMSRFE (MARFE) ratio larger than unity indicates the rival model produces more accurate forecasts than the benchmark model. The parameter. $h$ denotes the forecast horizon (in months).

**Table 4**
Modified Diebold–Mariano tests.

| Panel A: AR-CORE model is the benchmark model | | | | | |
|---|---|---|---|---|---|
| Models | h = 1 | h = 2 | h = 3 | h = 6 | h = 12 |
| | Quadratic-forecast-error loss function | | | | |
| AR-CORE vs. AR-CORE-BUBBLES/w = 0.5 | 0.0000 | 0.0845 | 0.2893 | 0.9814 | 0.9431 |
| AR-CORE vs. AR-CORE-BUBBLES/w = 0.58 | 0.0001 | 0.0380 | 0.1154 | 0.9475 | 0.9217 |
| AR-CORE vs. AR-CORE-BUBBLES/w = 0.67 | 0.0005 | 0.0489 | 0.1346 | 0.7721 | 0.7190 |
| AR-CORE vs. AR-CORE-BUBBLES/w = 0.75 | 0.0041 | 0.0408 | 0.1406 | 0.7365 | 0.4642 |
| | Absolute-forecast-error loss function | | | | |
| AR-CORE vs. AR-CORE-BUBBLES/w = 0.5 | 0.0000 | 0.1388 | 0.6055 | 0.9961 | 0.9928 |
| AR-CORE vs. AR-CORE-BUBBLES/w = 0.58 | 0.0000 | 0.0013 | 0.0567 | 0.9892 | 0.9716 |
| AR-CORE vs. AR-CORE-BUBBLES/w = 0.67 | 0.0001 | 0.0002 | 0.0766 | 0.7482 | 0.6610 |
| AR-CORE vs. AR-CORE-BUBBLES/w = 0.75 | 0.0006 | 0.0016 | 0.1302 | 0.6227 | 0.3592 |
| Panel B: AR-SHOCKS model is the benchmark model | | | | | |
| Models | h = 1 | h = 2 | h = 3 | h = 6 | h = 12 |
| | Quadratic-forecast-error loss function | | | | |
| AR-SHOCKS vs. AR-SHOCKS-BUBBLES/w = 0.5 | 0.0000 | 0.0461 | 0.2011 | 0.9729 | 0.9594 |
| AR-SHOCKS vs. AR-SHOCKS-BUBBLES/w = 0.58 | 0.0001 | 0.0432 | 0.1102 | 0.9524 | 0.9215 |
| AR-SHOCKS vs. AR-SHOCKS-BUBBLES/w = 0.67 | 0.0003 | 0.0503 | 0.1157 | 0.7339 | 0.5325 |
| AR-SHOCKS vs. AR-SHOCKS-BUBBLES/w = 0.75 | 0.0033 | 0.0481 | 0.1748 | 0.7524 | 0.4855 |
| | Absolute-forecast-error loss function | | | | |
| AR-SHOCKS vs. AR-SHOCKS-BUBBLES/w = 0.5 | 0.0000 | 0.0018 | 0.3986 | 0.9965 | 0.9932 |
| AR-SHOCKS vs. AR-SHOCKS-BUBBLES/w = 0.58 | 0.0000 | 0.0000 | 0.0353 | 0.9889 | 0.9727 |
| AR-SHOCKS vs. AR-SHOCKS-BUBBLES/w = 0.67 | 0.0001 | 0.0040 | 0.0226 | 0.7112 | 0.5245 |
| AR-SHOCKS vs. AR-SHOCKS-BUBBLES/w = 0.75 | 0.0005 | 0.0100 | 0.2672 | 0.6529 | 0.3831 |

The AR-CORE (AR-SHOCKS) forecasting model is estimated by the OLS technique. The AR-CORE-BUBBLES (AR-SHOCKS-BUBBLES) model, which includes all bubbles as potential predictors, is estimated by the elastic net estimator. The estimations are done for four different windows of training data covering 50% up to 75% of the data, as indicated by the parameter $w$. The forecasting models are used to compute forecasts of the remaining test data. The null hypothesis of the modified Diebold–Mariano test is that the forecasts obtained from the benchmark (AR-CORE or AR-SHOCKS) model and forecasts obtained from the rival (AR-CORE-BUBBLES or AR-SHOCKS-BUBBLES) model are equally accurate. The alternative hypothesis is that the rival forecasts are more accurate than the benchmark forecasts. The test is applied to the relative forecast errors. The table plots the p-value of the test. $h$ denotes the forecast horizon (in months).

and the AR-SHOCKS with the and the AR-SHOCKS-BUBBLES model. The AR-CORE and the AR-SHOCKS benchmark models are estimated by means of the OLS estimator. The general pattern that emerges from the results is that the AR-CORE-BUBBLES and AR-SHOCKS-BUBBLES rival models outperform the benchmark models for the forecast horizons of up to and including three months, where, depending on whether bubbles predictors are included in the penalized forecasting models (see Table 1), the superior performance of the rival forecasting model at the one-month-ahead forecast horizon does not necessarily stem from the bubble predictors, but from the shrinking of the coefficients of the benchmark model.

