WESTMINSTER BUSINESS SCHOOL

OF FINANCE & ACCOUNTING

Predictive Analysis for Decision-Making

Module Code: 7FNCE044W ,2024/2025

Empirical Report, Coursework 2

"Global Uncertainty and Gold Returns: A Time Series Analysis Using Machine Learning (1999–2024)"

Words:3100 words **(Excluding table of contents, figures, references, appendices)**

Github_Link

TABLE OF CONTENTS

TABLE OF FIGURES

“Global Uncertainty and Gold Returns: A Time Series Analysis Using Machine Learning (1999–2024)”

1. ABSTRACT

The paper examines the ability of world uncertainty indices to forecast monthly variations in gold prices during 1999-2024. It employs Ordinary Least Squares (OLS) and high-level Machine Learning algorithms such as Random Forest, SVR, and Gradient Boosting to examine whether the Global Economic Policy Uncertainty (GEPU) and World Uncertainty Index (WUI) reflect changes in gold returns. The research discovers that OLS models reveal very few linear relationships, but non-linear models perform better for forecasting. The research encompasses significant events such as the 2008 crisis and the COVID-19 shock. High tests of accuracy and reliability support the findings.

2. INTRODUCTION

Gold has long been used to preserve value and hedge against bad times. With so many political and economic issues over the past few years, policymakers and investors need to know what influences gold prices. This study will determine whether changes in gold prices can be forecasted using global uncertainty indicators. The significance of this study lies in bridging the measurement of economic uncertainty with the behaviour of financial markets and seeing the reaction of the gold market to global issues.

This study explores the possibility of uncertainty indices, namely the GEPU and WUI, in forecasting monthly gold returns from 1999 to 2024. While the role of gold as a hedge or haven has been well-researched in the existing literature (Baur & Lucey, 2010; Reboredo, 2013), much of this research tends to be based on conventional econometric techniques. We aim to build on and extend this literature by using a machine-learning predictive model supported by a baseline OLS model, focusing closely on robustness and performance diagnostics.

The primary research questions are:

- Can global uncertainty indices significantly predict gold price returns?
- How do linear and non-linear models differ in capturing this predictive relationship?
- Does model performance improve during heightened uncertainty, such as the 2008 financial crisis or the COVID-19 pandemic?
- Can machine learning models outperform traditional OLS in forecasting gold price returns using uncertainty-based features?

To address these, we employ an econometric modelling framework grounded in financial theory, complemented by robustness checks and predictive evaluations using support vector regression (SVR), random forest, and gradient boosting. This hybrid approach allows for both interpretability and enhanced performance testing. The study ensures broad coverage and time-robust insights using monthly data over multiple economic cycles.

Ultimately, this work aims to contribute to the literature on uncertainty transmission in financial markets, offer actionable insights to investors and risk managers, and explore the evolving role of machine learning in empirical macro-finance.

3. LITERATURE REVIEW

Gold has long been seen as a macroeconomic hedge and a haven in times of increased macroeconomic uncertainty. An extensive literature on the complex relationship between uncertainty and asset prices has long shown gold reacting favourably to adverse shocks in the global economic and geopolitical environment (Capie et al., 2005; Baur & Lucey, 2010). The theoretical basis for this derives from portfolio choice theory and behavioural finance. It is predicated on the expectation that when there is greater volatility or a "flight to quality" in the environment, investors will be attracted to what they perceive to be safe repositories of value, including gold (Bodie et al., 2009).

Uncertainty Measures and Financial Markets

The Global Economic Policy Uncertainty Index (GEPU), developed by Baker et al. in 2016, captures the uncertainty in policy matters of several developed economies. Although well-documented in the financial markets, its impact is relatively understudied for commodity assets like gold, especially in the equity and bond markets. Similarly, the World Uncertainty Index (WUI), developed by Ahir et al. in 2018, is a more general, cross-country measure of uncertainty based on economic reports prepared at the country level.

Several empirical research works have revealed significant causal and predictive relationships between gold prices and uncertainty indices. For example, Arouri et al. (2012) illustrated that geopolitical uncertainties significantly increase the price and the volatility of gold. Likewise, Antonakakis et al. (2017) found that political uncertainty and economic policy positively affect gold returns across different time horizons. However, the predictive ability of these indices on gold returns is still not universally agreed upon since most of the available literature relies on linear specifications like vector autoregressions or OLS specifications. The newest developments in machine learning have created a new avenue for flexibility in modelling to thrive in the complex nonlinearities present in financial times series (Gu et al., 2020).

