



Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.



## Assessing the safe haven property of the gold market during COVID-19 pandemic

Afees A. Salisu<sup>a,b,\*</sup>, Ibrahim D. Raheem<sup>a,b</sup>, Xuan Vinh Vo<sup>b,c</sup>

<sup>a</sup> Centre for Econometric & Allied Research, University of Ibadan, Ibadan, Nigeria

<sup>b</sup> Institute of Business Research, University of Economics Ho Chi Minh City, Viet Nam

<sup>c</sup> Institute of Business Research and CFVG Ho Chi Minh City, University of Economics Ho Chi Minh City, Viet Nam

### ARTICLE INFO

#### JEL classifications:

G10  
G11  
G15  
F10

#### Keywords:

Gold  
Safe haven  
Predictability  
Uncertainty  
COVID-19

### ABSTRACT

This study examines the safe haven prowess of gold against some exogenous shocks due to the COVID-19 pandemic. We further make a comparison of our findings with those obtained for the period before it. Our results confirm the potential of gold market to serve as a safe haven during the pandemic albeit with a higher effectiveness before the pandemic. Further results suggest that gold consistently offers better safe haven properties than the US stocks as well as other precious metals like Silver, Palladium and Platinum regardless of the period. Finally, we find that the predictive model that accounts for uncertainties outperforms the benchmark model that ignores the same both for the in- and out-of-sample forecast analyses.

### 1. Motivation

Gold is described as a multifaceted asset due to its many attributes: currency, commodity, and risk aversion (Wu, Tong, Yang, & Derbalic, 2019). Studies have focused more on the last attribute as gold acts as hedging asset in portfolio diversification and safe haven in period of economic uncertainties and turbulent markets environment. More importantly, gold has been found to retain its values during unsettled market episodes. (Salisu, Raheem, & Ndako, 2020; Salisu & Adediran, 2020; Shahzad, Bouri, Roubaud, & Kristoufek, 2020).

The coronavirus (COVID-19) outbreak that started from China, spiralled into a pandemic, and subject the global economic into a shock the world has never witnessed. In an attempt to combat the diseases, health policy guidelines has advised people to practice “stay at home”, a situation that has led to acute drop in production of output. This has triggered sharp disruptions in the demand and supply dynamics. Some countries have already experiencing recession and there are projections many more will follow suit. Assets and equity prices have plummeted, so as internationally traded commodity. For instance, crude oil price collapsed into negative in the second half of April 2020. There has been high volatility in the financial markets, which as further aggravated the

intensity of uncertainty. Thus, this pandemic has triggered several exogenous shocks.

Judging by the indicators of COVID-19, it is assumed that we have passed the peak and things are gradually going back to normalcy, as countries are beginning to open up. Evidences have suggested that post-crisis periods are associated with high volatilities and uncertainties in financial markets (Antonakakis, Chatziantoniou, & Filis, 2017), and this can spillover to other sectors of the economy (Summer, Johnson, & Soenen, 2010). In a bid to protect their investments, investors will start to readjustment their portfolios in favour of commodities that are considered to have safe haven property, prominent among which is gold.

Several studies have examined the safe haven property of gold (Bouri, Shahzad, Roubaud, Kristoufek, & Lucey, 2020; Shahzad et al., 2020); and other precious metals (Li & Lucey, 2017; Peng, 2019; Sakemoto, 2018). A common feature of these studies is their focus on a specific crisis, such as the financial/stock market crisis (Ciner, Gurdgiev, & Lucey, 2013; Dee, Li, & Zheng, 2013; Gutiérrez, Franco, & Campuzano, 2013; He, He, O'Connor, & Thijssen, 2018; Iqbal, 2017; Shahzad et al., 2020), debt crisis (Agyei-Ampomah, Gounopoulos, & Mazouz, 2014; Boubaker, Cunado, Gil-Alana, & Gupta, 2020; Bredin, Conlon, & Poti, 2015), exchange rate crisis (Joy, 2011; Reboredo, 2013b; Reboredo

\* Corresponding author at: Centre for Econometric & Allied Research, University of Ibadan, Ibadan, Nigeria.

E-mail addresses: [aa.salisu@cear.org.ng](mailto:aa.salisu@cear.org.ng) (A.A. Salisu), [i\\_raheem@ymail.com](mailto:i_raheem@ymail.com) (I.D. Raheem), [vinhvx@ueh.edu.vn](mailto:vinhvx@ueh.edu.vn) (X.V. Vo).

& Rivera-Castro, 2014a and b; Qureshi, Rehman, & Qureshi, 2018; Bedoui, Guesmi, Kalai, & Porcher, 2019). Some studies have also focused on non-crisis uncertainty such as economic uncertainty (Hood & Malik, 2013; Wu et al., 2019), and commodity price shock (Reboredo, 2013a; Salisu & Adediran, 2020).

Whereas, COVID-19 is a global phenomenon, that has subjected virtually all the countries to various forms of shocks (such as decline in output, stock prices, international reserves, to mention a few). It is argued that the severity of this pandemic is no match to any of the previous crises. The empirical inquiry this study seeks answer to is “does the safe haven property of gold extends to the COVID-19 pandemic induced shock?” To answer this question we built a dataset from 1 Jan. 2020 to 20 May 2020. We further make a comparison about the effectiveness of this safe haven property for pre- and post- announcement of the disease. This comparison becomes germane as studies have shown that the effectiveness of hedging prowess of gold is dependent upon the existing market conditions i.e. bear or bull (Iqbal, 2017). Limiting the scope of the study to COVID-19 era will reveal the true extent of gold hedging capacity. The novel contribution of this study is derived from its objective. No study we are aware of has focused on the safe haven of gold during the current COVID-19 crisis.

Our results confirm the potential of gold market to serve as a safe haven during the pandemic albeit with a higher effectiveness before the pandemic. Further results suggest that gold consistently offers better safe haven properties than the US stocks as well as other precious metals like Silver, Palladium and Platinum regardless of the period. Finally, we find that the predictive model that accounts for uncertainties outperforms the benchmark model that ignores the same both for the in- and out-of-sample forecast analyses.

