Computational Economics (2024) 64:487–513
https://doi.org/10.1007/s10614-023-10452-w

# Safe Havens, Machine Learning, and the Sources of Geopolitical Risk: A Forecasting Analysis Using Over a Century of Data

**Rangan Gupta**[1] · **Sayar Karmakar**[2] · **Christian Pierdzioch**[3]

Accepted: 2 August 2023 / Published online: 17 August 2023


**Abstract**
We use monthly data covering a century-long sample period (1915–2021) to study whether geopolitical risk helps to forecast subsequent gold volatility. We account not only for geopolitical threats and acts, but also for 39 country-specific sources of geopolitical risk. The response of subsequent volatility is heterogeneous across countries and nonlinear. We find that accounting for geopolitical risk at the country level improves forecast accuracy, especially when we use random forests to estimate our forecasting models. As an extension, we report empirical evidence on the predictive value of the country-level sources of geopolitical risk for two other candidate safe-haven assets, oil and silver, over the sample periods 1900–2021 and 1915–2021, respectively. Our results have important implications for the portfolio and risk-management decisions of investors who seek a safe haven in times of heightened geopolitical tensions.

**Keywords** Gold · Geopolitical risk · Forecasting · Returns · Volatility · Random forests

**JEL Classification** C22 · D80 · H56 · Q02

We would like to thank two anonymous referees for many helpful comments. However, any remaining errors are solely ours.

✉ Christian Pierdzioch
macroeconomics@hsu-hh.de

Rangan Gupta
rangan.gupta@up.ac.za

Sayar Karmakar
sayarkarmakar@ufl.edu

1 Department of Economics, University of Pretoria, Private Bag X20, Hatfield 0028, South Africa

2 Department of Statistics, University of Florida, 230 Newell Drive, Gainesville, FL 32601, USA

3 Department of Economics, Helmut Schmidt University, Holstenhofweg 85, P.O.B. 700822, 22008 Hamburg, Germany

## 1 Introduction

The role of gold as a "safe haven" in times of extreme jitters and disruptions in financial (stocks, bonds, and (crypto-)currencies) and commodity markets has been extensively studied in a large and significant literature and, thus, is a well-established research topic (see, for example, Baur and Lucey, 2010; Baur and McDermott, 2010; Reboredo, 2013a; Agyei-Ampomah et al., 2014; Gürgün and Ünalmis, 2014; Beckmann et al., 2015; Balcilar et al., 2020; Reboredo, 2013b; Low et al., 2016; Tiwari et al., 2020). More recently, several systematic studies (such as Balcilar et al., 2016 and 2017; Bouoiyour et al., 2018; Beckmann et al., 2019; Boubaker et al., 2020; Huynh; 2020; Bouri et al., 2021) have been undertaken to better understand the role played by crises, general economic uncertainty, uncertainty due to the COVID-19 pandemic, investor sentiment, i.e., global shocks that adversely affect the markets for risky assets and commodities, and, in the process, act as drivers of gold prices in the context of its safe haven property. Not surprisingly, the market for gold is by now the world's largest metal market in terms of U.S. dollar, valued at 170 billion U.S. dollars per year at current spot prices, with a production of over 3200 tons per annum (World Gold Council).

Building on the above-mentioned second line of research involving negative worldwide shocks, Baur and Smales (2020) find that gold serves as not only a hedge against geopolitical risk, but also its safe-haven property continues to hold under extreme geopolitical risk. A strong causal impact of geopolitical risk on gold returns has also been reported recently by Gozgor et al. (2019), Li et al. (2021), and Huang et al. (2023). Taken together, these findings hold important lessons given that central bankers, the financial press, and business investors have often cited geopolitical risk as a determinant of investment decisions (Caldara & Iacoviello, 2022). In this regard, when Gallup surveyed in 2017 more than 1000 investors, 75 percent of respondents expressed concerns about the economic impact of the various military and diplomatic conflicts taking place around the world, and in this regard geopolitical risk ranked ahead of political and economic uncertainties. It is, therefore, not surprising that Carney (2016) includes geopolitical risk, along with economic and policy uncertainties, in the "uncertainty trinity", which could have significant adverse economic and financial effects. Moreover, in the April 2017 Economic Bulletin of the European Central Bank, and the October 2017 World Economic Outlook of the International Monetary Fund, geopolitical uncertainty is highlighted as a salient risk to the economic outlook. Besides geopolitical risk being the dominant form of uncertainty, researchers have shown that it contains leading information for not only real economic variables of advanced and emerging economies (Cheng & Chiu, 2018; Clance et al., 2019; Caldara & Iacoviello, 2022), but also for financial markets of these economies (Balcilar et al., 2018; Bouri et al., 2019, forthcoming; Gupta et al., 2021, Salisu et al., 2022a; 2022b; Yang et al. 2021), and, in addition, for (co-)movements of commodity markets (Ding & Zhang, 2021; Tiwari et al., 2021).

Against this backdrop of in-sample evidence involving geopolitical risk and gold returns, our objective is to conduct an elaborate out-of-sample forecasting analysis of the role of such risk for the predictability of gold volatility, since in-sample evidence does not necessarily translate into out-of-sample forecasting gains (Rapach & Zhou, 2013). Besides the statistical significance of a forecasting exercise, real-time forecasting of gold volatility is of much more value to investors, relative to full-sample predictions, in designing their optimal portfolios involving gold due its its ability to offer diversification and hedging benefits during periods of turmoil and heightened uncertainties in financial and commodity markets, and the economy in general, emanating from geopolitical risk.

As far as the forecasting experiment is concerned, we analyze the role of global geopolitical risk, and since throughout history the realization of adverse geopolitical events has often been the catalyst for increased fears about future adverse events, we also disaggregate the overall geopolitical risk into threats and realization of adverse geopolitical events, i.e., acts. Importantly, we also study the contribution of country-specific geopolitical risk involving 39 economies to the forecastability of gold volatility. Given that many of the 39 economies considered here play an important role in the supply- and demand-side of the gold market, besides being vulnerable to consistent geopolitical risk at different points in time, the emphasis of our analysis to use country-specific data is likely to be more informative than overall geopolitical risk at the global level. Because analyzing so many predictors in a standard predictive regression framework comes at a cost of overparameterization, and, hence, poor out-of-sample performance, when we look at the role of the geopolitical risk of the 39 countries, we rely on two machine learning approaches. First, we use the least absolute shrinkage and selection operator (Lasso), proposed by Tibshirani (1996). The Lasso belongs to the spectrum of linear regression-analysis methods and performs both model selection and regularization in order to enhance the prediction accuracy and interpretability of the resulting forecasting model. Second, we switch to a nonlinear model and estimate random forests (Breiman, 2001), which, in turn, is a technique tailored to operate in settings featuring a large array of predictors, while simultaneously capturing the predictive value of any potential nonlinear links between the dependent variable and predictors, as well as any interaction effects between the predictors, as highlighted in the case of the relationship between gold market movements and geopolitical risk by Li et al. (2021) and Huang et al. (2023).

To the best of our knowledge, ours is the first paper to forecast real gold volatility over the monthly period from February 1915 to September 2021. Such a sample period allows us to cover the longest possible high-frequency (monthly) data available for gold, and the associated predictive impact of various historical global and country-specific geopolitical risk, and in the process makes our analysis immune to any sample-selection bias (Hollstein et al., 2021). Our paper can be considered to add to the relatively large literature associated with the forecasting gold-market developments based on a wide array of macroeconomic, financial, and behavioral predictors that rely on a large spectrum of linear and nonlinear univariate or multivariate models (see, for example, Pierdzioch et al., 2014a, 2014b, 2015a, 2015b, 2016a, 2020a; 2020b; Aye et al., 2015; Hassani et al., 2015; Sharma, 2016; Bonato et al., 2018; Nguyen et al., 2019; Dichtl, 2020, with the last paper in particular providing a detailed review). Our paper goes beyond this earlier research in that we use the information content of country-level disaggregate geopolitical risk. It also should be noted that geopolitical risk has been shown to lead several of the predictors considered thus far in this literature. The only somewhat related paper is that of Gupta et al. (2017), wherein the authors use a quantile predictive regression approach to analyze whether terror attacks predict gold returns. They find that terror attacks have predictive power for the lower and particularly the upper quantiles of the conditional distribution of gold returns over the sample period from January 1986 to December 2009, which clearly underlines the importance to account for a potential nonlinear link between gold returns and uncertainty in our forecasting experiment.

In line with the safe haven property, Baur and Smales (2020) and Huang et al. (2023) also provide evidence of the impact of geopolitical risk on the volatility of gold returns. The second-moment impact is not surprising, since higher (lower) geopolitical risk serves as negative (positive) news, and results in higher (lower) trading activity, which in turn can translate into higher (lower) volatility in the gold market (Baur, 2012). Realizing this, we also delve into the role of aggregate and disaggregate geopolitical risk in forecasting the conditional volatility of gold prices based on the same set of models used for gold returns. We obtain our metric of volatility by fitting to the gold returns data a standard generalized autoregressive conditional heteroskedasticity (GARCH) model and two recently-developed flexible variants of this model (Wu & Karmakar, 2021, 2023), namely the time-varying parameter GARCH (TVPGARCH) model and the non-parametric GARCH (NPGARCH) model. We then determine the "optimal" model as the one that produces the lowest forecast errors for the univariate process of volatility, i.e., squared returns. In this regard, it should be noted that Gkillas et al. (2020) have used in recent research a quantile-regression heterogeneous autoregressive realized volatility (QR-HAR-RV) model to show that global geopolitical risk have predictive power for realized gold-returns volatility (estimated from intraday data) mainly at a longer forecast horizon (when one accounts for the potential asymmetry of the loss function a forecaster might use to evaluate forecasts), over the daily period from December 1997 to May 2017. Clearly, our paper can be considered to add to this recent strand of research, and the large gold volatility forecasting literature in general,[1] by predicting the future evolution of more than a century of gold volatility, based on the information content of not only global but also country-specific geopolitical risk, using linear and nonlinear predictive regressions involving machine learning.

Our research is also linked to recent research by Li et al. (2023), who report that the information embedded in country-specific geopolitical risk is valuable for forecasting gold volatility. They use GARCH-Mixed Data Sampling (MIDAS) models to analyze the link between geopolitical risk and daily gold volatility. To this end, they conduct a country-by-country analysis and, in addition, extract information from the array of country-specific geopolitical risks by means of dimension-reduction techniques (like principal-components analysis). While we also study a large array of country-specific geopolitical risks, we use the Lasso estimator to reduce the dimension of one variant of our forecasting model. Importantly, we also study random forests, which do not require reducing the dimension of a forecasting model and,

[1] See Pierdzioch et al. (2016b), Salisu et al. (2020), and Luo et al. (2022) for detailed reviews in terms of predictors and alternative econometric frameworks used to forecast gold price volatility.

in addition, account in a data-driven way for potential nonlinear links between gold volatility and country-specific geopolitical risks as well as interaction effects among the latter. Moreover, we study a much longer sample period (1915–2021) than Li et al. (2023), who study data for the period 1985–2021. Finally, we model gold volatility by means of innovative conditional volatility models that have been developed in recent research (Wu & Karmakar, 2021, 2023).

The remainder of the paper is organized as follows: Sect. 2 outlines our dataset, Sect. 3 presents the methodologies, Sect. 4 discusses the results, and Sect. 5 concludes.

## 2 Data

We retrieved the real gold price in U.S. dollars from Macrotrends,[2] with the data starting from January 1915. With us working with log-returns, implies that the effective sample of our analyses covered February 1915. The sample period ends in September 2021, governed by data availability at the time of writing of this paper.