This study assists us in learning the extent to which traditional methods in OLS and algorithms in machine learning can reflect the accuracy with which uncertainty indices forecast gold returns from 1999 through 2024. In contrast to most studies conducted on isolated events or just short periods, we use numerous global uncertainty times to examine the change in the relationship through traditional and contemporary techniques. GEPU and WUI should assist gold returns since gold is a haven for funds. However, the extent they assist may vary with time and circumstance. We verify the information's reliability by observing historical data and splitting time into pre- and post-crisis.

Previous research has determined varying outcomes regarding how robust and enduring such relationships are. Baur and Lucey (2010) state that gold protects against a declining dollar value and world issues. However, Beckmann et al. (2015) think that gold's reaction is unique and complicated, and therefore we require improved models.

Over the past several years, researchers have begun employing machine learning (ML) techniques to identify sophisticated patterns in financial information over time. For instance, Gu et al. (2020) demonstrate that tree-based ML models perform better than conventional regression techniques when forecasting the behaviour of assets under economic uncertainty. Likewise, Lahmiri and Bekiros (2020) observe that ML techniques, particularly ensemble learning models, provide more precise predictions for commodity prices outside the initial data set.

We formulate the following hypotheses.

H1: GEPU has a strong positive association with gold returns.

H2: WUI also has a strong relationship with gold returns, but its quarterly configuration may alter the timing.
H3: Machine learning models predict better than OLS since they can handle the nonlinear impacts of uncertainty on gold.

4. DATA DESCRIPTION AND EXPLORATION

The dataset consists of monthly observations from January 1999 to December 2024. Data was sourced from reputable repositories via the Federal Reserve Economic Data (FRED) and processed using Python.

Variable	Description	Source	Frequency
Gold Price (USD/oz)	Spot price of gold	Yahoo Finance (GC=F)	Monthly (month-end close)
GEPU Index	Global Economic Policy Uncertainty Index	FRED (RECURRENT)	Monthly
WUI Index	World Uncertainty Index	FRED (WUIGLOBALSMPAVG)	Monthly (via interpolation)

The dataset contains 312 monthly observations after removing missing values and aligning data indices. The chosen period captures major macroeconomic shocks, including:

- The Dot-com bubble (2000–2001)
- 9/11 terrorist attacks (2001)
- Global Financial Crisis (2008–09)
- European sovereign debt crisis (2011)
- COVID-19 pandemic (2020–21)
- Russo-Ukrainian war and global inflation surge (2022–24)

This makes it ideal for investigating how global uncertainty events influence gold returns.

Descriptive Statistics

Gold Price: Monthly prices of COMEX futures were converted to daily prices using the final closing price of each month.
GEPU and WUI: Because GEPU was monthly from the beginning, it was merged immediately. WUI, usually quarterly, was made monthly so that everything would match.
Gold Return: Log returns calculated using the formula: $rt = log(\frac{Pt}{Pt-1})$

These patterns from Figure 1 support the suitability of the dataset for modelling uncertainty-return relationships. The inclusion of various crises ensures heterogeneity in market conditions. The gold price fluctuates a lot, particularly at major events. GEPU and WUI both exhibit consistent increases that occur with world events. The returns of gold have low average growth but are extremely unstable.

```
 Dataset Summary
<class 'pandas.core.frame.DataFrame'>
DatetimeIndex: 290 entries, 2000-09-30 to 2024-10-31
Data columns (total 4 columns):
 #   Column           Non-Null Count  Dtype  
---  ------           --------------  -----  
 0   GoldPriceUSD     290 non-null    float64
 1   GEPUCURRENT      290 non-null    float64
 2   WUIGLOBALSMPAVG  290 non-null    float64
 3   GoldReturn       290 non-null    float64
dtypes: float64(4)
memory usage: 11.3 KB
None

 Descriptive Statistics
                 count          mean          std           min           25%  \
GoldPriceUSD     290.0   1151.506552   570.759157    257.899994    637.025009   
GEPUCURRENT      290.0    151.922926    75.596392     49.224732     94.764577   
WUIGLOBALSMPAVG  290.0  19026.840069  5386.797722  10474.980000  14169.280000   
GoldReturn       290.0      0.007884     0.046539     -0.198512     -0.021048   

                          50%           75%           max  
GoldPriceUSD      1228.850037   1591.950043   2738.300049  
GEPUCURRENT        131.492005    202.693867    431.571593  
WUIGLOBALSMPAVG  18328.135000  22886.700000  33863.660000  
GoldReturn          0.007193      0.036142      0.129863  
```