Our results support the hypothesis set out in this study. Specifically, we show that gold serve as a safe haven during the pandemic albeit with a higher effectiveness before the pandemic. In comparison to other precious metals, gold offers better safe haven properties. The rest of the study is structured as follows. Section 2 provides discussions on methodology. Empirical results are presented in Section 3, while the concluding remark is rendered in Section 4.

## 2. Methodology

Here, we formulate an empirical model that allows us to assess the safe haven property of gold. The formulation relies on two assumptions: one, that investment in the gold market is expected to retain or increase in value during times of market turbulence (or high market volatility); two and flowing from one, since the COVID-19 pandemic is global, we anticipate a positive correlation between a measure of uncertainty in a competitive market (in this case, stock market) and returns on investment in the gold market. In other words, as the stock market uncertainty increases, investors are tempted to look elsewhere for safe investments and if truly the gold market possesses a safe haven property, it should be a safe destination for investments during the pandemic and by extension its returns should improve as trading in the market improves. The model is specified in such as to control for endogeneity bias (that may result from omitting other predictors of gold returns), conditional heteroscedasticity effect (due to the use of high frequency data) and persistence (which is typical of most financial and economic time series) (see Westerlund & Narayan, 2012, 2015):

$$r_t = \alpha + \sum_{i=1}^5 \lambda_i^{adj} uc_{t-i} + \gamma(uc_t - \rho_0 uc_{t-1}) + \eta_t \tag{1}$$

where  $r_t$  is the log return of gold price;  $uc$  is the market uncertainty index using the Chicago Board Options Exchange (CBOE), the Volatility Index, or VIX, as a proxy; the  $\eta_t$  is zero mean idiosyncratic error term gold returns; and the coefficient  $\lambda_i^{adj}$  measures the relative impact of market uncertainty on gold returns and we allow for up to five lags given the data frequency (daily five-day of the week) as well as the need to capture

more dynamics in the estimation process. Thus, the underlying null hypothesis of no predictability involves a joint (Wald) test -  $\sum_{i=1}^5 \lambda_i^{adj} = 0$ . Note that the original specification of (1) is given as  $r_t = \alpha + \lambda uc_{t-1} + \varepsilon_t$ , however, to resolve any probable endogeneity bias resulting from the correlation between  $uc_t$  and  $\varepsilon_t$  as well as any potential persistence effect, we follow the approach of Lewellen (2004) and Westerlund and Narayan (2012, 2015). Thus, the parameter  $\lambda^{adj}$  is derived as  $\lambda^{adj} = \lambda - \gamma(1 - \rho_0)$  (where  $\rho_0$  measures the degree of persistence in  $uc_t$ ) and is described as the bias adjusted OLS estimator of Lewellen (2004) which corrects for any persistence effect in the predictive model. The additional term  $\gamma(uc_t - \rho_0 uc_{t-1})$  corrects for any endogeneity bias resulting from the correlation between  $uc_t$  and  $\varepsilon_t$  as well as any inherent unit root problem in the predictor series. Accounting for endogeneity bias here is important since there could be several determinants of gold returns which are suppressed in Eq. (1). Such omissions could introduce endogeneity bias resulting from probable correlations between  $uc_t$  and  $\varepsilon_t$ . To resolve the conditional heteroscedasticity effect, Westerlund and Narayan (2012, 2015) suggest pre-weighting all the data by  $1/\hat{\sigma}_\eta$  and estimating the resulting equation with the Ordinary Least Squares (OLS). This modified OLS estimator is described as the Feasible Quasi GLS estimator in Westerlund and Narayan (2012, 2015) and it is technically computed as:

$$\lambda_{adj}^{FQGLS} = \frac{\sum_{t=q_m+2}^T \tau_t^2 p_{t-1}^d s_t^d}{\sqrt{\sum_{t=q_m+2}^T \tau_t^2 (p_{t-1}^d)^2}} \tag{2}$$

where  $\tau_t = 1/\sigma_{\eta, t}$  is used in weighting all the data in Eq. (2) and  $p_t^d = p_t - \sum_{s=2}^T p_s/T$ .

We further test whether the inclusion of the uncertainty index in the valuation of gold returns will produce better forecast accuracy at least relative to the historical average model which is a typical (baseline) predictive model for most financial and economic series. Since the two models are nested as the historical average is a restricted version of Eq. (1), their forecast performance comparison can easily be implemented using the Clark and West (2007) [CW] test.<sup>2</sup> The CW test is used to determine the statistical significance of the difference between the forecast errors of the two nested (restricted and unrestricted) models, with the underlying procedure defined by:

$$\hat{f}_{t+h} = \left( r_{t+h} - \hat{r}_{1,t,t+h} \right)^2 - \left[ \left( r_{t+h} - \hat{r}_{2,t,t+h} \right)^2 - \left( \hat{r}_{1,t,t+h} - \hat{r}_{2,t,t+h} \right)^2 \right] \tag{3}$$

where  $h$  denotes the forecast period;  $\left( r_{t+h} - \hat{r}_{1,t,t+h} \right)^2$  and  $\left( r_{t+h} - \hat{r}_{2,t,t+h} \right)^2$  are respectively the squared errors for the restricted and the unrestricted models; and  $\left( \hat{r}_{1,t,t+h} - \hat{r}_{2,t,t+h} \right)^2$  is the adjusted squared error incorporated in the CW test to correct for any noise that may characterize the forecasts of larger models. In Eq. (1), the sample average of  $\hat{f}_{t+h}$  can be expressed as  $MSE_1 - (MSE_2 - adj.)$  where  $MSE_1 = P^{-1} \sum \left( r_{t+h} - \hat{r}_{1,t,t+h} \right)^2$ ,  $MSE_2 = P^{-1} \sum \left( r_{t+h} - \hat{r}_{2,t,t+h} \right)^2$ ,  $adj. = P^{-1} \sum \left( \hat{r}_{1,t,t+h} - \hat{r}_{2,t,t+h} \right)^2$ , and  $P$  is the number of forecast periods considered in the computation of the averages. In testing for equality of forecast performance between restricted (the historical average) and unrestricted models, the generated  $\hat{f}_{t+h}$  series is regressed on a constant

<sup>1</sup> See Westerlund and Narayan (2012, 2015) for computational details.