In order to inspect whether the differences between the accuracy of forecasts are statistically significant, we report in Table 4 the results of the modified Diebold–Mariano test. The null hypothesis is that the benchmark and rival forecasts are equally accurate, while the alternative hypothesis is that the rival forecasts outperform the benchmark forecasts. We present results for a quadratic-forecast-error loss function and an absolute-forecast error loss function (we use relative forecast errors in both cases). We reject in the null hypothesis for the one-month-ahead and two-months-ahead forecast horizons and, mainly under the assumption of an absolute-forecast error loss function, also for the three-months-ahead forecast horizon in some cases. These test results are in line with the results we report in Table 3.

We summarize the RMSRFE ratios that we obtain when we consider four extensions in Table 5. First, we consider an extension in which we estimate both the AR-SHOCKS benchmark model and the AR-SHOCKS-BUBBLES model by means of the elastic net estimator, such that we can directly compare penalized versions of both forecasting models. Second, we compute the RMSRFE ratios based on the anti-logs of the forecasts and the actual realizations of realized volatility. Third, we use oil price returns to replace the shocks, which renders it possible to extend the sample period back to 1973:01. Fourth, we report the results we obtain when we change the mixing parameter from 0.5 to 0.25, bringing the elastic net somewhat closer to the Ridge regression estimator and, thereby, inflating the number of bubble predictors included in the penalized forecasting model. All four extensions are in line with the results of the baseline scenario that we consider in Table 3 in that the rival forecasting models tend to produce more accurate forecasts than the respective benchmark models for forecast horizons up to and including three months.

We document in Table 6 the results we obtain when we differentiate the bubble predictors according to the "type" of bubble: short-term, medium-term, and long-term bubbles as well as positive and negative bubbles. Two results emerge. First, irrespective of the type of bubble being studied, the RMSRFE ratios exceed unity for forecast horizons up to and including three months. Second, the rival model outperforms the benchmark model when we consider short-term bubbles also at the two long forecast horizons. The results for medium-term and long-term bubbles is weaker in this regard, but tends to strengthen in the length of the training-data window. Similarly, we observe that the AR-SHOCKS-BUBBLES rival model outperforms the AR-SHOCKS benchmark model when we study one of the two long training-data windows in the case of the negative bubbles also at the two long forecast horizons.

Finally, we differentiate between "good" and "bad" realized volatility, whereby the former is based on the square root of daily squared positive log-returns only, and the latter uses only negative log-returns. The results that we summarize in Table 7 for this scenario show that the AR-SHOCKS-BUBBLES rival model outperforms the AR-SHOCKS benchmark model in the case of good realized volatility: (i) for the three short and intermediate forecast horizons, and, (ii) for the two long forecast horizons also when we study a long training-data window. Similarly, the results for bad realized volatility strengthen for the two long training-data windows in favor of the AR-SHOCKS-BUBBLES rival model when we consider a forecast horizon up to and including three months.

**Table 5**
Extensions.

| Models | h = 1 | h = 2 | h = 3 | h = 6 | h = 12 |
|---|---|---|---|---|---|
| | Penalized AR-SHOCKS model | | | | |
| AR-SHOCKS vs. AR-SHOCKS-BUBBLES/w = 0.5 | 1.0449 | 1.0451 | 1.0225 | 0.5798 | 0.4649 |
| AR-SHOCKS vs. AR-SHOCKS-BUBBLES/w = 0.58 | 1.0606 | 1.0515 | 1.0380 | 0.4935 | 0.3047 |
| AR-SHOCKS vs. AR-SHOCKS-BUBBLES/w = 0.67 | 1.0404 | 1.0556 | 1.0355 | 0.9258 | 0.9681 |
| AR-SHOCKS vs. AR-SHOCKS-BUBBLES/w = 0.75 | 1.0347 | 1.0327 | 1.0287 | 0.9221 | 0.9849 |
| | Anti-log of realized volatility | | | | |
| AR-SHOCKS vs. AR-SHOCKS-BUBBLES/w = 0.5 | 1.1241 | 1.1517 | 1.1287 | 0.5405 | 0.6367 |
| AR-SHOCKS vs. AR-SHOCKS-BUBBLES/w = 0.58 | 1.0965 | 1.1678 | 1.1675 | 0.5440 | 0.3382 |
| AR-SHOCKS vs. AR-SHOCKS-BUBBLES/w = 0.67 | 1.0740 | 1.1174 | 1.1632 | 0.9473 | 0.9986 |
| AR-SHOCKS vs. AR-SHOCKS-BUBBLES/w = 0.75 | 1.0839 | 1.1226 | 1.1252 | 0.9496 | 1.0019 |
| | AR-RETURNS model | | | | |
| AR-RETURNS vs. AR-RETURNS-BUBBLES/w = 0.5 | 1.1129 | 1.1068 | 0.9986 | 0.7341 | 0.8271 |
| AR-RETURNS vs. AR-RETURNS-BUBBLES/w = 0.58 | 1.0861 | 1.1315 | 1.1144 | 0.8854 | 0.8619 |
| AR-RETURNS vs. AR-RETURNS-BUBBLES/w = 0.67 | 1.0857 | 1.1110 | 1.1074 | 0.9203 | 0.9722 |
| AR-RETURNS vs. AR-RETURNS-BUBBLES/w = 0.75 | 1.0658 | 1.0900 | 1.0698 | 0.9387 | 1.0233 |
| | Mixing parameter 0.25 | | | | |
| AR-SHOCKS vs. AR-SHOCKS-BUBBLES/w = 0.5 | 1.0879 | 1.1032 | 1.0719 | 0.5048 | 0.5356 |
| AR-SHOCKS vs. AR-SHOCKS-BUBBLES/w = 0.58 | 1.0458 | 1.1437 | 1.1393 | 0.5685 | 0.3256 |
| AR-SHOCKS vs. AR-SHOCKS-BUBBLES/w = 0.67 | 1.0930 | 1.1328 | 1.1247 | 0.9448 | 1.0040 |
| AR-SHOCKS vs. AR-SHOCKS-BUBBLES/w = 0.75 | 1.0834 | 1.1275 | 1.1187 | 0.9465 | 1.0275 |