As far as our predictors are concerned, we rely on the work of Caldara and Iacoviello (2022), who create a new measure of adverse geopolitical events based on an analysis of newspaper articles covering geopolitical tensions. Their data starts in 1900. The geopolitical risk (GPR) index summarizes the results of an automated text search of the electronic archives of three newspapers (The New York Times, Chicago Tribune, and The Washington Post), with the authors calculating the index by counting the number of articles related to adverse geopolitical events in each newspaper for each month (as a share of the total number of news articles). The authors organize the search in eight categories: War Threats (Category 1), Peace Threats (Category 2), Military Buildups (Category 3), Nuclear Threats (Category 4), Terror Threats (Category 5), Beginning of War (Category 6), Escalation of War (Category 7), and Terror Acts (Category 8). Building on the results for these search categories, Caldara and Iacoviello (2022) also derive two sub-indexes, namely an index of geopolitical threats (GPT), which includes words belonging to categories 1–5 above, and an index of geopolitical acts (GPA), which covers words belonging to categories 6–8.[3]

Besides these global indexes, Caldara and Iacoviello (2022) construct country-specific indexes for 39 different countries,[4] For each of the 39 countries, the authors

---

[2] Internet address: https://www.macrotrends.net/.

[3] The data is available for download from the following internet page: https://www.matteoiacoviello.com//gpr.htm, where interested readers also can find further information on the construction of the index, graphs of the data, and links to the relevant literature.

[4] The countries are: North America: Canada, Mexico, US; South America: Argentina, Brazil, Chile, Colombia, Peru, Venezuela; Europe (North and East): Denmark, Finland, Norway, Russia, Sweden, Ukraine, the United Kingdom; Europe (South and West): Belgium, France, Germany, Italy, The Netherlands, Portugal, Spain, Switzerland; Middle East and Africa: Israel, Saudi Arabia, South Africa, Turkey; Asia and Oceania: Australia, China, Hong Kong, Japan, South Korea, The Philippines, Taiwan, Indonesia, India, Malaysia, Thailand. Data on the country-specific indexes can be downloaded from the following internet page: https://www.matteoiacoviello.com/gpr_country.htm.

build the country-specific index (calculated as a monthly share of newspaper articles) by counting the monthly share of all newspaper articles that meet the criterion for inclusion in the GPR index and, in addition, either mention the country name or the names of its major cities.

## 3 Methodologies

### 3.1 Forecasting Models

The purpose of a forecasting exercise is to use training data to estimate a general model

$$y_{t+h} = F(x_{1,t}, x_{2,t}, \dots, x_{n,t}), \tag{1}$$

and then to use the estimates along with updated predictors to compute a forecast of the dependent variable. As for notation, $y_t$ denotes the dependent variable, $x_{i,t}, i = 1, 2, \dots, n$ denote the $n$ predictors, $F$ is a function that links the dependent variable to the predictors, $t$ denotes time, $h$ denotes the forecast horizon, a hat over a variable denotes a forecast. The dependent variable in the context of our forecasting exercise is volatility, where we forecast average volatility when our forecast horizon is larger than one month, that is, when we set $h > 1$. Also, we construct the data such that the number of forecasts is the same for all forecast horizons.

In practice, a popular approach is to start with a small number of predictors and to use a simple linear approximation of the function, $F$. When only one predictor is considered, the resulting forecasting model is of the format

$$y_{t+h} = \beta_0 + \beta_1 x_{1,t} + e_{t+h}, \tag{2}$$

where $e_t$ denotes a disturbance term and $\beta_i, i = 0, 1$ are the coefficients to be estimated. Estimation of this model can be done by the ordinary least-squared (OLS) technique. When the only predictor used to set up this model is given by lagged returns (volatility) then a simple autoregressive model obtains. We shall use such simple autoregressive model as one of our benchmark models.

The advantage of having a simple benchmark model is that we can compare its forecasting performance with the performance of slightly more complex models of the format:

$$y_{t+h} = \beta_0 + \beta_1 x_{1,t} + \beta_2 x_{2,t} + e_{t+h}, \tag{3}$$

$$y_{t+h} = \beta_0 + \beta_1 x_{1,t} + \beta_2 x_{2,t} + \beta_3 x_{3,t} + e_{t+h}, \tag{4}$$

These somewhat more complex models retain the simple linear structure of the benchmark model, but add additional predictors on the right-hand side. In the context of our analysis, the additional predictors are either geopolitical risk (the case of one additional predictor) or geopolitical risk as decomposed into threats and acts (the case of two additional predictors). These two extended forecasting models can also be estimated by the OLS technique.

We can also use the various country-level sources of geopolitical risk to set up an even more advanced extended linear forecasting model, which can then be described by an equation of the following format:

$$y_{t+h} = \beta_0 + \beta_1 x_{1,t} + \beta_2 x_{2,t} + \beta_3 x_{3,t} + \beta_4 x_{4,t} + \cdots + \beta_n x_{n,t} + e_{t+h}. \tag{5}$$

While this forecasting model still can be estimated by the OLS techniques, things are likely to get unedifying when the number, $n$, of predictors in this equation gets large, which is the case when we study the various country-level sources of geopolitical risk. A model shrinkage and predictor selection technique can help at this stage of the empirical analysis. The technique that we shall use to this end is the least absolute shrinkage and selection operator (Lasso) estimator (Tibshirani, 1996). The Lasso technique chooses the coefficients, $\beta_i, i = 1, 2, \ldots n$ to be estimated so as to minimize the following expression (for a detailed discussion of the Lasso, see Hastie et al., 2009):

$$\sum_{t=1}^{N} \left( y_{t+h} - \beta_0 - \sum_{i=1}^{n} \beta_i x_{i,t} \right)^2 + \lambda \left( |\beta_1| + |\beta_2| + \cdots + |\beta_n| \right), \tag{6}$$

where $N$ denotes the number of observations used for estimation of the model. The Lasso estimator, thus, adds to the standard quadratic loss function of the OLS technique a penalty term. The shrinkage parameter, $\lambda$, defines the weight attached to this penalty terms, and the penalty term itself is given by the sum of the absolute values of the coefficients. In other words, the Lasso estimator uses the L1 norm of the coefficient vectors to shrink the dimension of the estimated forecasting model. Depending on the magnitude of the shrinkage parameter, $\lambda$, the Lasso estimator shrinks the magnitude of the coefficients, or even sets to zero some of the coefficients. In the latter case, the Lasso estimator can be interpreted as a predictor-selection technique.

The Lasso estimator, as all other forecasting models that we have discussed so far, retains the assumption that the link between the dependent variable and its predictors is of a linear form, that is, the assumption is that the function, $F$, is linear. Moreover, the Lasso estimator retains the assumption that the various predictors enter the forecasting model in a additive and separable format. These two assumptions are problematic in a situation when the links between the dependent variable and its various predictors may be nonlinear, and when accounting for potential interaction effects between the predictors could help to improve forecasting performance. As we shall demonstrate when we describe our empirical results in Sect. 4, nonlinearities are widespread in our data. Moreover, the country-specific sources of geopolitical risk are likely to interact because political agents do not act in isolation and the risks originating in one country are likely to infect allies and adversaries.

In order to capture potential nonlinearities in the data as well as interaction effects among the predictors, we use random forests (Breiman, 2001). A random forest belongs to the class of ensemble machine-learning technique because it additively combines a large number of individual regression trees, $T$. Hence, the basic idea is to approximate the function $F$, by means of an ensemble of $m$ regression trees as follows:

$$F(x_{1,t}, x_{2,t}, \dots, x_{n,t}) = \sum_i T_i, \quad i = 1, 2, \dots m \tag{7}$$

As every tree in a natural forest, an individual regression tree consists of a root and several nodes and branches, which subdivide the space of the predictors, $\mathbf{x} = (x_1, x_2, \dots)$, into $l$ non-overlapping regions, $R_l$. These regions, which can be interpreted as the terminal leaves of a regression tree, are formed by applying a search-and-split algorithm in a recursive top-down fashion (for a textbook exposition, see Hastie et al., 2009). In order to describe this search-and-split algorithm, we start at the root of a regression tree. Our aim is to define a node in an optimal way such that we subdivide the space of predictors into a left region (that is, a branch), $R_1$, and a right region, $R_2$. To this end, we iterate over all predictors and use (in the simplest case) every realization of a predictor as a candidate splitting point. For every combination of a predictor and a splitting point, $\{s, p\}$, we then define the left and right branches, $R_1(s, p) = \{x_s | x_s \leq p\}$ and $R_2(s, p) = \{x_s | x_s > p\}$. In order to find the optimal combination of a predictor and a splitting point, we minimize a standard squared-error loss function:

$$\min_{s,p} \left\{ \min_{\bar{y}_1} \sum_{x_s \in R_1(s,p)} (y_i - \bar{y}_1)^2 + \min_{\bar{y}_2} \sum_{x_s \in R_2(s,p)} (y_i - \bar{y}_2)^2 \right\}, \tag{8}$$

where $i$ identifies those realizations of the dependent variable that belong to a half-plane, $\bar{y}_k = \text{ mean } \{y_i x_s \in R_k(s, p)\}, k = 1, 2$ denotes the region-specific mean of the dependent variable, and where where we have dropped the time index and the index for the forecast horizon to keep the notation as simple as possible. Hence, the outer minimization runs over all combinations of $\{s, p\}$, and for every single one of those combinations the inner minimization optimally selects the branch-specific means of the dependent variable so as to minimize the branch-specific squared error loss. The result of this inner and outer minimization yields an optimal top-level optimal splitting predictor, an optimal top-level splitting point (and, thus, two branches), and the two branch-specific means of the dependent variable.

We already can use this rudimentary regression tree to forecast the dependent variable. To this end, we simply update the predictors, decide on whether the updated realization of the optimal top-level predictor belongs to the left or the right branch, and use the corresponding brach-specific mean of the dependent variable as our forecast. We can attempt to produce a better forecast, however, upon growing a larger regression tree. To this end, we apply the search-and-split algorithm to both the left and the right top-level branches, which gives us two second-level optimal splitting predictors and optimal splitting points, and four second-level branch-specific means of the dependent variable. Applying the search-and-split algorithm multiple times in an by now obvious way, we grow an increasingly complex regression tree. We stop this growth process when a regression tree has a preset maximum number of terminal nodes or every terminal branch has a minimum number of observations.

We now can use the complex regression tree that we have built in this way to send the predictors down the tree from its top level to the various leaves along the optimal partitioning points and branches. Equipped with this information, we then compute the optimal means of the dependent variable for the terminal regions and, hence, model the link between the dependent variable and the various predictors as follows[5]:

$$T\left(\mathbf{x}_i, \{R_l\}_1^L\right) = \sum_{l=1}^{L} \bar{y}_l \mathbf{1}(\mathbf{x}_i \in R_l), \tag{9}$$

where $L$ denotes the number of regions and $\mathbf{1}$ denotes the indicator function. Upon updating the predictors, sending them down the tree, and using the optimal means of the dependent variable, we then can use this equation to compute a forecast of the dependent variable.