<sup>2</sup> An alternative approach which involves the Diebold & Mariano test is not considered since it is more suitable for non-nested models (see Clark & McCracken, 2001; Diebold & Mariano, 1995; Harvey, Leybourne, & Newbold, 1997).

term only and using the resulting t-statistic for a zero coefficient to determine significance. We reject the hypothesis of a zero coefficient if this statistic is greater than +1.282 (for a one sided 0.10 test) or + 1.645 (for a one sided 0.05 test) (see Clark & West, 2007). The rejection of the null hypothesis implies the preference for the uncertainty-based model for gold returns. We also employ single forecast measures such as the Root Mean Square Error (RMSE), Mean Absolute Deviation (MAD) and Mean Absolute Percentage Error (MAPE) to complement the Clark & West test results.

Since we are confronted with limited data sample between the period COVID-19 was announced (December 31, 2019) and the period of conducting this research (May 28, 2020), we can only use the 75:25 data split for the in-sample and out-of-sample forecast evaluation. Some studies have used 50:50 and 25:75 data splits, however, these cannot be implemented here given the available data scope for our study. We consider two out-of-sample forecast horizons, 10-day and 20-day ahead forecast horizons and the recursive approach to forecasting is adopted.

### 3. Results

#### 3.1. Preliminary analyses

Since this study involves comparative analyses, we dichotomize the sample size into pre- and post-COVID-19 periods. Table 1 presents the results of the descriptive statistics. An overview of the table shows that return on gold is higher in the post-announcement of the COVID-19, as compared to the preceding period. This increase could be explained by the high volatility, as depicted in the standard deviation statistic. This is the first pointer to confirm the hedging property of gold. Expectedly, uncertainty is higher during the turbulent period as compared to the tranquil market condition. There is evidence of conditional heteroscedasticity for VIX, irrespective of the lag-length. There is weak evidence for serial correlation and conditional heteroscedasticity for gold. Figs. 1 and 2 show the trend of gold returns with VIX (pre- and post-announcement, respectively). The figures suggest that both series move in the same direction. While the dispersion between the series is wide in the first half of pre-announcement period, the exact opposite

**Table 1**  
Summary statistics and residual based tests.

|                    | Pre-COVID<br>Announcement |         | Post-COVID<br>Announcement |          |            |
|--------------------|---------------------------|---------|----------------------------|----------|------------|
|                    | Gold<br>return            | VIX     | Gold<br>return             | VIX      |            |
| Mean               | 0.0154                    | 15.0844 | 0.1443                     | 33.1079  |            |
| Std.               | 0.6978                    | 2.6298  | 1.4162                     | 18.5992  |            |
| Skw                | -0.1385                   | 0.7649  | 0.4901                     | 0.6708   |            |
| Kurt               | 5.1160                    | 2.5738  | 8.8629                     | 2.6207   |            |
| Autocorrelation    | k =                       | 0.0056  | 0.8690                     | 1.7790   | 17.601***  |
|                    | 2                         |         |                            |          |            |
|                    | k =                       | 4.9598  | 2.6216                     | 2.8022   | 22.427***  |
|                    | 4                         |         |                            |          |            |
| Heteroscedasticity | k =                       | 5.1649  | 4.4544                     | 14.666** | 24.633***  |
|                    | 2                         |         |                            |          |            |
|                    | k =                       | 0.1592  | 9.3194***                  | 0.2678   | 15.9213*** |
|                    | 2                         |         |                            |          |            |
|                    | k =                       | 0.1590  | 4.3412***                  | 0.2089   | 8.9520***  |
|                    | 2                         |         |                            |          |            |
|                    | k =                       | 0.2033  | 2.0404*                    | 2.6935** | 7.1367***  |
|                    | 4                         |         |                            |          |            |

Note: Std is standard deviation, Skw is skewness, Kurt is Kurtosis. For autocorrelation and heteroscedasticity tests, the reported values are the Ljung-Box test Q-statistics for the former and the ARCH-LM test F-statistics in the case of the latter. We consider three different lag lengths (k) of 2, 4, and 6 for robustness. The null hypothesis for the autocorrelation test is that there is no serial correlation, while the null for the ARCH-LM (F distributed) test is that there is no conditional heteroscedasticity. \*\*\* and \*\* & \* imply the rejection of the null hypothesis in both cases at 1%, 5% & 10% levels of significance, respectively.

post-announcement period. Result of the unit root test is presented in Table 2. Irrespective of the period and test, gold is stationary at levels, thus confirming its mean reverting tendencies. On the flip side, VIX is found to be first difference stationary for the post-announcement period, confirming its high level of volatility. Two structural break dates were found: 2020-03-13 and 2020-04-02. The first date could be attributed to World Health Organisation's declaration of COVID-19 a pandemic. The second date could reflect the effect of recording over 5000 and 250,000 deaths and confirmed cases, respectively. Results for persistence and endogeneity, presented in Table 3, confirm the existence of persistence and absence of endogeneity. The presence of the persistence effect is a necessary condition for the consideration of the bias-adjusted estimator of Lewellen (2004) and Westerlund and Narayan (2012, 2015) used in this study.

#### 3.2. Main results

Table 4 presents the results of the safe haven prowess of gold due to uncertainties associated with the COVID-19 pandemic as well as those before it. The results confirm the potential of gold market as a safe haven during the pandemic. Confirming the position of the literature, the safe haven effectiveness of financial markets is dependent upon the market condition (see Junttila, Pesonen, & Raatikainen, 2018). We show that the effectiveness is higher for pre-COVID-19 announcement as compared to the post announcement period. Similar results are obtained by previous studies such as (Hood & Malik, 2013) who concluded that safe haven property is weak during period of extreme uncertainty as the 2008 financial crisis, a shock likened to the current pandemic.