Panel A: The benchmark model is the penalized AR-SHOCKS model estimated by the elastic net estimator. The AR-SHOCKS-BUBBLES rival model is estimated by the elastic net estimator. Panel B: The RMSRFE is computed for the anti-log of realized volatility. The AR-SHOCKS benchmark model is estimated by the OLS estimator and the AR-SHOCKS-BUBBLES rival model is estimated by the elastic net estimator. Panel C: The AR-RETURNS model, estimated by the OLS estimator, is the benchmark model and the AR-RETURNS-BUBBLES model, estimated by the elastic net estimator, is the rival model. The sample period starts in 1973/01. Panel D: The AR-SHOCKS-BUBBLES rival model is estimated by the elastic net estimator with a mixing parameter equal to 0.25. All Panels: The estimations are done for four different windows of training data covering 50% up to 75% of the data, as indicated by the parameter $w$. The forecasting models are used to compute forecasts of the remaining test data. The RMSRFE ratio is defined by dividing the RMSRFE of the benchmark model by the RMSRFE of the rival model. A RMSRFE ratio larger than unity indicates the rival model produces more accurate forecasts than the benchmark model. The parameter. $h$ denotes the forecast horizon (in months).

**Table 6**
Results by type of bubbles.

| Models | h = 1 | h = 2 | h = 3 | h = 6 | h = 12 |
|---|---|---|---|---|---|
| | Short-term bubbles | | | | |
| AR-SHOCKS vs. AR-SHOCKS-BUBBLES/w = 0.5 | 1.0738 | 1.0850 | 1.0663 | 1.0200 | 1.0060 |
| AR-SHOCKS vs. AR-SHOCKS-BUBBLES/w = 0.58 | 1.0769 | 1.0881 | 1.0646 | 1.0192 | 1.0208 |
| AR-SHOCKS vs. AR-SHOCKS-BUBBLES/w = 0.67 | 1.0622 | 1.0977 | 1.0782 | 1.0187 | 1.0356 |
| AR-SHOCKS vs. AR-SHOCKS-BUBBLES/w = 0.75 | 1.0480 | 1.0800 | 1.0927 | 1.0409 | 1.0685 |
| | Medium-term bubbles | | | | |
| AR-SHOCKS vs. AR-SHOCKS-BUBBLES/w = 0.5 | 1.0718 | 1.0627 | 1.0414 | 0.9597 | 0.9581 |
| AR-SHOCKS vs. AR-SHOCKS-BUBBLES/w = 0.58 | 1.0755 | 1.0673 | 1.0511 | 0.9633 | 1.0002 |
| AR-SHOCKS vs. AR-SHOCKS-BUBBLES/w = 0.67 | 1.0576 | 1.0765 | 1.0957 | 0.9773 | 1.0161 |
| AR-SHOCKS vs. AR-SHOCKS-BUBBLES/w = 0.75 | 1.0682 | 1.1028 | 1.1093 | 0.9742 | 1.0232 |
| | Long-term bubbles | | | | |
| AR-SHOCKS vs. AR-SHOCKS-BUBBLES/w = 0.5 | 1.0587 | 1.0585 | 1.0485 | 0.7839 | 0.4656 |
| AR-SHOCKS vs. AR-SHOCKS-BUBBLES/w = 0.58 | 1.0270 | 1.0773 | 1.0881 | 0.8402 | 0.2847 |
| AR-SHOCKS vs. AR-SHOCKS-BUBBLES/w = 0.67 | 1.0573 | 1.0684 | 1.0735 | 0.9681 | 1.0270 |
| AR-SHOCKS vs. AR-SHOCKS-BUBBLES/w = 0.75 | 1.0410 | 1.0681 | 1.0620 | 0.9731 | 1.0425 |
| | Positive bubbles | | | | |
| AR-SHOCKS vs. AR-SHOCKS-BUBBLES/w = 0.5 | 1.0882 | 1.0753 | 1.0610 | 0.7833 | 0.8812 |
| AR-SHOCKS vs. AR-SHOCKS-BUBBLES/w = 0.58 | 1.0700 | 1.0883 | 1.1024 | 0.9075 | 0.9555 |
| AR-SHOCKS vs. AR-SHOCKS-BUBBLES/w = 0.67 | 1.0650 | 1.0789 | 1.0860 | 0.9314 | 0.9786 |
| AR-SHOCKS vs. AR-SHOCKS-BUBBLES/w = 0.75 | 1.0493 | 1.0727 | 1.0530 | 0.9620 | 1.0134 |
| | Negative bubbles | | | | |
| AR-SHOCKS vs. AR-SHOCKS-BUBBLES/w = 0.5 | 1.0588 | 1.0810 | 1.0871 | 0.7910 | 0.4721 |
| AR-SHOCKS vs. AR-SHOCKS-BUBBLES/w = 0.58 | 1.0661 | 1.0877 | 1.0845 | 0.7815 | 0.4100 |
| AR-SHOCKS vs. AR-SHOCKS-BUBBLES/w = 0.67 | 1.0569 | 1.0825 | 1.1020 | 1.0495 | 1.0434 |
| AR-SHOCKS vs. AR-SHOCKS-BUBBLES/w = 0.75 | 1.0650 | 1.0837 | 1.0874 | 1.0068 | 1.0570 |