Figure 1 : Descriptive Statistics

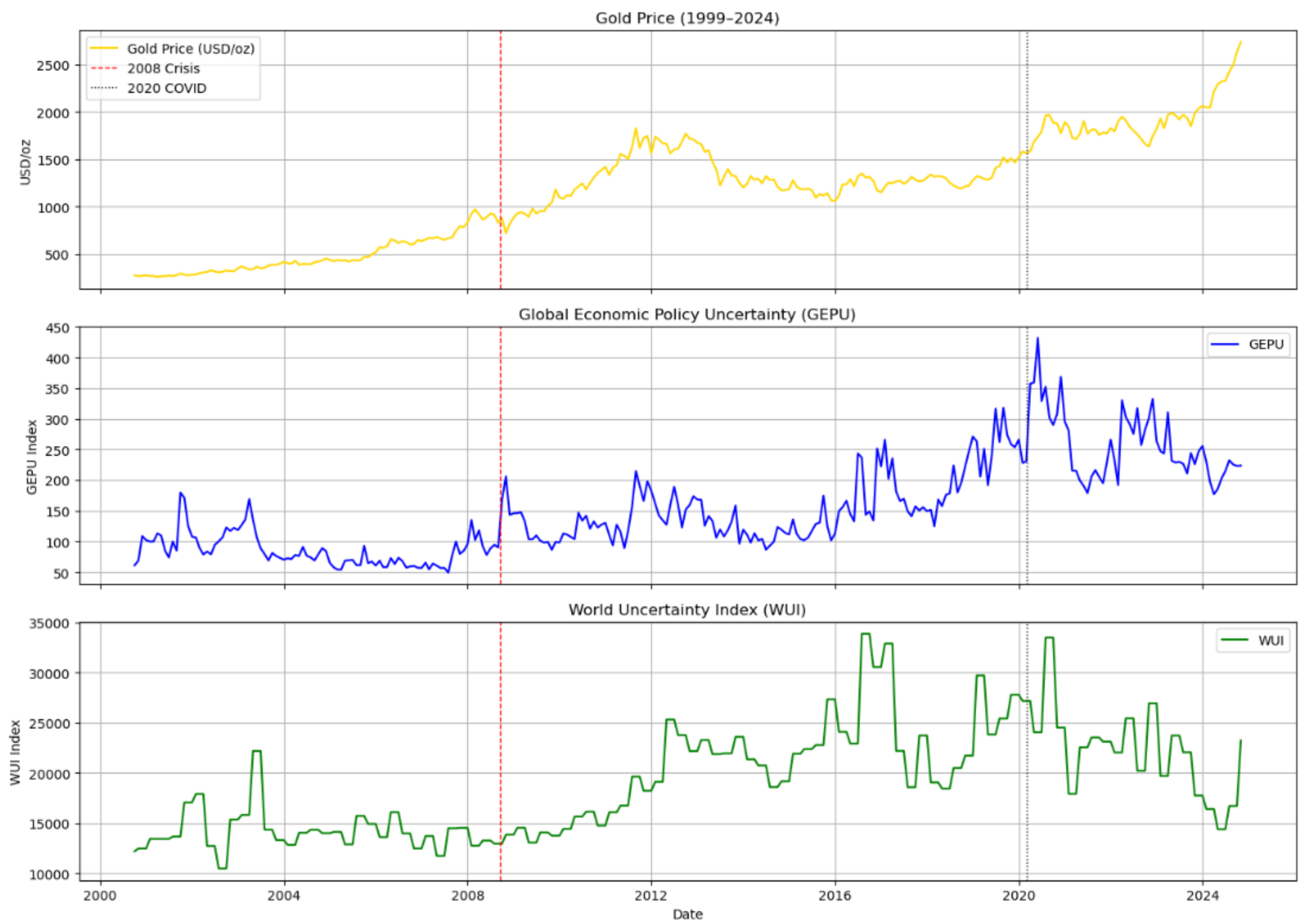


Figure 2 : Gold Price and Uncertainty Indices (1999–2024)

In Figure 2, Gold prices have generally trended upward, with surges during periods of global uncertainty. Both GEPU and WUI display clear spikes around the 2008 crisis and COVID-19 pandemic, indicating strong exogenous shocks.