The forecasting evaluation result is presented in subsequent tables (Tables 5, 6 & 7). It should be recalled that the data is divided into: in- and out-of-sample sizes using 75% and 25%, respectively. The results of the in-sample are presented in Table 5, which confirm that the uncertainty induced model is our preferred model relative to the historical average model, and the outcome remains the same for both pre- and post-announcement periods particularly judging by the RMSE and MAD where the preferred model reports the least value likewise the Clark and West test shows statistical significance at the 1% level. The out-of-sample forecast evaluation of both the uncertainty-based model and the historical average also produces the same conclusion as the in-sample forecasts irrespective of the period [whether pre-announcement or post-announcement]. Thus, we may conclude that accounting for uncertainties when modelling gold returns will offer better forecast estimates than those that ignore the same.

#### 3.3. Additional results

##### 3.3.1. Safe haven behaviour of stocks during the pandemic

Some robustness checks were conducted. First, we test further whether the US stocks will share the similar safe haven potential with gold by examining how the US stocks respond to uncertainty particularly during turbulent times like the current pandemic. Thus, the Eq. (1) for gold still suffices except that the dependent variable is now replaced with US stock return using S&P 500 index. The results of the predictability analyses are presented in Table 8. Unlike gold which possesses safe haven properties both for the pre- and post-announcement periods of COVID-19, the ability of US stocks to provide safe investment is noticed during the pandemic albeit with lower magnitude compared to gold. These results replicate those of existing studies (such as He et al., 2018; and Shahzad et al., 2020). The results of the out-of-sample analyses are presented in Table 9, which confirm the outperformance of the uncertainty-based model over the historical average. In other words, regardless of the choice of market, whether gold or stocks, the consideration of uncertainties is crucial in the valuation of assets. This conclusion is in line with the standard theories of asset pricing such as Capital Asset Pricing Model and Arbitrage Pricing Theory both of which suggest the need to account for one form of risk or the other in the

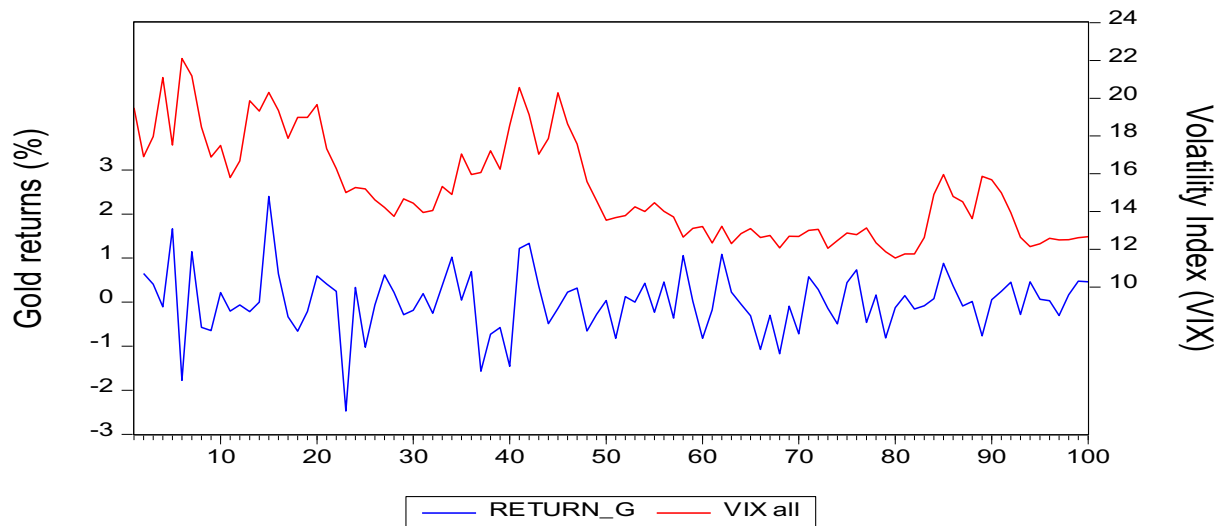


Fig. 1. Trends in gold returns and VIX index [Pre-COVID Announcement]. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

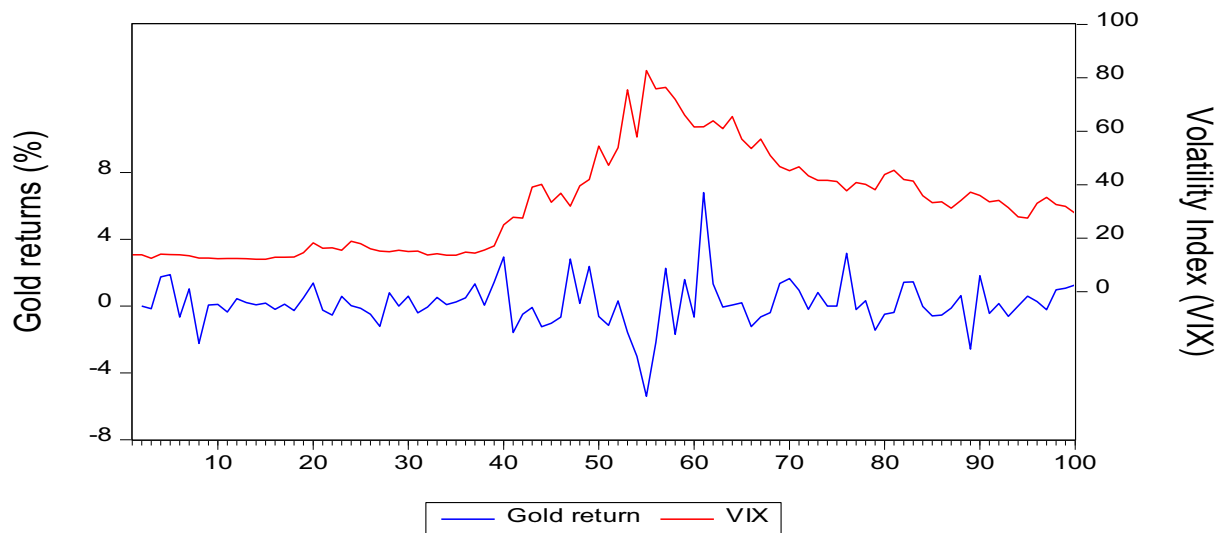


Fig. 2. Trends in gold returns and VIX index [Post-COVID Announcement]. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 2

Results of Unit root tests.