The AR-SHOCKS forecasting model is estimated by the OLS technique. The AR-SHOCKS-BUBBLES model, which includes the short-term, medium-term, or long-term bubbles as potential predictors, is estimated by the elastic net estimator. The estimations are done for four different windows of training data covering 50% up to 75% of the data, as indicated by the parameter $w$. The forecasting models are used to compute forecasts of the remaining test data. The RMSRFE ratio is defined by dividing the relative RMSRFE of the benchmark (AR-SHOCKS) model by the relative RMSRFE of the rival (AR-SHOCKS-BUBBLES) model. A RMSRFE ratio larger than unity indicates the rival model produces more accurate forecasts than the benchmark model. The parameter. $h$ denotes the forecast horizon (in months).

**Table 7**
Results for good and bad realized volatility.

| Models | h = 1 | h = 2 | h = 3 | h = 6 | h = 12 |
|---|---|---|---|---|---|
| | Good realized volatility | | | | |
| AR-SHOCKS vs. AR-SHOCKS-BUBBLES/w = 0.5 | 1.1589 | 1.1362 | 1.1247 | 0.9767 | 0.7715 |
| AR-SHOCKS vs. AR-SHOCKS-BUBBLES/w = 0.58 | 1.1153 | 1.1353 | 1.1898 | 0.8619 | 0.4849 |
| AR-SHOCKS vs. AR-SHOCKS-BUBBLES/w = 0.67 | 1.1100 | 1.1330 | 1.1720 | 1.0410 | 1.0889 |
| AR-SHOCKS vs. AR-SHOCKS-BUBBLES/w = 0.75 | 1.0599 | 1.1197 | 1.1119 | 1.0238 | 1.1090 |
| | Bad realized volatility | | | | |
| AR-SHOCKS vs. AR-SHOCKS-BUBBLES/w = 0.5 | 0.4086 | 0.5399 | 0.5707 | 0.7787 | 0.8303 |
| AR-SHOCKS vs. AR-SHOCKS-BUBBLES/w = 0.58 | 0.3826 | 0.5195 | 0.5717 | 0.9264 | 0.9304 |
| AR-SHOCKS vs. AR-SHOCKS-BUBBLES/w = 0.67 | 1.1624 | 1.2123 | 1.1562 | 0.9845 | 1.0207 |
| AR-SHOCKS vs. AR-SHOCKS-BUBBLES/w = 0.75 | 1.1657 | 1.1929 | 1.0966 | 0.9439 | 0.9728 |

The AR-SHOCKS forecasting model is estimated by the OLS technique. The AR-SHOCKS-BUBBLES model, which includes all bubbles as potential predictors, is estimated by the elastic net estimator. The estimations are done for four different windows of training data covering 50% up to 75% of the data, as indicated by the parameter $w$. The forecasting models are used to compute forecasts of the remaining test data. The RMSRFE ratio is defined by dividing the RMSFE of the benchmark (AR-SHOCKS) model by the RMSRFE of the rival (AR-SHOCKS-BUBBLES) model. A relative RMSRFE ratio larger than unity indicates the rival model produces more accurate forecasts than the benchmark model. The parameter. $h$ denotes the forecast horizon (in months).