An obvious problem of this approach to computing forecasts of the dependent variable is that the complex hierarchical structure of a regression tree gives results in an overfitting and data-sensitivity problem, which most likely deteriorates the forecasting performance of a regression tree. It is at this stage of the analysis that the concept of a random forest enters the stage. A random forest resolves the overfitting and data-sensitivity problem by combining a large number of individual regression trees to an ensemble of trees. This ensemble is formed by applying a three-step approach. In the first step, a large number of bootstrap samples is obtained by resampling from the data. In the second step, a regression tree is fitted to every single one of the bootstrapped samples. Importantly, the regression tree that is being fitted is a so-called random regression tree. A random regression tree uses for every splitting step only a random subset of the predictors. Injecting randomness into the splitting process in this way mitigates the effect of influential predictors on tree building. In the third step, we combine the large number of the resulting bootstrapped random trees and, use the decorrelated predictions of the dependent variable as obtained from the individual random regression trees. Moreover, averaging the predictions computed by means of the individual random regression trees stabilizes predictions.

We use the R language and environment for statistical computing (R Core Team, 2023) to set up our forecasting experiment. We use the R add-on package "glmnet" (Friedman et al., 2010) to implement the Lasso estimator, where we use 10-fold cross-validation to identify the optimal shrinkage parameter that minimizes the mean cross-validated error. We use the R add-on package "grf" (Tibshirani et al., 2021) to estimate random forests. In our forecasting experiment, a random forest is built from 1000 random regression trees (bootstrapping is done by sampling with replacement). 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

[5] It should be noted that Eq. (9) shows that random forests ensure that the forecasts of volatility are always non-negative, unlike volatility forecasts derived from models estimated by means of the ordinary least squares (OLS) technique. In our forecasting exercise, however, this is not a serious issue.

maximum imbalance of a split at a node of a tree). We refer a reader for further technical details to the extensive documentations of these packages.

### 3.2 Conditional Volatility Models

Because the purpose of our forecasting exercise is to forecast the conditional volatility of real log-returns of gold, we need to specify models that render it possible to estimate conditional volatility. To this end, we estimate three alternative models, and then identify the "optimal" model as the one that yields the lowest mean-squared error (MSE), $\sum(y_i^2 - \hat{\sigma}^2)^2$.

We start off with the standard GARCH(1,1) model:

$$y_t \sim N(0, \sigma_t^2) \quad \text{with} \quad \sigma_t^2 = \alpha_0 + \alpha_1 y_{t-1}^2 + \beta_1 \sigma_{t-1}^2. \tag{10}$$

We use the "fgarch" add-on package in R to obtain our maximum-likelihood parameter estimates of $\alpha_0, \alpha_1$ and $\beta_1$, so as eventually to obtain $\hat{\sigma}^2$.

We then estimate a time-varying parameter GARCH (TVPGARCH) as follows:

$$y_t \sim N(0, \sigma_t^2) \quad \text{with} \quad \sigma_t^2 = \alpha_0(t/n) + \alpha_1(t/n) y_{t-1}^2 + \beta_1(t/n) \sigma_{t-1}^2, \tag{11}$$

In order to estimate the time-varying parameter functions, $\alpha_0(\cdot), \alpha_1(\cdot)$, and $\beta_1(\cdot)$, we use the kernel-based method described in detail in Karmakar et al. (2021). For a suitable choice of kernel $K$ and bandwidth parameter, $b_n \in [0, 1]$, we estimate $\theta = (\alpha_0, \alpha_1, \beta)$ using

$$\hat{\theta}_{b_n}(t) = \text{ argmin}_{\theta \in \Theta} \sum_{i=1}^{n} K((t - i/n)/b_n) \ell(y_i, X_i, \theta), \qquad t \in [0, 1], \tag{12}$$

where $\ell(\cdot)$ is the corresponding negative log-likelihood or quasi log-likelihood for estimating the GARCH parameters, $X_i$ denotes the vector of covariates, which in the context of our empirical analysis will be $y_{i-1}$ for an univariate GARCH model. In particular, for our estimation problem, $\ell$ is of the following format:

$$\ell(y_i, X_i, \theta') = -\frac{1}{2} \log(\sigma_i^2) + y_i^2/\sigma_i^2 \quad \text{with} \quad \sigma_i^2 = \alpha_0 + \alpha_1 y_{i-1}^2 + \beta_1 \sigma_{i-1}^2, \tag{13}$$

and we choose an Epanechnikov Kernel to specify $K$. Finally, with the estimated function, $\alpha_0(\cdot), \alpha_1(\cdot)$, and $\beta_1(\cdot)$ in hand, we compute $\hat{\sigma}^2$.

As our third model, we estimate a non-parametric GARCH (NPGARCH) model. This is a relatively new model-free approach of fitting a GARCH model from the perspective of superior prediction performance than standard GARCH or GARCH-type models. Politis (2015) proposes a model-free approach that relaxes the normality assumption of traditional GARCH models, and rather tries to recover the error process from the reconstructed residuals. Building on that approach, Chen and Politis (2019) propose a method named GE-NoVas curated to ARCH models, which, in turn, paved the path for the parsimonious GE-NoVas approach described by Wu and Karmakar (2021). Adapted to the class of GARCH models, their approach led to the development of the parsimonious GARCH-NoVas (GA-NoVas in short) method by Wu and Karmakar (2023), which shows superior predictive performance among its class of model-free approaches, as well as significantly beating the traditional GARCH-based method. Formally the NPGARCH method can be described as follows. We start by rewriting the usual GARCH model

$$y_t = \psi_t \sigma_t \text{ with } \sigma_t^2 = a + \alpha_1 y_{t-1}^2 + \beta_1 \sigma_{t-1}^2, \text{ which implies } y_t = \psi_t \sqrt{\frac{a}{1-b_1} + \sum_{i=1}^{\infty} \alpha_1 \beta_1^{i-1} y_{t-i}^2} \tag{14}$$

As for the implementation, normality assumption on $\psi_t$ is relaxed by resampling

$$\hat{\psi}_t = \frac{y_t}{\sqrt{\frac{\beta y_t^2}{1-b_1} + \alpha s_{t-1}^2 + \sum_{i=1}^{q} \alpha_1 \beta_1^{i-1} y_{t-i}^2}}$$

for a large $q$ to generate future $y$ values. Above $s_{t-1}^2$ is the estimated variance using the data up to time period $t-1$. The parameters $\alpha, \beta, a, b$ are chosen to minimize $|Kurtosis(\hat{\psi}_{1:t}) - 3|$. In a way, one can think of this as a special time-varying method because the intercept term is changing, but it is clearly different from the one we use in the TVPGARCH framework.

The MSE obtained under the GARCH, TVPGARCH, and NPGARCH model for real log-returns of gold is equal to 3194.4030, 3420.2440, and 3184.9460, i.e., it is lowest for the NPGARCH model, although not by a large margin. We use the fitted variance computed by means of this model in the forecasting exercise we shall describe in Sect. 4 below when we forecast the volatility of real gold returns.[6]

## 4 Empirical Results

### 4.1 Baseline Results

Table 1 summarizes our baseline results in terms of the out-of-sample $R^2$ statistic, defined as $R^2 = 1 - \sum FE_R^2 / \sum FE_B^2$, where $FE_R$ ($FE_B$) denotes the forecast error of the rival (benchmark) model. A positive statistic, thus, indicates that the rival model performs better than the benchmark model. We report recursive-window and rolling-window estimates of the various models.As a sensitivity check, we report results for two training periods in the case of a recursive window, and two rolling-estimation window. Specifically, we use 60 (120) observations to initialize recursive-window estimation. Similarly, we use two alternative rolling-estimation windows of lengths 60 (120) observations.

[6] Complete estimation details of the parameters of the three models are available upon request from the authors.

**Table 1** Results for the out-of-sample $R^2$ statistic (gold)

| Benchmark/rival model | W | h = 1 | h = 3 | h = 6 | h = 12 |
|---|---|---|---|---|---|
| *Panel A: recursive* | | | | | |
| AR/AR + GPR | 60 | 0.0094 | 0.0216 | 0.0294 | 0.0365 |
| AR/AR + GPT and GPA | 60 | 0.0089 | 0.0179 | 0.0225 | 0.0240 |
| AR/AR + GPC (Lasso) | 60 | 0.0315 | – 0.0029 | – 0.0153 | – 0.0444 |
| AR/AR + GPC (RF) | 60 | 0.1897 | 0.2806 | 0.3906 | 0.4889 |
| AR + GPR/AR + GPT and GPA | 60 | – 0.0005 | – 0.0037 | – 0.0071 | – 0.0130 |
| AR + share/AR + GPC (Lasso) | 60 | 0.0224 | – 0.0250 | – 0.0461 | – 0.0840 |
| AR + share/AR + GPC (RF) | 60 | 0.1820 | 0.2647 | 0.3722 | 0.4695 |
| AR/AR + GPR | 120 | 0.0094 | 0.0216 | 0.0294 | 0.0365 |
| AR/AR + GPT and GPA | 120 | 0.0090 | 0.0180 | 0.0225 | 0.0240 |
| AR/AR + GPC (Lasso) | 120 | 0.0316 | – 0.0029 | – 0.0154 | – 0.0444 |
| AR/AR + GPC (RF) | 120 | 0.1897 | 0.2806 | 0.3906 | 0.4890 |
| AR + GPR/AR + GPT and GPA | 120 | – 0.0004 | – 0.0037 | – 0.0071 | – 0.0130 |
| AR + share/AR + GPC (Lasso) | 120 | 0.0224 | – 0.0250 | – 0.0461 | – 0.0840 |
| AR + share/AR + GPC (RF) | 120 | 0.1820 | 0.2647 | 0.3722 | 0.4696 |
| *Panel B: rolling* | | | | | |
| AR/AR + GPR | 60 | 0.2092 | – 0.1224 | – 0.0220 | – 0.0461 |
| AR/AR + GPT and GPA | 60 | 0.1319 | – 0.2565 | – 0.1137 | – 0.1369 |
| AR/AR + GPC (Lasso) | 60 | 0.5136 | – 0.5461 | – 0.4555 | – 0.9584 |
| AR/AR + GPC (RF) | 60 | 0.5579 | – 0.0458 | 0.0695 | 0.0216 |
| AR + GPR/AR + GPT and GPA | 60 | – 0.0979 | – 0.1195 | – 0.0897 | – 0.0869 |
| AR + share/AR + GPC (Lasso) | 60 | 0.3849 | – 0.3775 | – 0.4242 | – 0.8721 |
| AR + share/AR + GPC (RF) | 60 | 0.4409 | 0.0682 | 0.0896 | 0.0647 |
| AR/AR + GPR | 120 | – 0.0945 | – 0.0342 | – 0.0012 | – 0.0185 |
| AR/AR + GPT and GPA | 120 | 0.0072 | – 0.0555 | – 0.0350 | – 0.0905 |
| AR/AR + GPC (Lasso) | 120 | 0.1516 | – 0.2042 | – 0.1475 | – 0.0682 |
| AR/AR + GPC (RF) | 120 | 0.1950 | 0.0261 | 0.0531 | 0.1072 |
| AR + GPR/AR + GPT and GPA | 120 | 0.0929 | – 0.0206 | – 0.0338 | – 0.0707 |
| AR + share/AR + GPC (Lasso) | 120 | 0.2248 | – 0.1643 | – 0.1462 | – 0.0488 |
| AR + share/AR + GPC (RF) | 120 | 0.2644 | 0.0583 | 0.0542 | 0.1234 |

A positive out-of-sample $R^2$ statistic indicates that the rival model performs better than the benchmark model. W = Training window (recursive)/rolling window (rolling). h = Forecast horizon

Turning first to the results for returns and a recursive-estimation window, we find, when we extend a simple autoregressive (AR) benchmark model to include geopolitical risk (or the corresponding threats and acts), that the rival model outperforms the benchmark model by a small margin in case of a recursive-estimation window for both training periods and for all four forecast horizons that we study. In case of a rolling-estimation window, in contrast, most out-of-sample $R^2$ statistic take on negative values. We observe evidence of superior predictability mainly at the short forecast horizon. Similarly, the variants of the Lasso estimator produce better forecasts than the autoregressive benchmark model in terms of the out-of-sample $R^2$ statistic at the short forecast horizon only. For random forests, we observe positive out-of-sample $R^2$ statistics for both types of estimation windows, and for all four forecast horizons.[7] As a result, we observe that random forests produce the largest and, compared to the other models, most robust forecasting gains in terms of the out-of-sample $R^2$ statistic, irrespective of whether we study a recursive-estimation window or a rolling-estimation window.