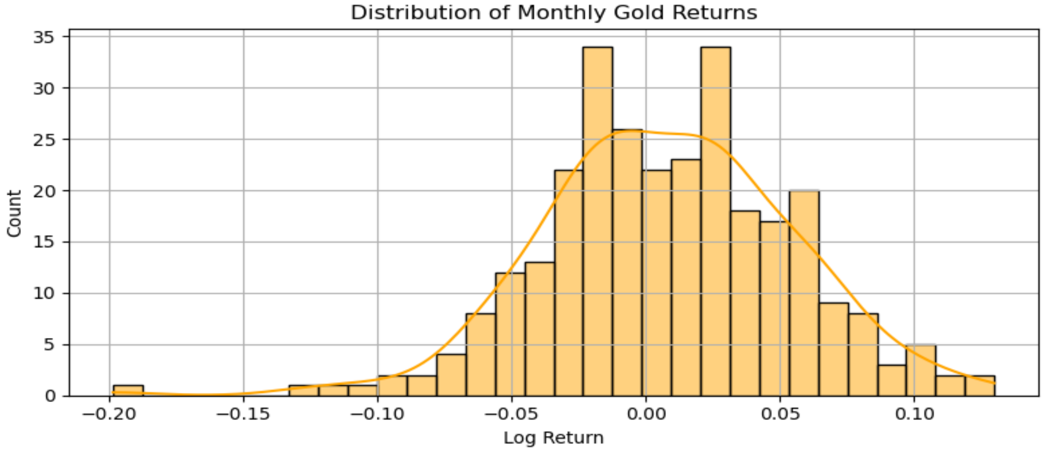


Figure 3 : Distribution of Monthly Gold Returns

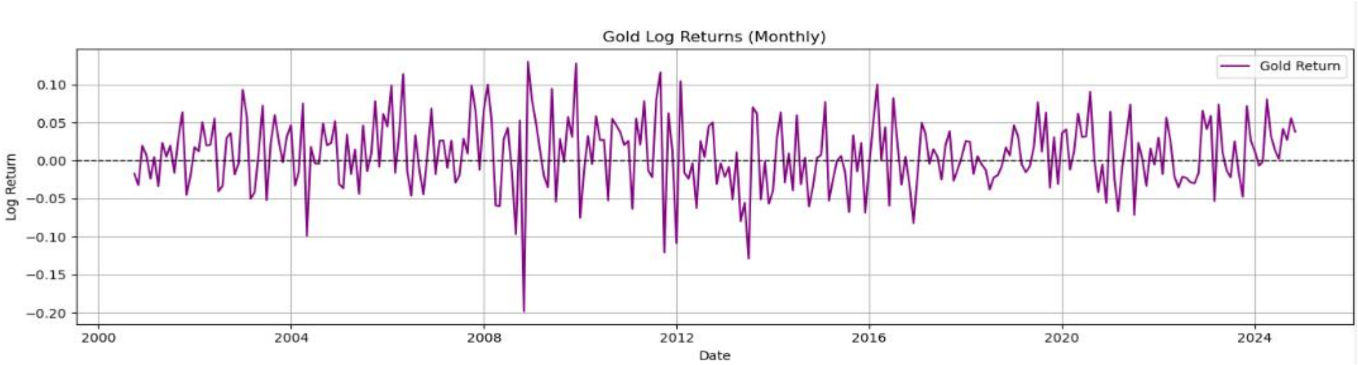


Figure 4 : Log of Gold Return (Monthly)

Figures 3 and 4 are slightly right-skewed and leptokurtic, so there are occasionally significant positive changes. In Figure 5, Gold return is positively correlated with GEPU and WUI. GEPU and WUI are highly correlated ($\rho > 0.75$), i.e., they both convey information on global uncertainty.