|                         | ADF                  |                     | NL                  |                      | PP                   |                     | KPSS               |                    | ERS                |                    | NP                 |                    | Break Dates          |
|-------------------------|----------------------|---------------------|---------------------|----------------------|----------------------|---------------------|--------------------|--------------------|--------------------|--------------------|--------------------|--------------------|----------------------|
|                         | Level                | FD                  | Level               | FD                   | Level                | FD                  | Level              | FD                 | Level              | FD                 | Level              | FD                 |                      |
| Pre-COVID Announcement  |                      |                     |                     |                      |                      |                     |                    |                    |                    |                    |                    |                    |                      |
| Gold return             | -10.595 <sup>a</sup> | -                   | -9.805 <sup>a</sup> | -                    | -11.810 <sup>a</sup> | -                   | 0.140 <sup>a</sup> | -                  | 2.053 <sup>a</sup> | -                  | 0.099 <sup>a</sup> | -                  | -                    |
| VIX                     | -3.5360 <sup>b</sup> | -                   | -3.183 <sup>b</sup> | -                    | -3.478 <sup>b</sup>  | -                   | 0.099 <sup>a</sup> | -                  | 5.082              | 3.040 <sup>a</sup> | 0.152              | 0.132 <sup>b</sup> | 09/04/19             |
| Post-COVID Announcement |                      |                     |                     |                      |                      |                     |                    |                    |                    |                    |                    |                    |                      |
| Gold return             | -8.690 <sup>a</sup>  | -                   | -12.431             | -                    | -9.129 <sup>a</sup>  | -                   | 0.032 <sup>a</sup> | -                  | 1.770 <sup>a</sup> | -                  | 0.098 <sup>a</sup> | -                  | -                    |
| VIX                     | -1.323               | -4.237 <sup>a</sup> | -2.0161             | -10.687 <sup>a</sup> | -1.023               | -14.12 <sup>a</sup> | 0.219              | 0.119 <sup>a</sup> | 15.725             | 2.065 <sup>a</sup> | 0.270              | 0.118 <sup>b</sup> | 13/03/20<br>02/04/20 |

Note: ADF test is the Augmented Dickey Fuller test; NL test is the Narayan and Liu (2015) test; PP is the Philips Perron test; KPSS is the Kwiatkowski, Phillips, Schmidt, and Shin (1992) test; ERS is the Elliot, Rothenberg, and Stock (1996) test and NP is the Ng and Perron (2001) test. FD denotes First Difference; “<sup>a</sup>” and “<sup>b</sup>” indicate the rejection of the null hypothesis of a unit root at 1%, 5% and 10%, respectively; the test regression for all the unit root tests includes intercept and trend; I(d) implies the order of integration, where d is the number of differencing required for a series to become stationary; All the variables are in their log forms. The breaks are determined using the Bai and Perron (2003) test.



**Table 3**  
Persistence and endogeneity test results for VIX.

|     | Persistence |            | Endogeneity |            |
|-----|-------------|------------|-------------|------------|
|     | Pre-COVID   | Post-COVID | Pre-COVID   | Post-COVID |
| VIX | 0.8741***   | 0.9572***  | -0.0679     | -0.0431    |

Note: \*\*\*, \*\*, & \* imply statistical significance of coefficients at 1%, 5%, and 10% levels, respectively.

**Table 4**  
Predictability results.

|                                | Pre-COVID Announcement           | Post-COVID Announcement           |
|--------------------------------|----------------------------------|-----------------------------------|
| $\sum_{i=1}^5 \lambda_i^{adj}$ | 0.0138**<br>(0.0057)<br>[2.4039] | 0.0063***<br>(0.0014)<br>[4.2257] |

Note: \*\*\* and \*\* & \* imply the rejection of the null hypothesis of no predictability at 1%, 5% & 10% levels of significance, respectively. Values in parentheses - () denote standard errors while those reported in square brackets - [] are for t-statistics.

**Table 5**  
In-Sample Forecast evaluation.

|              | Pre-COVID Announcement |                                   | Post-COVID Announcement |                                   |
|--------------|------------------------|-----------------------------------|-------------------------|-----------------------------------|
|              | Model 1                | Model 2                           | Model 1                 | Model 2                           |
| RMSE         | 0.3460                 | 0.7565                            | 1.0392                  | 1.4982                            |
| MAD          | 0.5000                 | 0.5570                            | 0.9193                  | 0.9708                            |
| MAPE         | 801.252                | 2160.096                          | 1255.636                | 758.942                           |
| Clark & West | -                      | 0.2099***<br>(0.0710)<br>[2.9535] | -                       | 1.2478***<br>(0.5252)<br>[2.3760] |

Note: Model 1 incorporates the Uncertainty (VIX) predictor while Model 2 is the Historical Average model. Thus, the former is the unrestricted model while the latter is the restricted model. RMSE is Root Mean Square Error; MAD is Mean Absolute Deviation & MAPE is Mean Absolute Percentage Error. The results for the Clark & West are reported for the model under the null (i.e. Model 2). The RMSE reported for Model 1 is the version of Clark and West (2007) which adjusts the difference in mean squared prediction errors to account for the additional predictors in the model. \*\*\* and \*\* & \* imply the rejection of the null hypothesis of equal forecast accuracy at 1%, 5% & 10% levels of significance, respectively. The null hypothesis of a zero coefficient is rejected if this statistic is greater than +1.282 (for a one sided 0.10 test), +1.645 (for a one sided 0.05 test) and + 2.00 for 0.01 test (for a one sided 0.01 test) (see Clark & West, 2007). Values in parentheses - () denote standard errors while those reported in square brackets - [] are for t-statistics.

valuation of assets and since the pandemic is associated with global uncertainty, it can serve as one form of systemic risk to financial market.