Table 2 summarizes the results ($p$-values) the Diebold and Mariano (1995) test, as modified by Harvey et al. (1997). The test results for a recursive-estimation window are significant to a stronger extent than those for a rolling-estimation window. In fact, most test results for the latter are insignificant, with the test results for random forests being an exception. Random forests also yield significant test results for a recursive-estimation window. The Lasso estimator performs poorly relative to the respective benchmark models for both estimation windows. Splitting geopolitical risk into threats and actuals rather than simply extending the autoregressive benchmark model to include geopolitical risk does not yield significant test results either. Threats and actuals only yield significant test results when the benchmark model is the autoregressive model without geopolitical risk added, and only for the recursive-estimation window,[8]

The relatively good and robust performance of random forests in comparison to the Lasso estimator leads us to hypothesize that it is not only accounting for the country-sources of geopolitical risk that leverages forecasting performance, but rather that accounting also for potentially nonlinear links between returns and volatility and the country-specific sources of geopolitical risk, as well as potential interactions between the latter, matters for forecasting performance. The results summarized in Table 3 support this hypothesis. Random forests clearly outperform, and consistently so across forecast horizons and types of estimation window, the Lasso estimator in terms of the out-of-sample $R^2$ statistic, where the relative forecast performance shows a tendency to increase in the forecast horizon.

Figure 1 illustrates that volatility is linked to geopolitical risk (i) in a heterogeneous way across countries (and, thus, accounting for country-specific sources of geopolitical risk matters), and, (ii) these links can be strongly nonlinear. The figure plots partial dependence functions that visualize how volatility responds to a

[7] Because the country-specific GPR indexes are expressed as a monthly share of newspaper articles, when we study the Lasso and random forests, we use for our model comparisons the GPR index expressed as a share for the benchmark model (though the share, of course, is perfectly correlated with the GPR index).

[8] A potential problem of the Diebold-Mariano test, when applied to forecasts obtained from a recursive-estimation window is that, under the null hypothesis, the difference between the forecast errors of nested rival and benchmark models vanishes asymptotically. The fact that the results we report in Table 2 are in line with the results we report in Table 1 helps to build confidence in the results of the Diebold-Mariano test. It should also be noted that this problem does not arise in the case of a rolling-estimation window. Moreover, it is important to emphasize that, in the case of random forests, a comparison of the linear benchmark models and random forests in terms of statistical tests is complicated by the nonlinear and complex structure of random forests. The nonlinear and complex structure of random forests implies that the linear models are not simple nested versions of the random-forest model. It should also be noted that the Diebold-Mariano test has been used in other recent research the role of geopolitical risk indices for forecasting gold volatility (Li et al., 2023) so that presenting results for this test implies that our results can be better compared to the results of this recent research.

variation in the share of geopolitical risk that can be attributed to the US and China, two major powers on the international political scene.[9] The partial dependence functions for the US show that returns increase at a relatively low value of geopolitical risk that originates in the US except at the three-months forecast horizon. At the one month and 6 months forecast horizon, returns stay at this higher level when geopolitical risk increases further. For the long forecast horizon, we observe that returns decrease again to a lower level for intermediate to high values of geopolitical risk. For all four forecast horizons, returns are more or less insensitive to high values of US geopolitical risk. As for China, we find a clear pattern that returns first tend to drop in the region of very low geopolitical risk, but they then start increasing as geopolitical risk increases, while the partial dependence functions become more or less flat for high values of geopolitical risk. The partial dependence functions further show that volatility is lower for higher than for lower values of US geopolitical risk for the intermediate and long forecast horizons. For the short forecast horizon, we find a partial dependence function that exhibits a U-shaped pattern at low und intermediate values of US geopolitical risk. The partial dependence functions for Chinese geopolitical risk, in sharp contrast, witness that volatility clearly is increasing when geopolitical risk increases from a low to an intermediate value, and then stays at this higher level when geopolitical risk increases further.

Another question is whether the contribution of geopolitical risk and its country-specific sources to forecasting performance is mainly a phenomenon that can be observed during a view historical episodes, or whether forecasting gains are more evenly spread across the century-long sample period that we study in our empirical research. The results that Li et al. (2023) report, for example, indicate that forecasting performance may differ across low vs. high gold volatility episodes. To answer this question, we plot in Fig. 2 the recursively estimated out-of-sample $R^2$ statistic. The message conveyed by? this figure is clear: The good performance of random forests is not centered in any specific subinterval of our sample period. For the recursive-estimation window, the good performance of random forests shows a discernible trend to increase over time, it is more or less stable for the rolling-estimation window. It changes sign in case of the latter only when we consider the short rolling-estimation window, and only for a forecast horizon of three months and the combination AR/AR + GPC (RF).

Finally, we compare in Table 4 at the end of the paper ("Appendix") the performance of the forecasting models across conditional volatility models. As we emphasized in Sect. 3.2, based on the MSE criterion, we choose the NPGARCH model over the GARCH and TVPGARCH models as our preferred model of conditional gold volatility. The results we report in Table 4 show that our forecasting models, in the overwhelming majority of cases, perform better in terms of the out-of-sample $R^2$ statistic for the NPGARCH model of gold volatility than for the rival conditional volatility models. The only exception arises for $h = 1$, where the GARCH model performs better than the NPGARCH model in terms of the out-of-sample $R^2$ statistic.

[9] The estimation of the partial dependence functions use the full sample of data. We computed the partial dependence functions by means of the R add-on package randomForestSRC (Ishwaran & Kogalur, 2021) Sampling is with replacement, the minimum node size is five and one third of the predictors are used for splitting.

**Table 2** Results of the Diebold–Mariano test (gold)

| Benchmark/rival model | W | h = 1 | h = 3 | h = 6 | h = 12 |
|---|---|---|---|---|---|
| *Panel A: recursive* | | | | | |
| AR/AR + GPR | 60 | 0.0000 | 0.0000 | 0.0000 | 0.0001 |
| AR/AR + GPT and GPA | 60 | 0.0000 | 0.0001 | 0.0036 | 0.0377 |
| AR/AR + GPC (Lasso) | 60 | 0.2753 | 0.5079 | 0.5309 | 0.5643 |
| AR/AR + GPC (RF) | 60 | 0.0003 | 0.0000 | 0.0000 | 0.0000 |
| AR + GPR/AR + GPT and GPA | 60 | 0.5782 | 0.7843 | 0.8205 | 0.8855 |
| AR + share/AR + GPC (Lasso) | 60 | 0.3367 | 0.5665 | 0.5899 | 0.6166 |
| AR + share/AR + GPC (RF) | 60 | 0.0005 | 0.0000 | 0.0000 | 0.0000 |
| AR/AR + GPR | 120 | 0.0000 | 0.0000 | 0.0000 | 0.0001 |
| AR/AR + GPT and GPA | 120 | 0.0000 | 0.0001 | 0.0036 | 0.0373 |
| AR/AR + GPC (Lasso) | 120 | 0.2751 | 0.5080 | 0.5311 | 0.5643 |
| AR/AR + GPC (RF) | 120 | 0.0003 | 0.0000 | 0.0000 | 0.0000 |
| AR + GPR/AR + GPT and GPA | 120 | 0.5713 | 0.7807 | 0.8187 | 0.8851 |
| AR + share/AR + GPC (Lasso) | 120 | 0.3367 | 0.5665 | 0.5900 | 0.6165 |
| AR + share/AR + GPC (RF) | 120 | 0.0005 | 0.0000 | 0.0000 | 0.0000 |
| *Panel B: rolling* | | | | | |
| AR/AR + GPR | 60 | 0.2254 | 0.9739 | 0.8020 | 0.8172 |
| AR/AR + GPT and GPA | 60 | 0.3825 | 0.9806 | 0.9295 | 0.9102 |
| AR/AR + GPC (Lasso) | 60 | 0.1403 | 0.9414 | 0.9796 | 0.9402 |
| AR/AR + GPC (RF) | 60 | 0.1193 | 0.7105 | 0.1484 | 0.2392 |
| AR + GPR/AR + GPT and GPA | 60 | 0.6389 | 0.9167 | 0.9155 | 0.9273 |
| AR + share/AR + GPC (Lasso) | 60 | 0.0734 | 0.9246 | 0.9818 | 0.9417 |
| AR + share/AR + GPC (RF) | 60 | 0.0461 | 0.0712 | 0.0393 | 0.1080 |
| AR/AR + GPR | 120 | 0.9582 | 0.9331 | 0.5227 | 0.6990 |
| AR/AR + GPT and GPA | 120 | 0.4642 | 0.8119 | 0.8236 | 0.8729 |
| AR/AR + GPC (Lasso) | 120 | 0.2391 | 0.9899 | 0.9983 | 0.9305 |
| AR/AR + GPC (RF) | 120 | 0.1754 | 0.2051 | 0.0445 | 0.0396 |
| AR + GPR/AR + GPT and GPA | 120 | 0.2138 | 0.6239 | 0.7984 | 0.8980 |
| AR + share/AR + GPC (Lasso) | 120 | 0.1747 | 0.9776 | 0.9984 | 0.8797 |
| AR + share/AR + GPC (RF) | 120 | 0.1321 | 0.0681 | 0.0346 | 0.0131 |

Results (*p*-values) of the Diebold-Mariano test for alternative forecast horizons. The alternative hypothesis is that the rival model performs better than the benchmark model. Results are based on robust standard errors. W = Training window (recursive)/rolling window (rolling). h = Forecast horizon

Notably, random forests produce the largest forecasting benefits relative to the other forecasting models, and the out-of-sample $R^2$ statistic is positive also for $h = 1$ when we study random forests. In sum, these results lend further support to our choice of the NPGARCH model as our preferred model of conditional gold volatility.

## 4.2 Extension

As an extension, we report at the end of the paper ("Appendix") results for two other commodities namely, silver and oil, that have also recently been shown to

**Table 3** Comparing the Lasso and random forests (gold)

| Test (window type) | W | h = 1 | h = 3 | h = 6 | h = 12 |
|---|---|---|---|---|---|
| Out-of-sample $R^2$ (recursive) | 60 | 0.1633 | 0.2827 | 0.3998 | 0.5107 |
| Out-of-sample $R^2$ (rolling) | 60 | 0.0910 | 0.3236 | 0.3607 | 0.5004 |
| Out-of-sample $R^2$ (recursive) | 120 | 0.1633 | 0.2827 | 0.3999 | 0.5107 |
| Out-of-sample $R^2$ (rolling) | 120 | 0.0511 | 0.1912 | 0.1748 | 0.1641 |

Benchmark model: AR+ GPC (Lasso). Rival model: AR+ GPC (RF). A positive out-of-sample $R^2$ statistic indicates that random forests perform better than the Lasso estimator. W = Training window (recursive)/rolling window (rolling). h = Forecast horizon

act as hedges, if not safe-havens, against geopolitical risk (see for example, Bouoiyour et al., 2019; Baur and Smales, 2020; Li et al., 2020; Qin et al., 2020; Smales, 2021).[10] It should be noted that, while ours is the first paper to forecast historical real-returns volatility of silver based on geopolitical risk, there exists a few papers that have forecasted the returns and volatility of the oil market, but based on data spanning the last three decades or so (see, for example, Liu et al. 2019; Plakandaras et al. 2019; Asai et al. 2020; Mei et al. 2020; Salisu et al. 2021a).