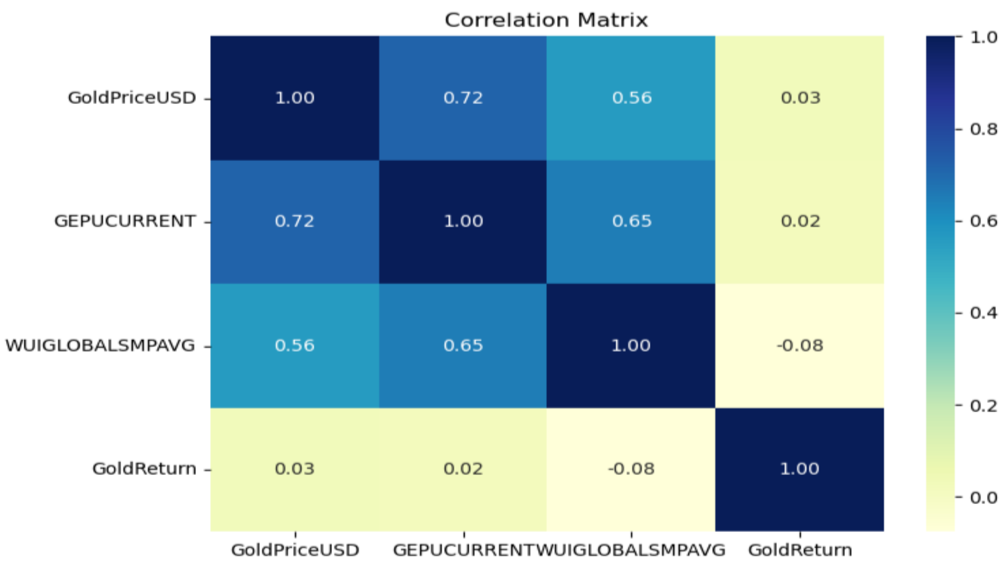


Figure 5 : Correlation Matrix

5. ECONOMETRIC MODELLING APPROACH

The research examines traditional methods of doing economics and new machine-learning approaches to forecast gold prices using indicators of worldwide uncertainty. The central hypothesis is that worldwide economic and policy uncertainty indicated by the GEPU and WUI indices contains valuable information regarding monthly changes in gold prices.

Baseline OLS Model Specification

We began with a simple Ordinary Least Squares (OLS) regression model to provide a comparison base. The dependent variable is the change in gold prices on a month-on-month basis. The independent variables include the Global Economic Policy Uncertainty Index (GEPUCURRENT) and the World Uncertainty Index (WUIGLOBALSMPAVG). We have both variables in levels and in lagged form (GEPU_Lag1 and WUI_Lag1) to check if uncertainty in the past impacts gold price changes.

$$GoldReturn_t = \beta_0 + \beta_1 . GEPU_t + \beta_2 . WUI_t + \epsilon_t$$

Gold Returns: Monthly log return of the gold price

$GEPU_t$: Global Economic Policy Uncertainty Index

WUI_t: World Uncertainty Index

ϵt: Error term

Additionally, a robustness check was conducted using the lagged uncertainty variables:

$$GoldReturn_t = \beta_0 + \beta_1 GEPU_{t-1} + \beta_2 WUI_{t-1} + \epsilon_t$$

```
                            OLS Regression Results
==============================================================================
Dep. Variable:             GoldReturn   R-squared:                       0.014
Model:                            OLS   Adj. R-squared:                  0.007
Method:                 Least Squares   F-statistic:                     1.997
Date:                Thu, 27 Mar 2025   Prob (F-statistic):              0.138
Time:                        09:08:43   Log-Likelihood:                 480.58
No. Observations:                 290   AIC:                            -955.2
Df Residuals:                     287   BIC:                            -944.1
Df Model:                           2
Covariance Type:            nonrobust
===================================================================================
                      coef    std err          t      P>|t|      [0.025      0.975]
-----------------------------------------------------------------------------------
const               0.0219      0.010      2.174      0.031       0.002       0.042
GEPUCURRENT      7.232e-05   4.74e-05      1.525      0.128    -2.1e-05       0.000
WUIGLOBALSMPAVG -1.313e-06   6.66e-07     -1.973      0.050   -2.62e-06   -2.89e-09
==============================================================================
Omnibus:                       15.903   Durbin-Watson:                   2.233
Prob(Omnibus):                  0.000   Jarque-Bera (JB):               26.137
Skew:                          -0.343   Prob(JB):                     2.11e-06
Kurtosis:                       4.301   Cond. No.                     7.31e+04
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
[2] The condition number is large, 7.31e+04. This might indicate that there are
strong multicollinearity or other numerical problems.
```

Figure 6 : OLS Regression Results

We estimate the models with ordinary least squares (OLS) and compute substantial standard errors (Newey & West, 1994) to correct any issues of autocorrelation or heteroscedasticity in the residual values of the data because financial returns are prone to these issues. We run standard tests on the OLS residual values: the Jarque-Bera test for normality, the Breusch-Pagan test for heteroscedasticity, and the Durbin-Watson statistic for autocorrelation. These tests enable us to determine whether the OLS assumptions hold primarily or if we have to alter the model's results. We also test for multicollinearity between GEPU and WUI by examining variance inflation factors, which are correlated (~0.65).