**3.3.2. The safe haven properties of other precious metals**

Further analyses are carried out for other precious metals such as Silver, Palladium and Platinum. This consideration is important as most precious metals including gold tend to share some inherent characteristics such as: (i) global acceptance; (ii) no possibility of default risks; (iii) store of value tendencies; and their protection attributes and properties (see Arnold & Auer, 2015), make them (precious metals) have heterogeneous hedging tendencies (Uddin, Shahzad, Boako, Hernandez, & Lucey, 2019). An attempt to isolate gold will only offer biased or partial outcomes if the results are not compared with those from other precious metals. This is the motivation for these additional analyses. The results of the in-sample predictability are rendered in Table 10 while those of out-of-sample forecast evaluation are presented in Table 11. We can infer two remarks from Table 10: (i) precious metals have different hedging prowess, with silver being the least hedging option; (ii) in comparison with statistics in Table 4, gold offer the best financial assets that could shield investors' portfolio from any exogenous shocks. This

**Table 6**  
Out-of-Sample Forecast evaluation [Pre-COVID Announcement].

|              | Model 1 |                                   | Model 2  |                                   |
|--------------|---------|-----------------------------------|----------|-----------------------------------|
|              | h = 10  | h = 20                            | h = 10   | h = 20                            |
| RMSE         | 0.2954  | 0.2758                            | 0.7253   | 0.6966                            |
| MAD          | 0.4685  | 0.4532                            | 0.5247   | 0.5031                            |
| MAPE         | 728.076 | 678.195                           | 2034.644 | 1950.854                          |
| Clark & West | -       | 0.2140***<br>(0.0646)<br>[3.3118] | -        | 0.1928***<br>(0.0583)<br>[3.3046] |

Note: Model 1 incorporates the Uncertainty (VIX) predictor while Model 2 is the Historical Average. Thus, the former is the unrestricted model while the latter is the restricted model. RMSE is Root Mean Square Error; MAD is Mean Absolute Deviation & MAPE is Mean Absolute Percentage Error. The results for the Clark & West are reported for the model under the null (i.e. Model 2). The RMSE reported for Model 1 is the version of Clark and West (2007) which adjusts the difference in mean squared prediction errors to account for the additional predictors in the model. \*\*\* and \*\* & \* imply the rejection of the null hypothesis of equal forecast accuracy at 1%, 5% & 10% levels of significance, respectively. The null hypothesis of a zero coefficient is rejected if this statistic is greater than +1.282 (for a one sided 0.10 test), +1.645 (for a one sided 0.05 test) and + 2.00 for 0.01 test (for a one sided 0.01 test) (see Clark & West, 2007). Two out-of-Sample forecast horizons are considered: 10-day & 20-day ahead forecast horizons. Values in parentheses - () denote standard errors while those reported in square brackets - [] are for t-statistics.

**Table 7**  
Out-of-Sample Forecast evaluation [Post-COVID Announcement].

|              | Model 1  |                                   | Model 2 |                                   |
|--------------|----------|-----------------------------------|---------|-----------------------------------|
|              | h = 10   | h = 20                            | h = 10  | h = 20                            |
| RMSE         | 1.2006   | 1.1299                            | 1.4706  | 1.4092                            |
| MAD          | 0.9928   | 0.9720                            | 0.9717  | 0.9273                            |
| MAPE         | 1186.005 | 1073.689                          | 759.618 | 724.923                           |
| Clark & West | -        | 0.9954***<br>(0.4697)<br>[2.1193] | -       | 0.8782***<br>(0.4227)<br>[2.0779] |

Note: Model 1 incorporates the Uncertainty (VIX) predictor while Model 2 is the Historical Average. Thus, the former is the unrestricted model while the latter is the restricted model. RMSE is Root Mean Square Error; MAD is Mean Absolute Deviation & MAPE is Mean Absolute Percentage Error. The results for the Clark & West are reported for the model under the null (i.e. Model 2). The RMSE reported for Model 1 is the version of Clark and West (2007) which adjusts the difference in mean squared prediction errors to account for the additional predictors in the model. \*\*\* and \*\* & \* imply the rejection of the null hypothesis of equal forecast accuracy at 1%, 5% & 10% levels of significance, respectively. The null hypothesis of a zero coefficient is rejected if this statistic is greater than +1.282 (for a one sided 0.10 test), +1.645 (for a one sided 0.05 test) and + 2.00 for 0.01 test (for a one sided 0.01 test) (see Clark & West, 2007). Two out-of-Sample forecast horizons are considered: 10-day & 20-day ahead forecast horizons. Values in parentheses - () denote standard errors while those reported in square brackets - [] are for t-statistics.

**Table 8**  
Predictability results.

|                                | Pre-COVID Announcement              | Post-COVID Announcement          |
|--------------------------------|-------------------------------------|----------------------------------|
| $\sum_{i=1}^5 \lambda_i^{adj}$ | -0.0278***<br>(0.0030)<br>[-9.2979] | 0.0039**<br>(0.0015)<br>[2.5397] |

Note: \*\*\* and \*\* & \* imply the rejection of the null hypothesis of no predictability at 1%, 5% & 10% levels of significance, respectively. Values in parentheses - () denote standard errors while those reported in square brackets - [] are for t-statistics.

**Table 9**  
Out-of-Sample Forecast evaluation.

|              | Pre-COVID Announcement            |                                   | Post-COVID Announcement            |                                    |
|--------------|-----------------------------------|-----------------------------------|------------------------------------|------------------------------------|
|              | Model 1 vs Model 2                |                                   | Model 1 vs Model 2                 |                                    |
|              | h = 10                            | h = 20                            | h = 10                             | h = 20                             |
| Clark & West | 0.9911***<br>(0.2410)<br>[4.1132] | 0.9052***<br>(0.2173)<br>[4.1657] | 15.9457***<br>(4.8775)<br>[3.2692] | 14.7262***<br>(4.3784)<br>[3.3633] |

Note: Model 1 incorporates the Uncertainty (VIX) predictor while Model 2 is the Historical Average. Thus, the former is the unrestricted model while the latter is the restricted model. \*\*\* and \*\* & \* imply the rejection of the null hypothesis of equal forecast accuracy at 1%, 5% & 10% levels of significance, respectively. The null hypothesis of a zero coefficient is rejected if this statistic is greater than +1.282 (for a one sided 0.10 test), +1.645 (for a one sided 0.05 test) and + 2.00 for 0.01 test (for a one sided 0.01 test) (see Clark & West, 2007). Two out-of-Sample forecast horizons are considered: 10-day & 20-day ahead forecast horizons. Values in parentheses – () denote standard errors while those reported in square brackets – [] are for t-statistics.