For oil volatility, we observe that random forests often perform better than the competing models, especially when we study a rolling-estimation window. For a recursive-estimation window, the forecasting gains tend to be smaller than for a rolling-estimation window. For silver volatility, in contrast, random forests do not lead to an improvement in forecasting performance.

While the focus of our empirical analysis is on gold volatility, it is interesting to note that a comparative analysis of the results for oil, silver, and gold indicates that the safe-haven asset gold is different from oil and silver and, thus, that not all safe-haven assets behave alike with regard to geopolitical risk and its (country-specific) components. Similarly, a comparison of the findings for gold and silver yields the result that even these two precious metals should not be regarded as a single asset class, a result that is in line with observations made by earlier researchers (see, for example, Batten et al., 2010). This result can be further quantitatively substantiated by comparing the relative performance of the various forecasting models when applied to forecast gold and silver volatility (note that the sample for oil volatility starts earlier than for the other two commodities, complicating a direct quantitative cross-asset comparison of model performance). To this end, we define relative forecast errors (*RFE*, that is, forecast errors scaled by actual realizations) to ensure better comparability across the two precious metals, and then compute a modified

[10] The oil data starts in January 1900 and ends in September 2021. The sample period for silver is the same as that for gold. The data source for the nominal West Texas Intermediate (WTI) oil price in U.S. dollars and the U.S. consumer price Index used to deflate the nominal price to get real values are from Global Financial Data (https://globalfinancialdata.com/), and the datasource for real silver price is Macrotrends. As with gold, we work with the real log returns of these two commodities, and the optimal conditional volatility models (i.e., based on the minimum MSE among the GARCH, TVP-GARCH and NPGARCH models) for oil is the NPGARCH model (with MSE equal to 19965.9300 relative to 32658.5000 (GARCH) and 26905.0200 (TVPGARCH)), while that for silver is the standard GARCH model (with an MSE of 38626.0900 relative to 39249.1100 (TVPGARCH) and 41390.3100 (NPGARCH)).

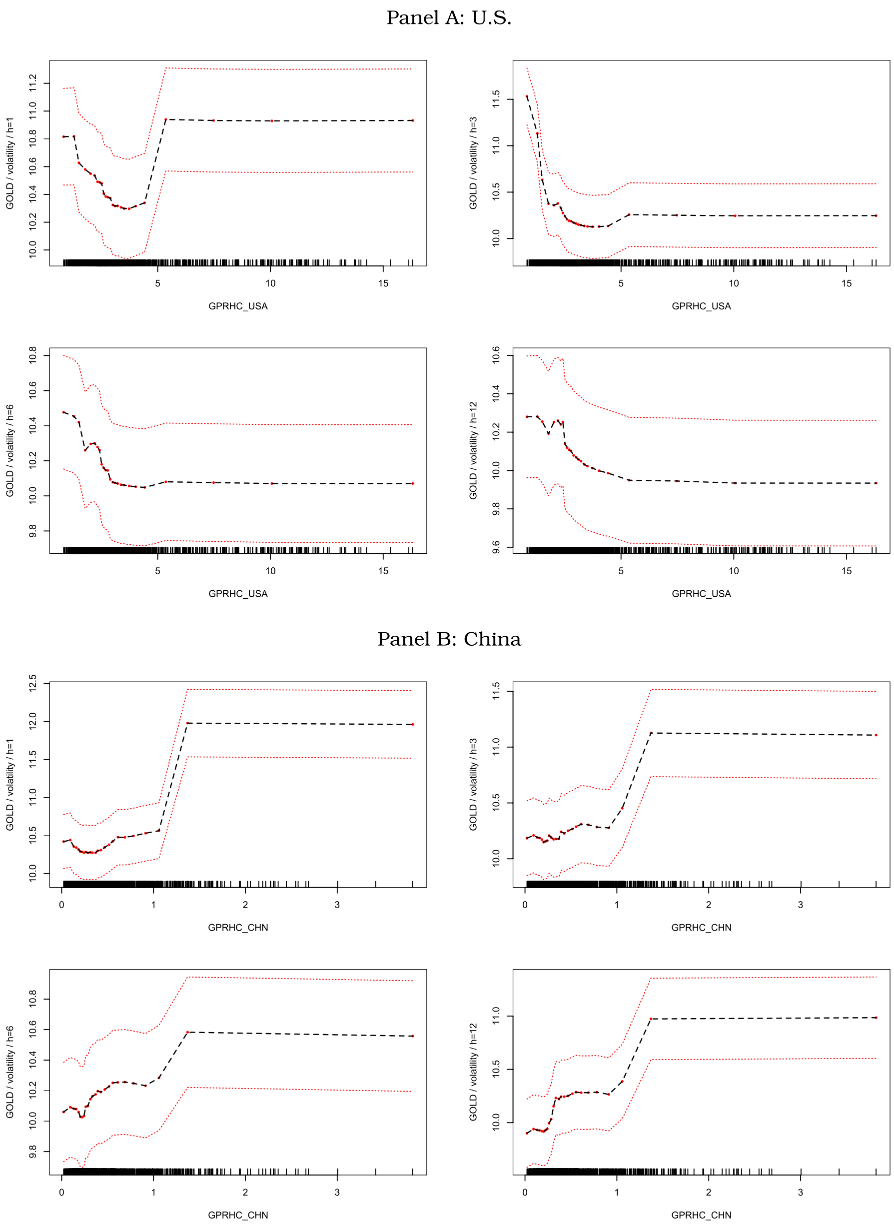


**Fig. 1** Examples of partial dependence functions (gold). *Note* The partial dependence functions (based on on-of-bag data) are estimated based on the full sample of data. Red points/black dashed lines: partial values. Dashed red lines: error band (plus/minus two standard errors). (Color figure online)

out-of-sample $R^2$ statistic as $R^2 = 1 - \sum RFE^2_{silver} / \sum RFE^2_{gold}$. The results (Table 5) show that the numerical value of the modified out-of-sample $R^2$ statistic decreases as the forecast horizon increases and eventually turns from a positive to a negative value for the long-forecast horizon, except when we combine a recursive-estimation window and the

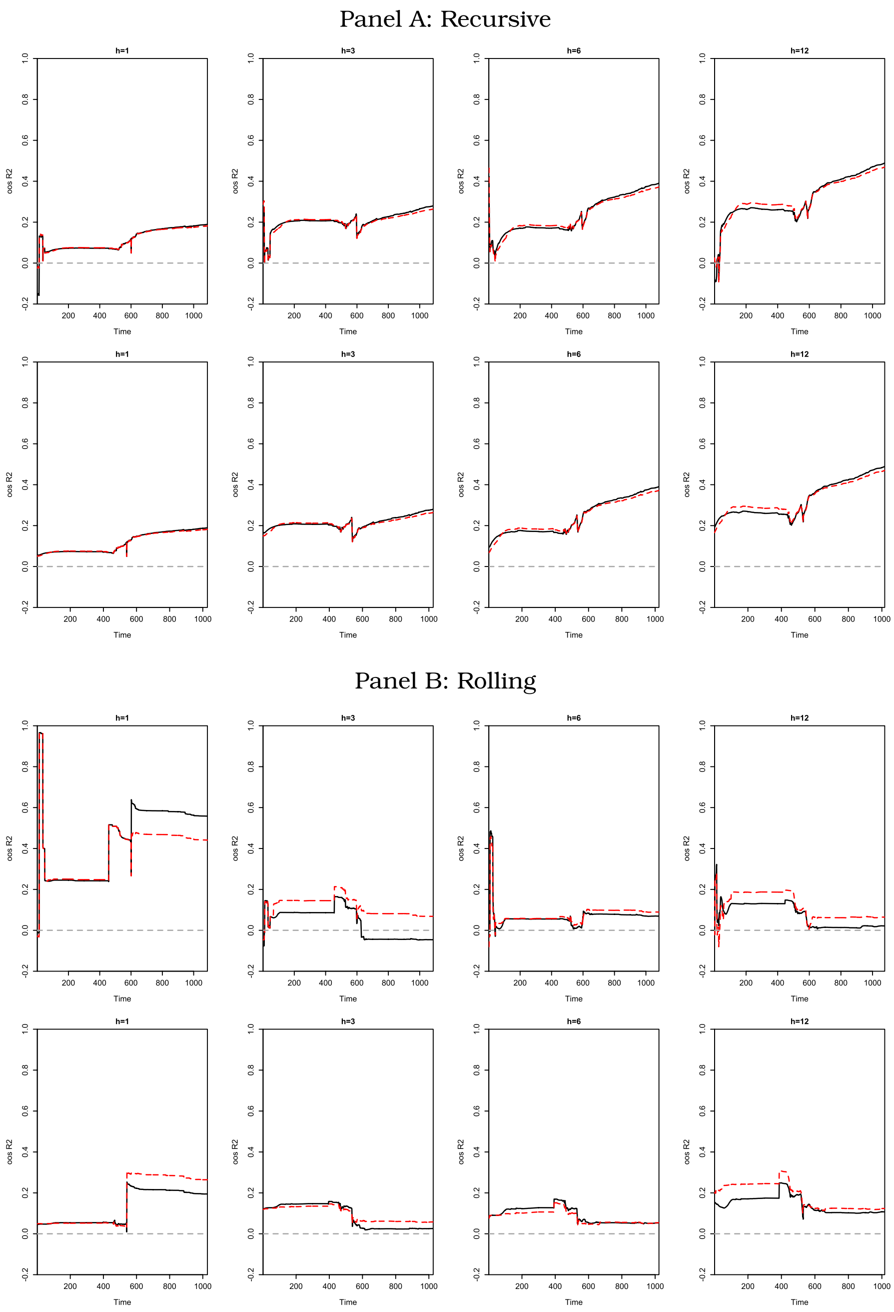


**Fig. 2** Stability of results (gold). *Note* The figure plots the recursively estimated out-of-sample $R^2$ statistics for a training window (rolling-window) of length 60 (120) observations in the upper (lower) rows of Panels A and B. The first 120 out-of-sample forecasts are used to initialize the estimations. The solid black line shows the results for the combination AR/AR + GPC (RF). The dashed red line shows the results for the combination AR + share/AR + GPC (RF). (Color figure online)

long training period. Interestingly, random forests are a main exception insofar as the modified out-of-sample $R^2$ statistic is negative irrespective of whether we study a recursive- or a rolling-estimation window (except at $h = 1$). Hence, the nonlinear random forests model tends to work consistently better for gold volatility than for silver volatility, while the results of this comparative analysis for the linear forecasting models depend to a stronger extent on the model configuration and the forecast horizon being studied. This result further underscores that viewing these two precious metals as belonging to a single asset class may be problematic.