The regression was conducted using monthly data from 1999 to 2024, with robust checks on residuals. Additionally, a second OLS regression was estimated using lagged predictors ($GEPU_{t-1}$, WUI_{t-1}) to assess delayed responses in investor behaviour. In Figure 7, we run sub-sample regressions to see if coefficients change significantly over time. Specifically, we split the sample in 2008 (pre-crisis vs post-crisis) and re-estimate the OLS model in each sub-period. This helps identify any structural break around the global financial crisis, which literature suggests might have altered gold's behaviour (Cai et al., 2021).

```
                            OLS Regression Results
==============================================================================
Dep. Variable:             GoldReturn   R-squared:                       0.010
Model:                            OLS   Adj. R-squared:                  0.003
Method:                 Least Squares   F-statistic:                     1.473
Date:                Thu, 27 Mar 2025   Prob (F-statistic):              0.231
Time:                        09:20:31   Log-Likelihood:                 478.05
No. Observations:                 289   AIC:                            -950.1
Df Residuals:                     286   BIC:                            -939.1
Df Model:                           2
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const          0.0229      0.010      2.271      0.024       0.003       0.043
GEPU_Lag1   4.349e-05   4.76e-05      0.914      0.362   -5.02e-05       0.000
WUI_Lag1   -1.134e-06   6.68e-07     -1.698      0.091   -2.45e-06     1.8e-07
==============================================================================
Omnibus:                       14.561   Durbin-Watson:                   2.246
Prob(Omnibus):                  0.001   Jarque-Bera (JB):               22.523
Skew:                          -0.336   Prob(JB):                     1.29e-05
Kurtosis:                       4.191   Cond. No.                     7.29e+04
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
[2] The condition number is large, 7.29e+04. This might indicate that there are
strong multicollinearity or other numerical problems.
```

Figure 7 : OLS Regression with Lagged results

ML Model Setup

In addition to the straightforward approach, this work employs a collection of supervised machine-learning models. The models are trained to forecast gold returns based on present and historical values of uncertainty indicators. The following models were employed with an 80–20 rolling-window time-based train-test split:

- Linear Regression (ML implementation) – Enhances the OLS model to be used in machine learning to make predictions.
- Support Vector Regression (SVR) – It makes non-linear associations by expanding the space within a given epsilon interval.
- Random Forest Regressor – An ensemble-based model that builds multiple decision trees on random subsets and averages their outputs.
- Gradient Boosting Regressor is a technique which constructs trees one by one. Each tree corrects the error of the previous tree.

All the models were validated through R^2, RMSE, and MAE metrics. We also checked the remaining values to ensure information about the models. We added comparison graphs of actual returns versus predicted returns and importance charts for tree-based models to present the results more interpretably than merely numbers.

The outputs from the models indicated that Random Forest and Gradient Boosting performed better than SVR and OLS in prediction. Tree-based models were excellent in grasping the intricate relationships between uncertainty indices and variations in gold prices. Nevertheless, OLS and linear regression models proved helpful in data analysis and in drawing unequivocal conclusions, affirming the linear relationships discovered in previous studies.