## 5. Concluding remarks

The objective of this paper is to forecast monthly realized volatility of returns of the WTI oil price based on the information content of multi-scale positive and negative stock market bubbles indicators of the G7 countries over 1973:01 to 2020:09. After successfully detecting major crashes and rallies, our main finding is that stock market bubbles have predictive value for realized volatility at a short to intermediate forecast horizon. This is particularly the case when the shrinkage estimators that we have studied tend to include more of the bubble predictors in the penalized forecasting model as the forecast horizon increases, where in some cases no stock market bubbles makes it into the penalized forecasting model at the one-month-ahead forecast horizon, especially when we have studied the Lasso estimator. In this regard, we have found that the elastic net estimator strikes a reasonable balance between the aim of model parsimony and the aim to exploit the predictive value of the stock market bubbles for the realized volatility of oil price returns. Our main finding also holds for various types of stock market bubbles, good and bad realized volatilities, and for several model extensions. Moreover, statistical tests show that the difference in accuracy between the benchmark and rival forecasts, while moderate in some cases, at short to intermediate forecast horizons are often statistically significant. Given that the pattern of results is similar across the various model configurations that we have studied we conclude that our empirical results lend support to the hypothesis that speculative stock market bubbles help to improve the accuracy of forecasts of the realized volatility of oil price returns (after controlling for the impact of the autoregressive component of realized volatility, various shocks, and oil price returns) mainly at short- and intermediate-forecast horizons.

Our main findings are important from the perspective of academics because we are able to validate econometrically the intuition that stock market bubbles can have repercussions in the form of oil-price fluctuations, beyond market-specific structural shocks. In addition to being important for academics, our results also yield important insights for investors. The predictive role of G7 stock market bubbles for oil-price volatility may help oil traders to optimize portfolio-allocation decisions. Moreover, volatility is a crucial parameter in the field of option pricing, where one needs reliable forecasts of the volatility of the underlying asset(s). In this regard, our main findings that stock market bubbles have predictive value for the subsequent realized volatility of oil-price movements at short- and intermediate forecast horizons can be utilized by investors for pricing oil derivatives. Finally, in light of the oil volatility-recession nexus, our findings imply that policymakers should be cognizant of not only the direct effect of boom-bust cycles in the G7 stock markets on real activity, but rather policymakers should also be aware of a potential indirect effect that originates via oil returns volatility, potentially resulting in a more persistent impact on real activity. This would be a pertinent issue when designing the strength and time-profile of macroeconomic (monetary and fiscal) policies to combat possible second-round adverse real effects of large stock-market swings in general and stock-market crashes in particular that are likely to result in higher oil price volatility.

While we are able to obtain favorable predictive outcomes for oil market volatility due to stock market bubbles, much research remains to be done. For example, in a recent paper, Balcilar et al. (2022) show that oil uncertainty, as captured by oil-returns volatility, can predict the connectedness of stock markets of advanced countries. Given this, and our observation of the synchronicity in the movement of the MS-LPPLS-CIs, an interesting avenue for future research is to analyze whether oil uncertainty can explain the connectedness of positive and negative bubbles at the short-, medium-, and long-run, besides, possibly even impacting the individual-level bubbles indicators. This would then tie the work with the possible role of uncertainty and volatility as an explanatory variable for the origin and bursts of bubbles, besides other factors (as discussed in detail by Sornette et al. (2018)).

Finally, while we present our results based on linear machine learning approaches, another area for further research would be to revisit our results using nonlinear methods in a big data setting. For example, one could utilize random forests (Breiman, 2001), which, in turn, would help to shed light on potential a nonlinear dependence between oil-market realized volatility and stock market bubbles, as well as any interaction effects between the predictors. Understandably, random forests is a relatively robust econometric method, but whether it would end up producing superior results relative to what we have thus far obtained utilizing linear models is indeed an empirical question for the future.

**Funding statement**

The authors declare that they did not receive any funding for this research.

**CRediT authorship contribution statement**

**Rangan Gupta:** Conceptualization, Data curation, Supervision, Validation, Writing – original draft, Writing – review & editing. **Joshua Nielsen:** Formal analysis, Methodology, Software, Writing – review & editing. **Christian Pierdzioch:** Formal analysis, Investigation, Methodology, Software, Writing – review & editing.

**Declaration of competing interest**

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Appendix A

### A.1. The MS-LPPLS model: Structure and estimation

We describe here in some detail the Multi-Scale Log-Periodic Power Law Singularity (MS-LPPLS) model. For our description of the details of the MS-LPPLS model, we heavily draw on the formal descriptions of the model provided in earlier works by Demirer et al. (2019), Caraiani et al. (2023), Gupta et al. (2023), and van Eyden et al. (2023).

We use the stable and robust calibration scheme developed by Filimonov and Sornette (2013) to set up a LPPLS model of the following format:

$$\ln E[p(t)] = A + B(t_c - t)^m + C(t_c - t)^m \cos(\omega \ln (t_c - t)^m - \phi), \tag{A.1}$$

where $t_c$ = denotes the critical time (the period of time when a bubble burst), $A$ = the expected log value of the time-series being studied (the stock price–dividend ratio) as observed at time $t_c$, $B$ = the amplitude of the power law acceleration, $C$ = the relative magnitude of the log-periodic oscillations, $m$ =the exponent of the power law growth, $\omega$ = the frequency of the log-periodic oscillations, and $\phi$ = a phase shift parameter.