## 5 Conclusion

Using country-specific sources of geopolitical risk is useful for forecasting gold volatility, especially when random forests, which account for potentially complex nonlinear data links and predictor interactions in a flexible and purely data-driven way, are used for model estimation and forecasting. Random forests often perform better than various linear alternative models, including forecasting models estimated using the Lasso estimator, a popular model shrinkage and predictor selection technique. Random forests also render it possible via the instrument of partial dependence functions to track more closely than is possible in the case of other methods how exactly gold returns and gold volatility respond to different levels of geopolitical risk. This information is particularly useful for forecasters and investors in the pricing of related derivatives as well as for devising hedging strategies involving gold investments as a safe haven in times of heightened geopolitical stress. Moreover, given that gold volatility tends to lead economic activity (Piffer & Podstawski, 2017; Çepni et al., 2021; Salisu et al., 2021b), the real effect of geopolitical events could end up being more persistent due to the existence of the indirect channel involving the gold market-geopolitical risk nexus. In light of this, accurate forecasts of the second-moment movements of gold returns due to geopolitical risks would help policymakers to design the size of their policy response to prevent deep recessions.

In future research, it will be interesting to extend our research along several avenues. First, one could study the contribution of geopolitical risk at the country-level to forecast accuracy in case of other commodities. Second, the implications of country-level geopolitical risk for other asset prices could also be investigated. In this regard, it would be interesting to study, for example, whether relative country-specific risk helps to improve the accuracy of forecasts of exchange-rate changes and volatility. Yet another avenue for future research would be to analyze the safe-haven property of gold (oil and silver) using other machine-learning techniques besides random forests used in this paper.

## Appendix

See Figs. 3 and 4, Tables 4 and 5.

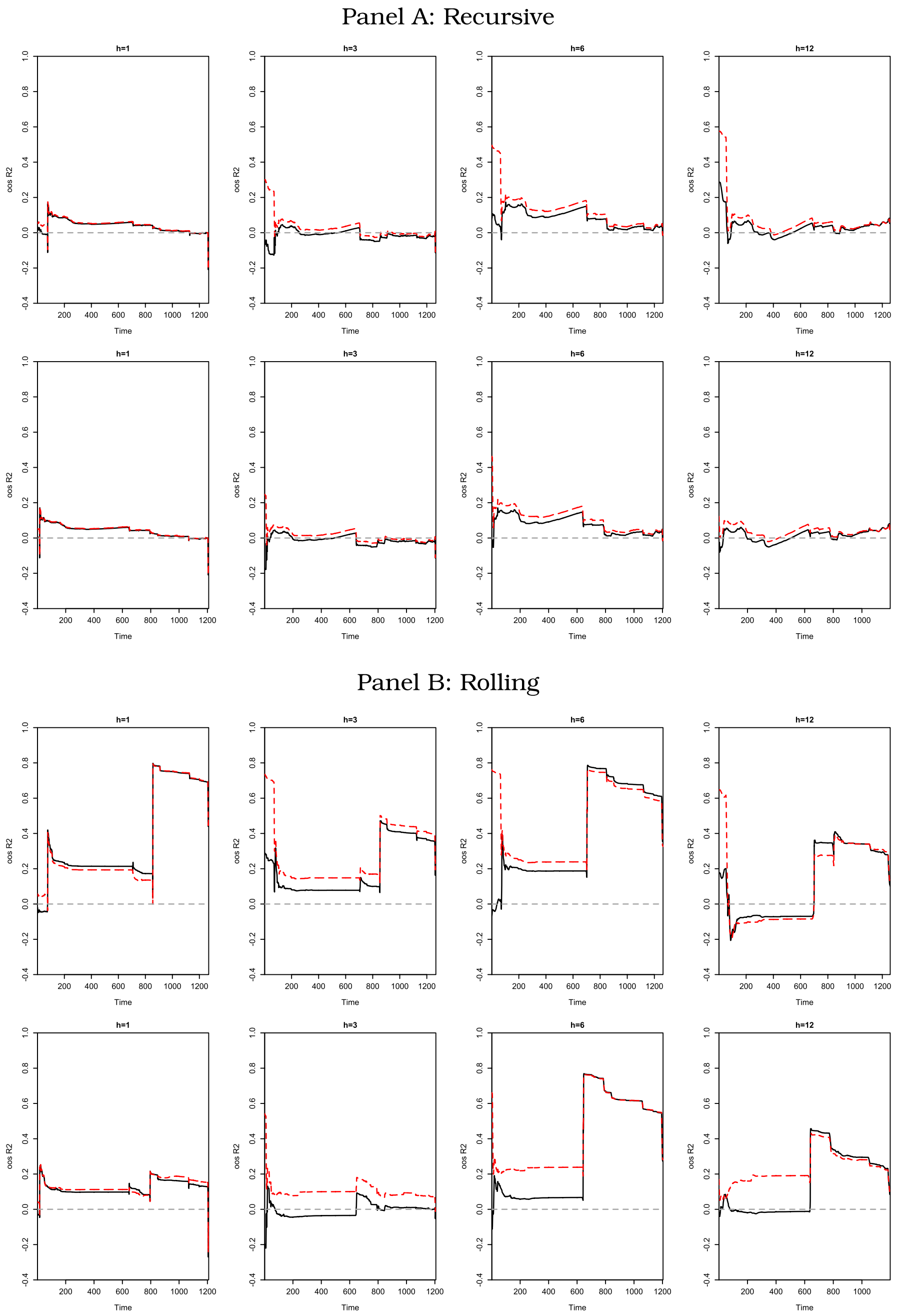


**Fig. 3** Results for the out-of-sample $R^2$ Statistic (Oil). *Note* The figure plots the recursively estimated out-of-sample $R^2$ statistics for a training window (rolling-window) of length 60 (120) observations in the upper (lower) rows of Panels A and B. The first 120 out-of-sample forecasts are used to initialize the estimations. The solid black line shows the results for the combination AR/AR + GPC (RF). The dashed red line shows the results for the combination AR + share/AR + GPC (RF). (Color figure online)

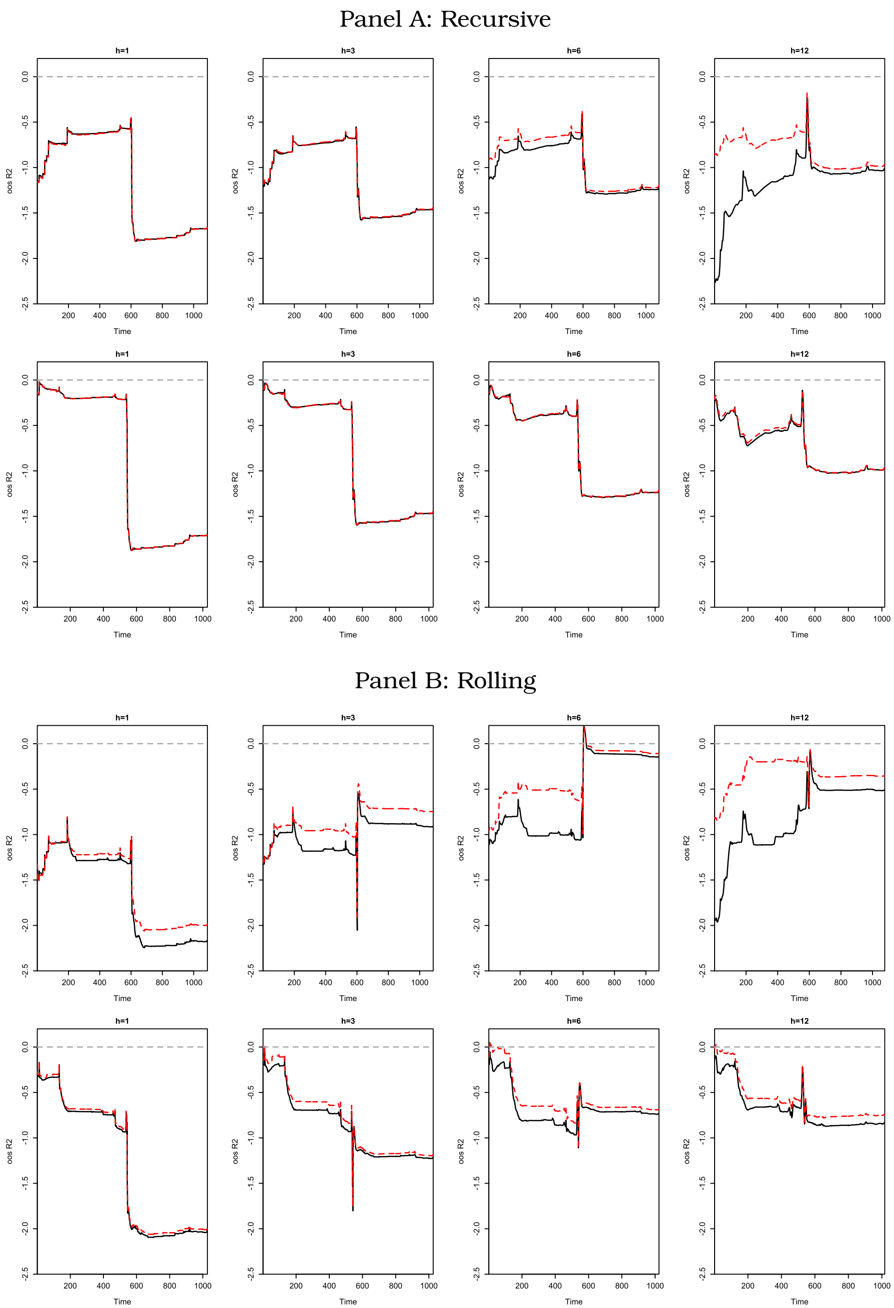


**Fig. 4** Results for the out-of-sample $R^2$ statistic (Silver). *Note*: The figure plots the recursively estimated out-of-sample $R^2$ statistics for a training window (rolling-window) of length 60 (120) observations in the upper (lower) rows of Panels A and B. The first 120 out-of-sample forecasts are used to initialize the estimations. The solid black line shows the results for the combination AR/AR + GPC (RF). The dashed red line shows the results for the combination AR + share/AR + GPC (RF). (Color figure online)