**Table 10**  
Predictability results.

|                         |                                | Palladium                          | Platinum                            | Silver                             |
|-------------------------|--------------------------------|------------------------------------|-------------------------------------|------------------------------------|
| Pre-COVID Announcement  | $\sum_{i=1}^5 \lambda_i^{adj}$ | 0.0948***<br>(0.0065)<br>[14.4949] | -0.2373***<br>(0.0672)<br>[-3.5318] | 0.0068<br>(0.007768)<br>[0.872976] |
| Post-COVID Announcement | $\sum_{i=1}^5 \lambda_i^{adj}$ | 0.0083<br>(0.0085)<br>[0.9706]     | 0.05724***<br>(0.0041)<br>[14.0426] | -0.0031*<br>(0.0018)<br>[-1.6886]  |

Note: \*\*\* and \*\* & \* imply the rejection of the null hypothesis of no predictability at 1%, 5% & 10% levels of significance, respectively. Values in parentheses - () denote standard errors while those reported in square brackets – [] are for t-statistics.

**Table 11**  
Out-of-Sample Forecast evaluation.

|              |           | Pre-COVID Announcement            |                                    | Post-COVID Announcement           |                                   |
|--------------|-----------|-----------------------------------|------------------------------------|-----------------------------------|-----------------------------------|
|              |           | Model 1 vs Model 2                |                                    | Model 1 vs Model 2                |                                   |
|              |           | h = 10                            | h = 20                             | h = 10                            | h = 20                            |
| Clark & West | Palladium | 0.1309<br>(0.1059)<br>[1.2357]    | 0.1206<br>(0.0953)<br>[1.2660]     | 2.1448*<br>(1.5472)<br>[1.3862]   | 1.8567*<br>(1.4013)<br>[1.3250]   |
|              | Platinum  | 48.2419*<br>(30.8232)<br>[1.5651] | 44.7887*<br>(27.5875)<br>[1.62352] | 55.9379*<br>(36.5221)<br>[1.5316] | 52.2739*<br>(32.6786)<br>[1.5996] |
|              | Silver    | 0.6235***<br>(0.1811)<br>[3.4428] | 0.5313***<br>(0.1645)<br>[3.2301]  | 3.4376***<br>(1.7101)<br>[2.0102] | 3.3240***<br>(1.5398)<br>[2.1587] |

Note: Model 1 incorporates the Uncertainty (VIX) predictor while Model 2 is the Historical Average. Thus, the former is the unrestricted model while the latter is the restricted model. \*\*\* and \*\* & \* imply the rejection of the null hypothesis of equal forecast accuracy at 1%, 5% & 10% levels of significance, respectively. The null hypothesis of a zero coefficient is rejected if this statistic is greater than +1.282 (for a one sided 0.10 test), +1.645 (for a one sided 0.05 test) and + 2.00 for 0.01 test (for a one sided 0.01 test) (see Clark & West, 2007). Two out-of-Sample forecast horizons are considered: 10-day & 20-day ahead forecast horizons. Values in parentheses – () denote standard errors while those reported in square brackets – [] are for t-statistics.

confirms position that gold holds a “special position”. Summarizing Table 11, our model, as against the historical average-based model, is more preferred in accurately forecasting the returns of other precious metals except for Palladium during the pre-COVID announcement.

## 4. Conclusion

This study aims to examine the safe haven properties of gold on some selected exogenous shocks, such as uncertainty due to the novel virus – COVID-19. For completeness, we further consider the following: (i) we carry out distinct analyses for two sub-periods of pre- and post-COVID announcement periods; (ii) we compare the safe haven potential of gold during the pandemic with that of US stocks and other precious metals; (iii) we evaluate the forecast performance of including uncertainty in the valuation of financial assets. Our results confirm the ability of gold market to safe investment options during the pandemic than other financial assets such as those considered in this study. The results of this exercise further enhance the dexterity of gold as the most useful hedging financial asset that protects investors’ portfolios. In other words, investors are better off shielding their investments by diversifying their portfolio to include the acquisition of gold. Finally, for financial analysts and policy makers who constantly confronted with the need to provide accurate forecasts for investment and policy decisions, this study further lends support to the inclusion of uncertainty in the valuation of stocks as well as risk-adjusted returns in order to produce better forecast outcomes. Future studies could extend the frontier of knowledge by conducting similar exercise for other countries, beyond the US market.

## CRedit authorship contribution statement

**Afees A. Salisu:** Conceptualization; Formal analysis; Investigation; Methodology; Project administration; Software; Validation; Visualization; Writing - original draft; Writing - review & editing. **Ibrahim D. Raheem:** Conceptualization; Investigation; Project administration; Validation; Visualization; Writing - original draft; Writing - review & editing. **Xuan Vinh Vo:** Conceptualization; Funding acquisition; Resources; Visualization; Writing - original draft; Writing - review & editing.