Like Filimonov and Sornette (2013), we next re-express Eq. (A.1) so as to reduce the complexity of the calibration process. To this end, we eliminate the nonlinear parameter, $\phi$, and use the expansions $C$ to $C_1 = C\cos\phi$ and $C_2 = C\cos\phi$ for the linear parameters, to obtain the following equation:

$$\ln E[p(t)] = A + B(f) + C_1(g) + C_2(h), \tag{A.2}$$

where

$$f = (t_c - t)^m,$$
$$g = (t_c - t)^m cos[\omega \ln (t_c - t)],$$
$$h = (t_c - t)^m sin[\omega \ln (t_c - t)].$$

We estimate Eq. (A.2) on the log of the price–dividend ratio so as to obtain estimates of the three nonlinear parameters, $\{t_c, m, \omega\}$, and the four linear parameters, $\{A, B, C_1, C_2\}$. Using the $L^2$ norm, we obtain

$$F(t_c, m, \omega, A, B, C_1, C_2) = \sum_{i=1}^{N} \left[\ln p(\tau_i) - A - B(f_i) - C_1(g_i) - C_2(h_i)\right]^2. \tag{A.3}$$

Because the estimates of the nonlinear parameters depend on the linear ones, we formulate the following cost function,

$$F_1(t_c, m, \omega) = \min_{A,B,C_1,C_2} F(t_c, m, \omega, A, B, C_1, C_2) = F(t_c, m, \omega, \hat{A}, \hat{B}, \hat{C}_1, \hat{C}_2), \tag{A.4}$$

and solve the following optimization problem:

$$\{\hat{A}, \hat{B}, \hat{C}_1, \hat{C}_2\} = \arg \min_{A,B,C_1,C_2} F(t_c, m, \omega, A, B, C_1, C_2), \tag{A.5}$$

which can be done analytically by solving the following matrix equation:

$$\begin{pmatrix} N & \sum f_i & \sum g_i & \sum h_i \\ \sum f_i & \sum f_i^2 & \sum f_i g_i & \sum f_i h_i \\ \sum g_i & \sum f_i g_i & \sum g_i^2 & \sum g_i h_i \\ \sum h_i & \sum f_i h_i & \sum g_i h_i & \sum h_i^2 \end{pmatrix} \begin{pmatrix} \hat{A} \\ \hat{B} \\ \hat{C}_1 \\ \hat{C}_2 \end{pmatrix} = \begin{pmatrix} \sum \ln p_i \\ \sum f_i \ln p_i \\ \sum g_i \ln p_i \\ \sum h_i \ln p_i \end{pmatrix}. \tag{A.6}$$

We then determine the three nonlinear parameters by solving the following nonlinear optimization problem:

$$\{\hat{t_c}, \hat{m}, \hat{\omega}\} = \arg \min_{t_c,m,\omega} F_1(t_c, m, \omega), \tag{A.7}$$

using the Sequential Least Squares Programming (SLSQP) search algorithm (Kraft, 1988).

We use the LPPLS confidence indicator (Sornette et al., 2015) to measure the prevalence of bubble patterns in the log price–dividend ratio. A large LPPLS confidence indicator (CI) signals the presence of a LPPLS bubble pattern. To this end, we calibrate the LPPLS model to shrinking time windows by shifting the initial observation, $t_1$, forward in time towards the final observation, $t_2$, with a step $dt$. For each LPPLS model fit, we compare the estimated parameters against established thresholds and take the qualified fits as a fraction of the total number of positive or negative fits, where $B < 0$ ($B > 0$) for a positive (negative) fit.

Building on the research by Demirer et al. (2019), we study bubbles of varying multiple time-scales. We sample the time series in steps of 5 trading days, define the nested windows $[t_1, t_2]$, and iterate through each window in steps of 2 trading days, resulting in a weekly resolution that we use to obtain the following indicators:

- Short-term bubble: A number $\in [0, 1]$ which denotes the fraction of qualified fits for estimation windows of length $dt := t_2 - t_1 \in [30 : 90]$ trading days per $t_2$. This indicator is comprised of $(90 - 30)/2 = 30$ fits.
- Medium-term bubble: A number $\in [0, 1]$ which denotes the fraction of qualified fits for estimation windows of length $dt := t_2 - t_1 \in [30 : 90]$ trading days per $t_2$. This indicator is comprised of $(300 - 90)/2 = 105$ fits.
- Long-term bubble: A number $\in [0, 1]$ which denotes the fraction of qualified fits for estimation windows of length $dt := t_2 - t_1 \in [30 : 90]$ trading days per $t_2$. This indicator is comprised of $(745 - 300)/2 = 223$ fits.

In a final step, we use the following filter conditions to check which fits are qualified:

$$m \in [0.01, 0.99],$$
$$\omega \in [2, 15],$$
$$t_c \in [\max(t_2 - 60, t_2 - 0.5(t_2 - t_1)), \min(252, t_2 + 0.5(t_2 - t_1))],$$
$$O > 2.5,$$
$$D > 0.5,$$

where

$$O = \frac{\omega}{2\pi} \ln\left(\frac{t_c - t_1}{t_c - t_2}\right),$$
$$D = \frac{m|B|}{\omega|C|}.$$

## Appendix B. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.eneco.2024.107432.

## References

Aboura, S., Chevallier, J., 2013. Leverage vs. Feedback: Which effect drives the oil market? Finance Res. Lett. 10 (3), 131–141.