**Table 4** Comparing model performance across conditional volatility models (gold)

| Model | W | h = 1 | h = 3 | h = 6 | h = 12 |
|---|---|---|---|---|---|
| *Panel A: recursive* | | | | | |
| NPGARCH vs. GARCH: AR(1) | 60 | − 0.8771 | 0.4003 | 0.7133 | 0.8411 |
| NPGARCH vs. GARCH: AR(1) + GPR | 60 | − 0.8571 | 0.4140 | 0.7217 | 0.8449 |
| NPGARCH vs. GARCH: AR(1) + GPT and GPA | 60 | − 0.8494 | 0.4157 | 0.7212 | 0.8433 |
| NPGARCH vs. GARCH: AR(1) + GPR (Lasso) | 60 | − 0.7252 | 0.4489 | 0.7561 | 0.8621 |
| NPGARCH vs. GARCH: AR(1) + GPC (RF) | 60 | 0.3548 | 0.8251 | 0.9205 | 0.9560 |
| NPGARCH vs. GARCH: AR(1) | 120 | − 0.8771 | 0.4003 | 0.7134 | 0.8411 |
| NPGARCH vs. GARCH: AR(1) + GPR | 120 | − 0.8570 | 0.4141 | 0.7217 | 0.8449 |
| NPGARCH vs. GARCH: AR(1) + GPT and GPA | 120 | − 0.8493 | 0.4157 | 0.7212 | 0.8433 |
| NPGARCH vs. GARCH: AR(1) + GPR (Lasso) | 120 | − 0.7252 | 0.4489 | 0.7561 | 0.8621 |
| NPGARCH vs. GARCH: AR(1) + GPC (RF) | 120 | 0.3548 | 0.8251 | 0.9205 | 0.9560 |
| NPGARCH vs. TVPGARCH: AR(1) | 60 | 0.2150 | 0.7075 | 0.8291 | 0.8475 |
| NPGARCH vs. TVPGARCH: AR(1) + GPR | 60 | 0.2224 | 0.7137 | 0.8340 | 0.8524 |
| NPGARCH vs. TVPGARCH: AR(1) + GPT and GPA | 60 | 0.2229 | 0.7135 | 0.8334 | 0.8511 |
| NPGARCH vs. TVPGARCH: AR(1) + GPR (Lasso) | 60 | 0.4502 | 0.7466 | 0.8743 | 0.8984 |
| NPGARCH vs. TVPGARCH: AR(1) + GPC (RF) | 60 | 0.6552 | 0.8763 | 0.9374 | 0.9635 |
| NPGARCH vs. TVPGARCH: AR(1) | 120 | 0.2150 | 0.7074 | 0.8291 | 0.8474 |
| NPGARCH vs. TVPGARCH: AR(1) + GPR | 120 | 0.2223 | 0.7137 | 0.8340 | 0.8523 |
| NPGARCH vs. TVPGARCH: AR(1) + GPT and GPA | 120 | 0.2229 | 0.7135 | 0.8334 | 0.8511 |
| NPGARCH vs. TVPGARCH: AR(1) + GPR (Lasso) | 120 | 0.4501 | 0.7465 | 0.8743 | 0.8984 |
| NPGARCH vs. TVPGARCH: AR(1) + GPC (RF) | 120 | 0.6552 | 0.8763 | 0.9374 | 0.9635 |
| *Panel B: rolling* | | | | | |
| NPGARCH vs. GARCH: AR(1) | 60 | − 1.3246 | 0.7732 | 0.9182 | 0.9684 |
| NPGARCH vs. GARCH: AR(1) + GPR | 60 | − 0.7386 | 0.7536 | 0.9160 | 0.9700 |
| NPGARCH vs. GARCH: AR(1) + GPT and GPA | 60 | − 0.8872 | 0.7159 | 0.8894 | 0.9565 |
| NPGARCH vs. GARCH: AR(1) + GPR (Lasso) | 60 | − 0.1024 | 0.6449 | 0.9049 | 0.9380 |
| NPGARCH vs. GARCH: AR(1) + GPC (RF) | 60 | 0.5561 | 0.8662 | 0.9468 | 0.9697 |
| NPGARCH vs. GARCH: AR(1) | 120 | − 0.7670 | 0.6200 | 0.8496 | 0.9329 |
| NPGARCH vs. GARCH: AR(1) + GPR | 120 | − 0.7740 | 0.6425 | 0.8593 | 0.9363 |
| NPGARCH vs. GARCH: AR(1) + GPT and GPA | 120 | − 0.5694 | 0.6445 | 0.8560 | 0.9348 |
| NPGARCH vs. GARCH: AR(1) + GPR (Lasso) | 120 | − 0.1521 | 0.7238 | 0.8806 | 0.9431 |
| NPGARCH vs. GARCH: AR(1) + GPC (RF) | 120 | 0.4705 | 0.8454 | 0.9296 | 0.9610 |
| NPGARCH vs. TVPGARCH: AR(1) | 60 | − 0.2809 | 0.8068 | 0.9041 | 0.9507 |
| NPGARCH vs. TVPGARCH: AR(1) + GPR | 60 | 0.0042 | 0.7872 | 0.8991 | 0.9507 |
| NPGARCH vs. TVPGARCH: AR(1) + GPT and GPA | 60 | − 0.0718 | 0.7660 | 0.8939 | 0.9512 |
| NPGARCH vs. TVPGARCH: AR(1) + GPR (Lasso) | 60 | 0.5370 | 0.7882 | 0.9232 | 0.9292 |
| NPGARCH vs. TVPGARCH: AR(1) + GPC (RF) | 60 | 0.6978 | 0.8916 | 0.9507 | 0.9709 |
| NPGARCH vs. TVPGARCH: AR(1) | 120 | 0.2178 | 0.7918 | 0.8967 | 0.9240 |
| NPGARCH vs. TVPGARCH: AR(1) + GPR | 120 | 0.1642 | 0.7979 | 0.9032 | 0.9306 |
| NPGARCH vs. TVPGARCH: AR(1) + GPT and GPA | 120 | 0.2550 | 0.8012 | 0.9076 | 0.9344 |
| NPGARCH vs. TVPGARCH: AR(1) + GPR (Lasso) | 120 | 0.5493 | 0.8208 | 0.9416 | 0.9556 |
| NPGARCH vs. TVPGARCH: AR(1) + GPC (RF) | 120 | 0.6924 | 0.8957 | 0.9460 | 0.9677 |

**Table 4** (continued)

A positive modified out-of-sample $R^2$ statistic indicates that a forecasting model performs better for the NPGARCH model than for the rival model of conditional gold volatility. W = Training window (recursive)/rolling window (rolling). h = Forecast horizon

**Table 5** Comparing model performance (gold and silver)

| Model | W | h = 1 | h = 3 | h = 6 | h = 12 |
|---|---|---|---|---|---|
| *Panel A: recursive* | | | | | |
| AR(1) | 60 | 0.8901 | 0.6137 | 0.2836 | – 0.1779 |
| AR(1) + GPR | 60 | 0.8775 | 0.5566 | 0.1623 | – 0.4748 |
| AR(1) + GPT and GPA | 60 | 0.8811 | 0.5625 | 0.1239 | – 0.5246 |
| AR(1) + GPR (Lasso) | 60 | 0.8303 | 0.5284 | – 0.2626 | – 2.6662 |
| AR(1) + GPC (RF) | 60 | 0.0948 | – 1.1860 | – 3.3895 | – 5.3836 |
| AR(1) | 120 | 0.9006 | 0.6630 | 0.4003 | 0.0542 |
| AR(1) + GPR | 120 | 0.8981 | 0.6596 | 0.4039 | 0.0722 |
| AR(1) + GPT and GPA | 120 | 0.9018 | 0.6655 | 0.3790 | 0.0585 |
| AR(1) + GPR (Lasso) | 120 | 0.8547 | 0.6214 | – 0.0435 | – 2.2095 |
| AR(1) + GPC (RF) | 120 | 0.4273 | – 0.1539 | – 1.3245 | – 3.1901 |
| *Panel B: rolling* | | | | | |
| AR(1) | 60 | 0.9943 | 0.5865 | – 0.5083 | – 2.2670 |
| AR(1) + GPR | 60 | 0.9915 | 0.5391 | – 1.0653 | – 3.8144 |
| AR(1) + GPT and GPA | 60 | 0.9944 | 0.7461 | – 0.9735 | – 4.6681 |
| AR(1) + GPR (Lasso) | 60 | 0.6997 | 0.1910 | – 2.0114 | – 5.3348 |
| AR(1) + GPC (RF) | 60 | – 0.1009 | – 2.2273 | – 5.8749 | – 7.7669 |
| AR(1) | 120 | 0.9367 | 0.5476 | 0.2055 | – 0.7154 |
| AR(1) + GPR | 120 | 0.9459 | 0.4360 | 0.0104 | – 0.7962 |
| AR(1) + GPT and GPA | 120 | 0.9159 | 0.2646 | – 0.3477 | – 1.6775 |
| AR(1) + GPR (Lasso) | 120 | 0.6733 | – 0.0316 | – 3.2036 | – 10.4890 |
| AR(1) + GPC (RF) | 120 | 0.1448 | – 1.2914 | – 3.0796 | – 4.9427 |

A positive modified out-of-sample $R^2$ statistic indicates that a forecasting model performs better for silver than for gold. W = Training window (recursive)/rolling window (rolling). h = Forecast horizon

**Author Contributions** All authors contributed to the study conception, design, and analysis. The first draft of the manuscript was written by all authors, and and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

**Funding** Open Access funding enabled and organized by Projekt DEAL. The authors declare that no funds, grants, or other support were received during the preparation of this manuscript.

**Data Availability** The authors declare that they will make available the data and computer used to derive the results documented in this research upon request.

## Declarations

**Conflict of interest** The authors have no relevant financial or non-financial interests to disclose.

## References

Agyei-Ampomah, S., Gounopoulos, D., & Mazouz, K. (2014). Does gold offer a better protection against sovereign debt crisis than other metals? *Journal of Banking and Finance, 40*(C), 507–521.

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. *International Journal of Forecasting, 36*(3), 933–948.

Aye, G. C., Gupta, R., Hammoudeh, S., & Kim, W.-J. (2015). Forecasting the price of gold using dynamic model averaging. *International Review of Financial Analysis, 41*(C), 257–266.

Balcilar, M., Bonato, M., Demirer, R., & Gupta, R. (2017). The effect of investor sentiment on gold market return dynamics: Evidence from a nonparametric causality-in-quantiles approach? *Resources Policy, 51*(C), 77–84.

Balcilar, M., Bonato, M., Demirer, R., & Gupta, R. (2018). Geopolitical risks and stock market dynamics of the BRICS. *Economic Systems, 42*(2), 295–306.

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

Balcilar, M., Gupta, R., & Pierdzioch, C. (2016). Does uncertainty move the gold price? New evidence from a nonparametric causality-in-quantiles test. *Resources Policy, 49*(C), 74–80.

Batten, J. A., Ciner, C., & Lucey, B. M. (2010). The macroeconomic determinants of volatility in precious metals markets. *Resources Policy, 35*(2), 65–71.

Baur, D. G. (2012). Asymmetric volatility in the gold market. *Journal of Alternative Investments, 14*(4), 26–38.

Baur, D. G., & Lucey, B. M. (2010). Is gold a hedge or a safe haven? An analysis of stocks, bonds and gold. *Financial Review, 45*(2), 217–229.

Baur, D. G., & McDermott, T. K. (2010). Is gold a safe haven? International evidence. *Journal of Banking and Finance, 34*(8), 1886–1898.

Baur, D. G., & Smales, L. A. (2020). Hedging geopolitical risk with precious metals. *Journal of Banking and Finance, 117*(C), 105823.

Beckmann, J., Berger, T., & Czudaj, R. (2015). Does gold act a hedge or safe haven for stocks? A smooth transition approach. *Economic Modelling, 48*(C), 16–24.

Beckmann, J., Berger, T., & Czudaj, R. (2019). Gold Price dynamics and the role of uncertainty. *Quantitative Finance, 19*(4), 663–681.

Bonato, M., Demirer, R., Gupta, R., & Pierdzioch, C. (2018). Gold futures returns and realized moments: A forecasting experiment using a quantile-boosting approach. *Resources Policy, 57*(C), 196–212.

Boubaker, H., Cunado, J., Gil-Alana, L. A., & Gupta, R. (2020). Global crises and gold as a safe haven: Evidence from over seven and a half centuries of data. *Physica A: Statistical Mechanics and its Applications, 540*(C), 123093.

Bouoiyour, J., Selmi, R., Hammoudeh, S., & Wohar, M. E. (2019). What are the categories of geopolitical risks that could drive oil prices higher? Acts or threats? *Energy Economics, 84*(C), 104523.

Bouoiyour, J., Selmi, R., & Wohar, M. E. (2018). Measuring the response of gold prices to uncertainty: An analysis beyond the mean. *Economic Modelling, 75*(C), 105–116.