## References

- Agyei-Ampomah, S., Gounopoulos, D., & Mazouz, K. (2014). Does gold offer a better protection against losses in sovereign debt bonds than other metals? *Journal of Banking & Finance*, 40, 507–521.
- Antonakakis, N., Chatziantoniou, I., & Filis, G. (2017). Oil shocks and stock markets: Dynamic connectedness under the prism of recent geopolitical and economic unrest. *International Review of Financial Analysis*, 50, 1–26.
- Arnold, S., & Auer, B. R. (2015). What do scientists know about inflation hedging. *North American Journal of Economics and Finance*, 34, 187–214.
- Bai, J., & Perron, P. (2003). Computation and analysis of multiple structural change models. *Journal of Applied Econometrics*, 18(1), 1–22.
- Bedoui, R., Guesmi, K., Kalai, S., & Porcher, T. (2019). Diamonds versus precious metals: What gleams most against USD exchange rates? *Finance Research Letters*, 34, 101253.
- 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*, 540. <https://doi.org/10.1016/j.physa.2019.123093>
- Bouri, E., Shahzad, S. J. H., Roubaud, D., Kristoufek, L., & Lucey, B. (2020). Bitcoin, gold, and commodities as safe havens for stocks: New insight through wavelet analysis. *The Quarterly Review of Economics and Finance*. <https://doi.org/10.1016/j.qref.2020.03.004>
- Bredin, D., Conlon, T., & Poti, V. (2015). Does gold glitter in the long-run? Gold as a hedge and safe haven across time and investment horizon. *International Review of Financial Analysis*, 41, 320–328.
- Ciner, C., Gurdgiev, C., & Lucey, B. M. (2013). Hedges and safe havens: An examination of stocks, bonds, gold, oil and exchange rates. *International Review of Financial Analysis*, 29, 202–211.
- Clark, T. E., & McCracken, M. W. (2001). Tests of equal forecast accuracy and encompassing for nested models. *Journal of Econometrics*, 105(1), 85–110. Elsevier.
- Clark, T. E., & West, K. D. (2007). Approximately normal tests for equal predictive accuracy in nested models. *Journal of Econometrics*, 138, 291–311.
- Dee, J., Li, L., & Zheng, Z. (2013). Is gold a hedge or a safe haven? Evidence from inflation and stock market. *International Journal of Development and Sustainability*, 2, 1–16.
- Diebold, F. X., & Mariano, R. S. (1995). Comparing predictive accuracy. *Journal of Business & Economic Statistics*, 13, 253–263.
- Elliot, G., Rothenberg, T. J., & Stock, J. H. (1996). Efficient tests for an autoregressive unit root. *Econometrica*, 64, 813–836.
- Gutiérrez, M., Franco, G., & Campuzano, C. (2013). Gold prices: Analyzing its cyclical behavior. *Lecturas de Economía*, 79, 113–142.

- Harvey, D., Leybourne, S., & Newbold, P. (1997). Testing the equality of prediction mean squared errors. *International Journal of Forecasting*, 13, 281–291.
- He, Z., He, Z., O'Connor, F., & Thijssen, J. (2018). Is gold a sometime safe haven or an always hedge for equity investors? A Markov-switching CAPM approach for US and UK stock indices. *International Review of Financial Analysis*, 60, 30–37.
- Hood, M., & Malik, F. (2013). Is gold the best hedge and a safe haven under changing stock market volatility? *Review of Financial Economics*, 22, 47–52.
- Iqbal, J. (2017). Does gold hedge stock market, inflation and exchange rate risks? An econometric investigation. *International Review of Economics and Finance*, 48, 1–17.
- Joy, M. (2011). Gold and the US dollar: Hedge or haven? *Finance Research Letters*, 8, 120–131.
- Junttila, J., Pesonen, J., & Raatikainen, J. (2018). Commodity market based hedging against stock market risk in times of financial crisis: The case of crude oil and gold. *Journal of International Financial Markets, Institutions & Money*, 56, 225–280.
- Kwiatkowski, D., Phillips, P. C. D., Schmidt, P., & Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root? *Journal of Econometrics*, 54, 159–178.
- Lewellen, J. (2004). Predicting returns with financial ratios. *Journal of Financial Economics*, 74, 209–235.
- Li, S., & Lucey, B. M. (2017). Reassessing the role of precious metals as safe havens—what colour is your haven and why? *Journal of Commodity Markets*, 7, 1–14.
- Narayan, P. K., & Liu, R. (2015). A unit root model for trending time-series energy variables. *Energy Economics*, 50, 391–402.
- Ng, S., & Perron, P. (2001). Lag length selection and the construction of unit root tests with good size and power. *Econometrica*, 69, 519–1554.
- Peng, X. (2019). Do precious metals act as hedges or safe havens for China's financial markets? *Finance Research Letters*, 101353. <https://doi.org/10.1016/j.frl.2019.101353>
- Qureshi, S., Rehman, I. U., & Qureshi, F. (2018). Does gold act as a safe haven against exchange rate fluctuations? The case of Pakistan rupee. *Journal of Policy Modeling*, 40 (4), 685–708.
- Reboredo, J. C. (2013a). Is gold a hedge or safe haven against oil price movements? *Resources Policy*, 38, 130–137.
- Reboredo, J. C. (2013b). Is gold a safe haven or a hedge for the US dollar? Implications for risk management. *Journal of Banking & Finance*, 37(8), 2665–2676.
- Reboredo, J. C., & Rivera-Castro, M. A. (2014a). Can gold hedge and preserve value when the US dollar depreciates? *Economic Modelling*, 39, 168–173.
- Sakemoto, R. (2018). Do precious and industrial metals act as hedges and safe havens for currency portfolios? *Finance Research Letters*, 24, 256–262.
- Salisu, A. A., & Adediran, I. (2020). Gold as a hedge against oil shocks: Evidence from new datasets for oil shocks. *Resources Policy*, 66, 101606. <https://doi.org/10.1016/j.resourpol.2020.101606>
- Salisu, A. A., Raheem, I. D., & Ndako, U. B. (2020). The inflation hedging properties of gold, stocks and real estate: A comparative analysis. *Resources Policy*, 66, 101605. <https://doi.org/10.1016/j.resourpol.2020.101605>
- Shahzad, S. J. H., Bouri, E., Roubaud, D., & Kristoufek, L. (2020). Safe haven, hedge and diversification for G7 stock markets: Gold versus bitcoin. *Economic Modelling*, 87(1), 212–224.
- Summer, S., Johnson, R., & Soenen, L. (2010). Spillover effects among gold, stocks, and bonds. *Journal of CENTRUM Cathedra*, 3(2), 106–120.
- Uddin, G. S., Shahzad, A. J. H., Boako, G., Hernandez, J., & Lucey, B. M. (2019). Heterogeneous interconnections between precious metals: Evidence from asymmetric and frequency-domain spillover analysis. *Resources Policy*, 64, 101509. <https://doi.org/10.1016/j.resourpol.2019.101509>
- Westerlund, J., & Narayan, P. K. (2012). Does the choice of estimator matter when forecasting returns? *Journal of Banking and Finance*, 36, 2632–2640.
- Westerlund, J., & Narayan, P. K. (2015). Testing for predictability in conditionally heteroscedasticity stock returns. *Journal of Financial Econometrics*, 13, 342–375.
- Wu, S., Tong, M., Yang, Z., & Derbalic, A. (2019). Does gold or bitcoin hedge economic policy uncertainty? *Finance Research Letters*, 31, 171–178.