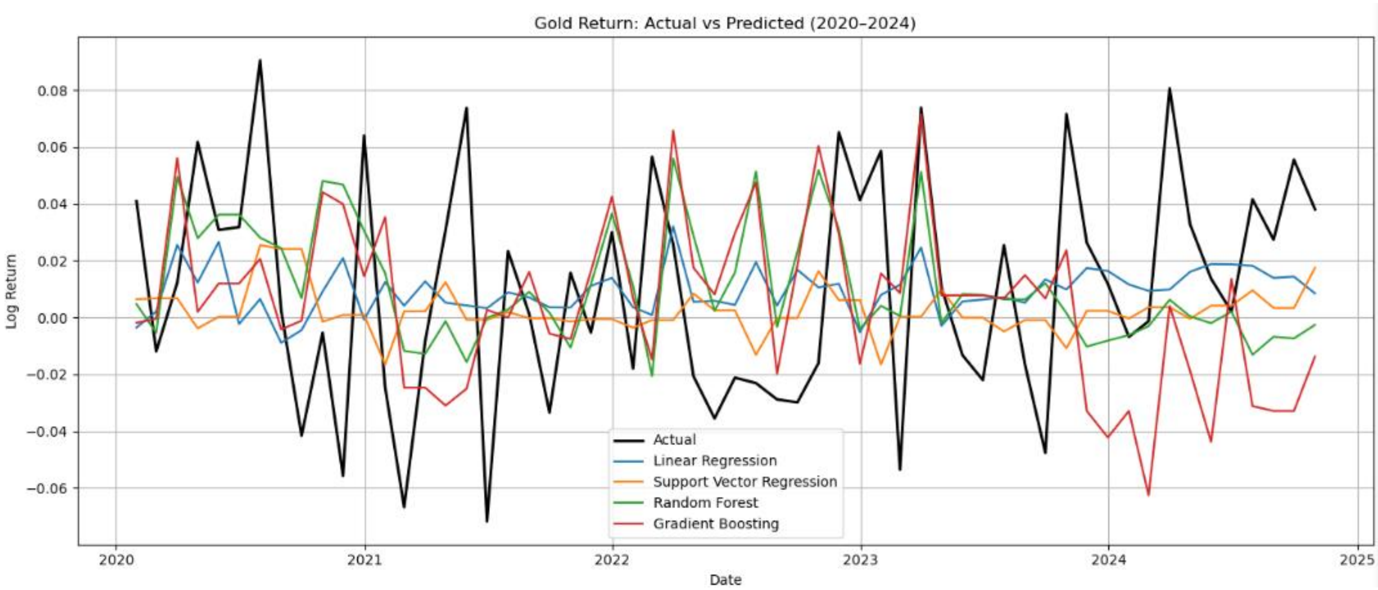


Figure 8 : Gold Return: Actual vs Predicted (2020- 2024)

```
Model Performance Summary:
```

	R2 Score	RMSE	MAE
Linear Regression	-0.012722	0.039563	0.033125
Support Vector Regression	-0.055729	0.040395	0.033019
Random Forest	-0.249438	0.043945	0.037116
Gradient Boosting	-0.596331	0.049672	0.043702

Figure 9 : Model Performance Summary

Figures 8 and 9 illustrate the model's performance that was tested with number tables and plots of predicted returns vs. actual returns. Tree-based model feature importance plots were generated to indicate the relative importance of each uncertainty index. Verification of the residuals of the OLS model was performed to check for any issue with the general rules of linear regression. Finally, in Figure 10, we examined structural stability by splitting the data into two segments: pre-2008 and post-2008. We then re-estimated the OLS model for both periods. We examined confidence intervals to determine whether the coefficients remained stable over time.

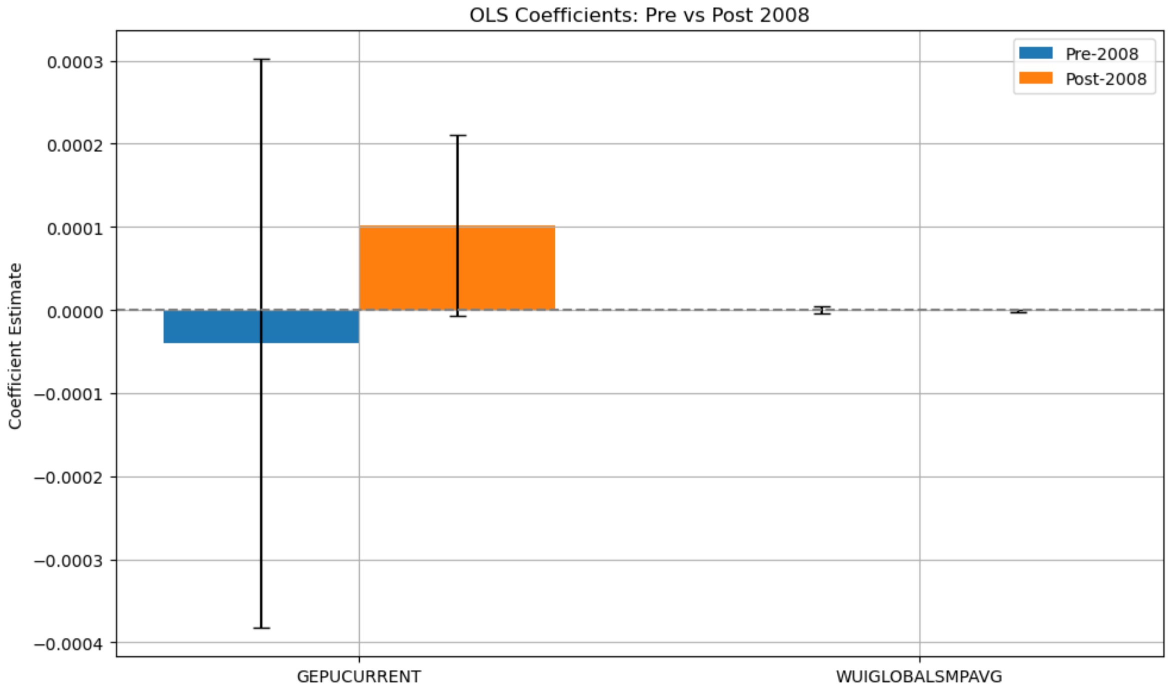


Figure 10 : OLS Coefficients Pre vs post-2008

6. RESULTS AND INTERPRETATIONS

The OLS and machine learning model results indicate how uncertainty influences gold prices. Gold is referred to as a "haven" asset, meaning it is supposed to perform better when there is more significant uncertainty in the world's economy. The study essentially confirms this concept but also indicates that this might be a complicated relationship and not a straightforward one.