Andersen, T.G., Bollerslev, T., 1998. Answering the skeptics: yes, standard volatility models do provide accurate forecasts. Internat. Econom. Rev. 39 (4), 885–905.

Antonakakis, N., Gupta, R., Kollias, C., Papadamou, S., 2017. Geopolitical risks and the oil-stock nexus over 1899–2016. Finance Res. Lett. 23, 165–173.

Asai, M., Gupta, R., McAleer, M., 2019. The impact of jumps and leverage in forecasting the co-volatility of oil and gold futures. Energies 12 (17), 3379.

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

Balcilar, M., Gupta, R., Jooste, C., Wohar, M.E., 2016. Periodically collapsing bubbles in the South African stock market? Res. Int. Bus. Finance 38, 191–201.

Balcilar, M., Gupta, R., Pierdzioch, C., 2022. Oil-price uncertainty and international stock returns: Dissecting quantile-based predictability and spillover effects using more than a century of data. Energies 15 (22), 8436.

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

Baumeister, C., Hamilton, J.D., 2019. Structural interpretation of vector autoregressions with incomplete identification: Revisiting the role of oil supply and demand shocks. Amer. Econ. Rev. 109 (5), 1873–1910.

Bernanke, B.S., 1983. Irreversibility, uncertainty, and cyclical investment. Q. J. Econ. 98 (1), 85–106.

Biswas, S., Hanson, A., Phan, T., 2020. Bubbly recessions. Am. Econ. J.: Macroecon. 12 (4), 33–70.

Black, F., 1976. Studies of stock price volatility changes. In: Proceedings of the 1976 Meeting of the Business and Economic Statistics Section. American Statistical Association, Washington, DC, USA, pp. 177–181.

Bonato, M., Gkillas, K., Gupta, R., Pierdzioch, C., 2020. Investor happiness and predictability of the realized volatility of oil price. Sustainability 12 (10), 4309.

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

Bouri, E., Gupta, R., Pierdzioch, C., Salisu, A.A., 2021. El Niño and forecastability of oil-price realized volatility. Theor. Appl. Climatol. 144 (3–4), 1173–1180.

Breiman, L., 2001. Random forests. Mach. Learn. 45 (1), 5–32.

Brunnermeier, M.K., Oehmke, M., 2013. Bubbles, financial crises, and systemic risk. In: Constantinides, George M., Harris, Milton, Stulz, Rene M. (Eds.), Handbook of the Economics of Finance. Elsevier, The Netherlands, pp. 1221–1288, 2(B) Chapter 18.

Büyükşahin, B., Robe, M.A., 2014. Speculators, commodities and cross-market linkages. J. Int. Money Finance 42, 38–70.

Caraiani, P., Gupta, R., Nel, J., Nielsen, J., 2023. Monetary policy and bubbles in G7 economies using a panel VAR approach: Implications for sustainable development. Econ. Anal. Policy 78, 133–155.

Çepni, O., Gupta, R., Pienaar, D., Pierdzioch, C., 2022. Forecasting the realized variance of oil-price returns using machine learning: Is there a role for U.S. state-level uncertainty? Energy Econ. 114, 106229.

Chan, J.C.C., Grant, A., 2016. Modeling energy price dynamics: GARCH versus stochastic volatility. Energy Econ. 54, 182–189.

Das, S., Demirer, R., Gupta, R., Mangisa, S., 2019. The effect of global crises on stock market correlations: Evidence from scalar regressions via functional data analysis. Struct. Change Econ. Dyn. 50, 132–147.

Degiannakis, S., Filis, G., 2017. Forecasting oil price realized volatility using information channels from other asset classes. J. Int. Money Finance 76, 28–49.

Demirer, R., Demos, G., Gupta, R., Sornette, D., 2019. On the predictability of stock market bubbles: Evidence from LPPLS confidence multi-scale indicators. Quant. Finance 19 (5), 843–858.

Demirer, R., Gkillas, K., Gupta, R., Pierdzioch, C., 2022. Risk aversion and the predictability of crude oil market volatility: A forecasting experiment with random forests. J. Oper. Res. Soc. 73 (8), 1755–1767.

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

Diebold, F.X., Mariano, R.S., 1995. Comparing predictive accuracy. J. Bus. Econom. Statist. 13 (3), 253–263.

Filimonov, V., Sornette, D., 2013. A stable and robust calibration scheme of the log-periodic power law model. Physica A 392 (17), 3698–3707.

Friedman, J., Tibshirani, R., Hastie, T., 2010. Regularization paths for generalized linear models via coordinate descent. J. Stat. Softw. 33 (1), 1–22.

Geman, H., Shih, Y.F., 2009. Modeling commodity prices under the CEV model. J. Altern. Invest. 11 (3), 65–84.