Bouri, E., Çepni, O., Gabauer, D., & Gupta, R. (2021). Return connectedness across asset classes around the COVID-19 outbreak? *International Review of Financial Analysis, 73*(C), 101646.
Bouri, E., Çepni, O., Gupta, R., & Jalkh, N. (Forthcoming). Geopolitical risks and stock market volatility in the G7 countries: A century of evidence from a time-varying nonparametric panel data model. *Handbook for the Economics of Terrorism, Edited by Atin Basuchoudhary and Gunther G. Schulze*.
Bouri, E., Demirer, R., Gupta, R., & Marfatia, H. A. (2019). Geopolitical risks and movements in Islamic bond and equity markets: A note. *Defence and Peace Economics, 30*(3), 367–379.
Breiman, L. (2001). Random forests. *Machine Learning, 45*, 5–32.
Caldara, D., & Iacoviello, M. (2022). Measuring geopolitical risk. *American Economic Review, 112*(4), 1194–1225.
Carney, M. (2016). Uncertainty, the economy and policy. Speech at the Bank of England, London, 30 June 2016. https://www.bis.org/review/r160704c.pdf.
Çepni, O., Dul, W., Gupta, R., & Wohar, M. E. (2021). The dynamics of U.S. REITs returns to uncertainty shocks: A proxy SVAR approach. *Research in International Business and Finance, 58*(C), 101433.
Chen, J., & Politis, D. N. (2019). Optimal multi-step-ahead prediction of arch/garch models and novas transformation. *Econometrics, 7*(3), 34.
Cheng, C. H. J., & Chiu, C.-W.J. (2018). How important are global geopolitical risks to emerging countries? *International Economics, 156*(C), 305–325.
Clance, M. W., Gupta, R., & Wohar, M. E. (2019). Geopolitical risks and recessions in a panel of advanced economies: Evidence from over a century of data. *Applied Economics Letters, 26*(16), 1317–1321.
Diebold, F. X., & Mariano, R. S. (1995). Comparing predictive accuracy. *Journal of Business and Economic Statistics, 13*(3), 253–263.
Ding, Z., & Zhang, X. (2021). The impact of geopolitical risk on systemic risk spillover in commodity market: An EMD-based network topology approach. *Complexity, 2021*(Complexity in Economics and Business), 2226944.
Dichtl, H. (2020). Forecasting excess returns of the gold market: Can we learn from stock market predictions? *Journal of Commodity Markets, 19*(C), 100106.
Friedman, J., Hastie, T., & Tibshirani, R. (2010). Regularization paths for generalized linear models via coordinate descent. *Journal of Statistical Software, 33*(1), 1–22.
Gkillas, K., Gupta, R., & Pierdzioch, C. (2020). Forecasting realized gold volatility: Is there a role of geopolitical risks? *Finance Research Letters, 35*(C), 101280.
Gozgor, G., Lau, C. K. M., Sheng, X., & Yarovaya, L. (2019). The role of uncertainty measures on the returns of gold. *Economics Letters, 185*(C), 108680.
Gupta, R., Majumdar, A., Nel, J., & Subramaniam, S. (2021). Geopolitical risks and the high-frequency movements of the US term structure of interest rates. *Annals of Financial Economics, 16*(3), 2150012.
Gupta, R., Majumdar, A., Pierdzioch, C., & Wohar, M. E. (2017). Do terror attacks predict gold returns? Evidence from a quantile-predictive-regression approach. *Quarterly Review of Economics and Finance, 65*(C), 276–284.
Gürgün, G., & Ünalmis, I. (2014). Is gold a safe haven against equity market investment in emerging and developing countries? *Finance Research Letters, 11*(4), 341–348.
Harvey, D., Leybourne, S., & Newbold, P. (1997). Testing the equality of prediction mean squared errors. *International Journal of Forecasting, 13*(2), 281–291.
Hassani, H., Silva, E. S., Gupta, R., & Segnon, M. K. (2015). Forecasting the price of gold. *Applied Economics, 47*(39), 4141–4152.
Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning: Data mining, inference, and prediction* (Vol. 2). New York, NY, USA: Springer.
Hollstein, F., Prokopczuk, J., Tharann, B., & Wese Simen, C. (2021). Predictability in commodity markets: Evidence from more than a century. *Journal of Commodity Markets, 24*(C), 100171.
Huang, J., Li, Y., Suleman, M. T., & Zhang, H. (2023). Effects of geopolitical risks on gold market return dynamics: Evidence from a nonparametric causality-in-quantiles approach. *Defence and Peace Economics, 34*(3), 308–322.
Huynh, T. L. D. (2020). The effect of uncertainty on the precious metals market: New insights from transfer entropy and neural network VAR. *Resources Policy, 66*(C), 101623.
Ishwaran, H., & Kogalur, U. B. (2021). Fast unified random forests for survival, regression, and classification (RF-SRC). *R Package version, 2*(12), 1.

Karmakar, S., Richter, S., & Wu, W. B. (2021). Simultaneous inference for time-varying models. *Journal of Econometrics*. https://doi.org/10.1016/j.jeconom.2021.03.002

Li, B., Chang, C.-P., Chu, Y., & Sui, B. (2020). Oil prices and geopolitical risks: What implications are offered via multi-domain investigations? *Energy and Environment, 31*(3), 492–516.

Li, Y., Huang, J., & Chen, J. (2021). Dynamic spillovers of geopolitical risks and gold prices: New evidence from 18 emerging economies. *Resources Policy, 70*(C), 101938.

Liu, J., Ma, F., Tang, Y., & Zhang, Y. (2019). Geopolitical risk and oil volatility: A new insight. *Energy Economics, 84*(C), 104548.

Li, X., Guo, Q., Liang, C., & Umar, M. (2023). Forecasting gold volatility with geopolitical risk indices. *Research in International Business and Finance, 64*(C), 101857.

Low, R. K. Y., Yao, Y., & Faff, R. (2016). Diamonds vs. precious metals: What shines brightest in your investment portfolio? *International Review of Financial Analysis, 43*(C), 1–14.

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

Mei, D., Ma, F., Liao, Y., & Wang, L. (2020). Geopolitical risk uncertainty and oil future volatility: Evidence from MIDAS models. *Energy Economics, 86*(C), 104624.

Nguyen, D. B. B., Prokopczuk, M., & Wese Simen, C. (2019). The risk premium of gold. *Journal of International Money Finance, 94*(C), 140–159.

Plakandaras, V., Gupta, R., & Wong, W.-K. (2019). Point and density forecasts of oil returns: The role of geopolitical risks. *Resources Policy, 62*(C), 580–587.

Pierdzioch, C., & Gupta, R. (2020a). Uncertainty and forecasts of U.S. recessions. *Studies in Nonlinear Dynamics and Econometrics, 24*(4), 20180083.

Pierdzioch, C., & Risse, M. (2020b). Forecasting precious metal returns with multivariate random forests. *Empirical Economics, 58*(3), 1167–1184.

Pierdzioch, C., Risse, M., & Rohloff, S. (2014a). On the efficiency of the gold market: Results of a real-time forecasting approach. *International Review of Financial Analysis, 32*(C), 95–108.

Pierdzioch, C., Risse, M., & Rohloff, S. (2014b). The international business cycle and gold-price fluctuations. *Quarterly Review of Economics and Finance, 54*(2), 292–305.

Pierdzioch, C., Risse, M., & Rohloff, S. (2015a). A real-time quantile-regression approach to forecasting gold returns under asymmetric loss. *Resources Policy, 45*(C), 299–306.

Pierdzioch, C., Risse, M., & Rohloff, S. (2015b). Forecasting gold-price fluctuations: A real-time boosting approach. *Applied Economics Letters, 22*(1), 46–50.

Pierdzioch, C., Risse, M., & Rohloff, S. (2016a). A quantile-boosting approach to forecasting gold returns. *North American Journal of Economics and Finance, 35*(C), 38–55.

Pierdzioch, C., Risse, M., & Rohloff, S. (2016b). A boosting approach to forecasting the volatility of gold-price fluctuations under flexible loss. *Resources Policy, 47*(C), 95–107.

Piffer, M., & Podstawski, M. (2017). Identifying uncertainty shocks using the price of gold. *Economic Journal, 128*(616), 3266–3284.

Politis, D. N. (2015). The model-free prediction principle. In: *Model-free prediction and regression* (pp. 13–30). Berlin/Heidelberg, Germany: Springer.

Qin, Y., Hong, K., Chen, J., & Zhang, Z. (2020). Asymmetric effects of geopolitical risks on energy returns and volatility under different market conditions. *Energy Economics, 90*(C), 104851.

R Core Team. (2023). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Rapach, D. E., & Zhou, G. (2013). Forecasting stock returns. In G. Elliott & A. Timmermann (Eds.), *Handbook of Economic Forecasting* (Vol. 2A, pp. 328–383). Amsterdam: Elsevier.

Reboredo, J. C. (2013a). Is gold a safe haven or a hedge for the US dollar? Implications for risk management. *Journal of Banking and Finance, 37*(8), 2665–2676.

Reboredo, J. C. (2013b). Is gold a hedge or safe haven against oil price movements? *Resources Policy, 38*(2), 130–137.

Salisu, A. A., Cunado, J., & Gupta, R. (2022a). Geopolitical risks and historical exchange rate volatility of the BRICS. *International Review of Economics and Finance, 77*(C), 179–190.

Salisu, A. A., Gupta, R., Bouri, E., & Ji, Q. (2020). The role of global economic conditions in forecasting gold market volatility: Evidence from a GARCH-MIDAS approach. *Research in International Business and Finance, 54*(C), 101308.

Salisu, A. A., Gupta, R., Karmakar, S., & Das, S. (2021b). Forecasting output growth of advanced economies over eight centuries: The role of gold market volatility as a proxy of global uncertainty. *Resources Policy, 75*(C), 102527.

Salisu, A. A., Lukman, L., & Tchankam, J. P. (2022b). Historical geopolitical risk and the behaviour of stock returns in advanced economies. *European Journal of Finance, 28*(9), 889–906.
Salisu, A. A., Pierdzioch, C., & Gupta, R. (2021a). Geopolitical risk and forecastability of tail risk in the oil market: Evidence from over a century of monthly data. *Energy, 235*(C), 121333.
Sharma, S. S. (2016). Can consumer price index predict gold price returns? *Economic Modelling, 55*(C), 269–278.
Smales, L. A. (2021). Geopolitical risk and volatility spillovers in oil and stock markets. *Quarterly Review of Economics and Finance, 80*(C), 358–366.
Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society, Series B, 58*(1), 267–288.
Tibshirani, J., Athey, S., Sverdrup, E., & Wager, S. (2021). grf: Generalized Random Forests. R package version 2.0.2. https://CRAN.R-project.org/package=grf.
Tiwari, A. K., Aye, G. C., Gupta, R., & Gkillas, K. (2020). Gold-oil dependence dynamics and the role of geopolitical risks: Evidence from a Markov-switching time-varying copula model. *Energy Economics, 88*(C), 104748.
Tiwari, A. K., Boachie, M. K., Suleman, M. T., & Gupta, R. (2021). Structure dependence between oil and agricultural commodities returns: The role of geopolitical risks. *Energy, 219*(C), 119584.
Wu, K., & Karmakar, S. (2021). Model-free time-aggregated predictions for econometric datasets. *Forecasting, 3*(4), 920–933.
Wu, K., & Karmakar, S. (2023). A model-free approach to do long-term volatility forecasting and its variants. *Financial Innovation, 9*, 59.
Yang, M., Zhang, Q., Yi, A., & Peng, P. (2021). Geopolitical risk and stock market volatility in emerging economies: Evidence from GARCH-MIDAS model. *Discrete Dynamics in Nature and Society, 2021*(Discrete Dynamics in Economic and Business Systems), 1159358.