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

Goswami, S., Gupta, R., Wohar, M.E., 2020. Historical volatility of advanced equity markets: The role of local and global crises. Finance Res. Lett. 34, 101265.

Gupta, R., Nel, J., Nielsen, J., 2023. US monetary policy and BRICS stock market bubbles. Finance Res. Lett. 51, 103435.

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

Gupta, R., 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., Pierdzioch, C., 2022. Forecasting the realized variance of oil-price returns: a disaggregated analysis of the role of uncertainty and geopolitical risk. Environ. Sci. Pollut. Res. 29 (29), 52070–52082.

Gupta, R., Pierdzioch, C., 2023. Do U.S. economic conditions at the state level predict the realized volatility of oil-price returns? A quantile machine-learning approach. Financial Innov. 9 (1), 24.

Harvey, D., Leybourne, S., Newbold, P., 1997. Testing the equality of prediction mean squared errors. Int. J. Forecast. 13, 281–291.

Hastie, T., Tibshirani, R., Friedman, J., 2009. The Elements of Statistical Learning: Data Mining, Inference, and Prediction, second ed. Springer, New York, NY, USA.

Hyndman, R., Athanasopoulos, G., Bergmeir, C., Caceres, G., Chhay, L., O'Hara-Wild, M., Razbash, S.Petropoulos.F., Wang, E., Yasmeen, F., 2023. Forecast: Forecasting functions for time series and linear models. r package version 8.21. URL: https://pkg.robjhyndman.com/forecast.

Hyndman, R.J., Khandakar, Y., 2008. Automatic time series forecasting: The forecast package for r. J. Stat. Softw. 26 (3), 1–22. http://dx.doi.org/10.18637/jss.v027.i03.

Johansen, A., Ledoit, O., Sornette, D., 2000. Crashes as critical points. Int. J. Theor. Appl. Finance 2, 219–255.

Johansen, A., Sornette, D., Ledoit, O., 1999. Predicting financial crashes using discrete scale invariance. J. Risk 1 (4), 5–32.

Jordà, Ò., Schularick, M., Taylor, A.M., 2015. Leveraged bubbles. J. Monetary Econ. 76, S1–S20.

Kilian, L., 2009. Not all oil price shocks are alike: disentangling demand and supply shocks in the crude oil market. Amer. Econ. Rev. 99 (3), 1053–1069.

Kraft, D., 1988. A Software Package for Sequential Quadratic Programming. Technical Report DFVLR-FB 88-28, Institut für Dynamik der Flugsysteme, Oberpfaffenhofen.

Ludvigson, S.C., Ma, S., Ng, S., 2021. Uncertainty and business cycles: Exogenous impulse or endogenous response? Am. Econ. J.: Macroecon. 13 (4), 369–410.

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

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

Müller, U.A., Dacorogna, M.M., Davé, R.D., Olsen, R.B., Pictet, O.V., 1997. Volatilities of different time resolutions – analyzing the dynamics of market components. J. Empir. Financ. 4 (2–3), 213–239.

Poon, S.-H., Granger, C.W.J., 2003. Forecasting volatility in financial markets: A review. J. Econ. Lit. 41 (2), 478–539.

R Core Team, 2023. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria, URL: https://www.R-project.org/.

Reinhart, C.M., Rogoff, K.S., 2009. This Time Is Different: Eight Centuries of Financial Folly. Princeton University Press, Princeton, USA.

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

Salisu, A.A., Gupta, R., Olaniran, A., 2021. The effect of oil uncertainty shock on real GDP of 33 countries: A global VAR approach. Appl. Econ. Lett. 30 (3), 269–274.

Skrobotov, A., 2023. Testing for explosive bubbles: a review. Depend. Model. 11 (1), 20220152.

Sornette, D., 2003. Why Stock Markets Crash: Critical Events in Complex Financial Systems. Princeton University Press, New Jersey.

Sornette, D., Cauwels, P., Smilyanov, G., 2018. Can we use volatility to diagnose financial bubbles? Lessons from 40 historical bubbles. Quant. Finance Econ. 2 (1), 486–590.

Sornette, D., Demos, G., Zhang, Q., Cauwels, P., Filimonov, V., Zhang, Q., 2015. Real-time prediction and post-mortem analysis of the Shanghai 2015 stock market bubble and crash. J. Invest. Strateg. 4, 77–95.

Tang, K., Xiong, W., 2012. Index investment and the financialization of commodities. Financ. Anal. J. 68, 54–74.

Tibshirani, R., 1996. Regression shrinkage and selection via the lasso. J. R. Stat. Soc. Ser. B Stat. Methodol. 58 (1), 267–288.

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

van Eyden, R., Gupta, R., Nielsen, J., Bouri, E., 2023. Investor sentiment and multi-scale positive and negative stock market bubbles in a panel of G7 countries. J. Behav. Exp. Finance 38, 100804.

Zhang, Q., Sornette, D., Balcilar, M., Gupta, R., Ozdemir, Z.A., Yetkiner, H., 2016. LPPLS bubble indicators over two centuries of the S & P 500 index. Physica A 458 (C), 126–139.