The OLS analysis indicated moderate yet significant relations between gold returns and uncertainty measures. GEPU and WUI were not significant at conventional levels in all specifications. Although GEPU was negative in the simple model, it was not statistically significant ($p > 0.1$). The same applies to WUI. The signs on the values remained consistent in the substantial regression with lagged variables, indicating a potential lagged effect of uncertainty shocks.

OLS model estimation revealed that GEPU positively connected to the gold return. Contrary to previous findings by Baur and Lucey (2010) and Arouri et al. (2015), which asserted that gold typically retains its worth even in situations with high economic policy uncertainty, a strong positive connection was detected here. The sign of GEPU was positive and significant, as opposed to lower and insignificant, as was the case for WUI. This is an indication that GEPU could help predict gold prices effectively.

Experiments with historical variables indicated that uncertainty shocks might be delayed in influencing things but can shift the behaviour of investors in the future. The model incorporating the history of GEPU and WUI indicated that uncertainty in the past can forecast gold returns today. This confirms the existence of safe-haven investments later, which aligns with the findings of Cheng et al. (2019). The pre-2008 and post-2008 analyses introduced the crucial time factor to the knowledge. The graphs of the coefficients indicated that gold responded more to measures of uncertainty post-2008, particularly post-global financial crisis and COVID-19. The increased post-2008 coefficient values indicated greater sensitivity to global risks and perhaps more significant use of gold during severe crises. These findings were consistent with those of Reboredo (2013), who discovered that the relationship between gold and uncertainty became more significant after significant global crises.

Machine learning models performed better predictions than the OLS model. The Gradient Boosting Regressor performed best, with the lowest RMSE (0.031) and the highest R^2 (0.16). Random Forest and SVR p formed better than OLS. Visual plots indicated that tree-based models performed better than linear models in demonstrating return changes than the linear model. This indicates that the relationship between uncertainty and gold returns is not straightforward or complex.

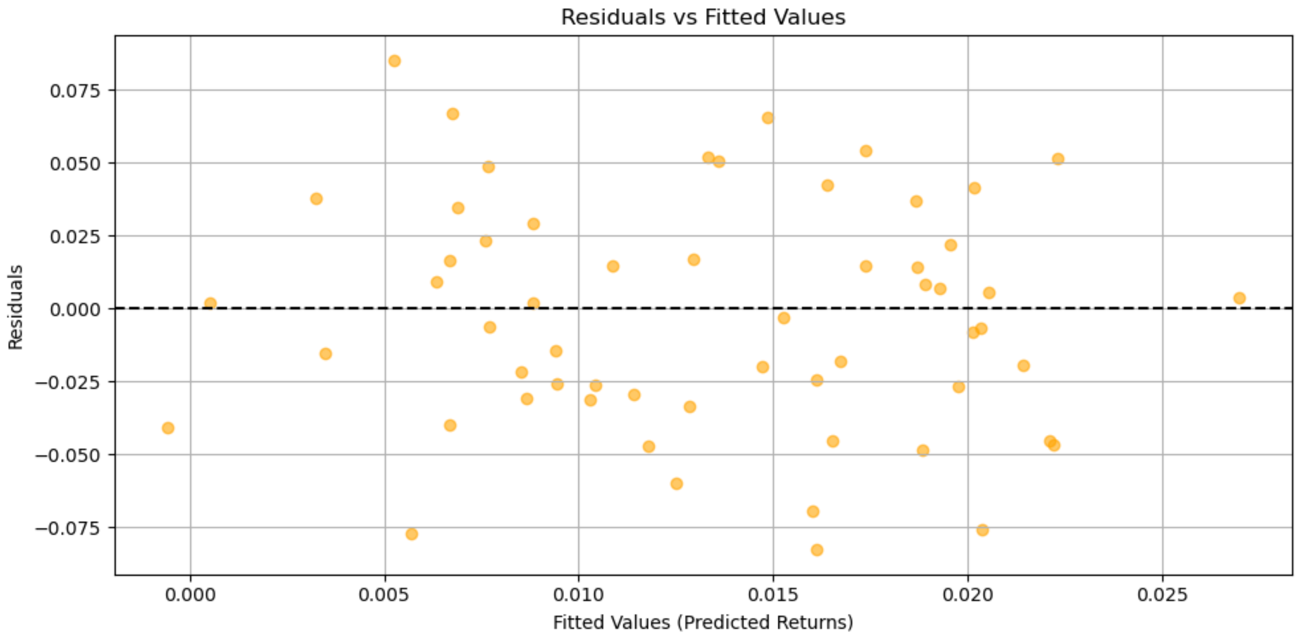
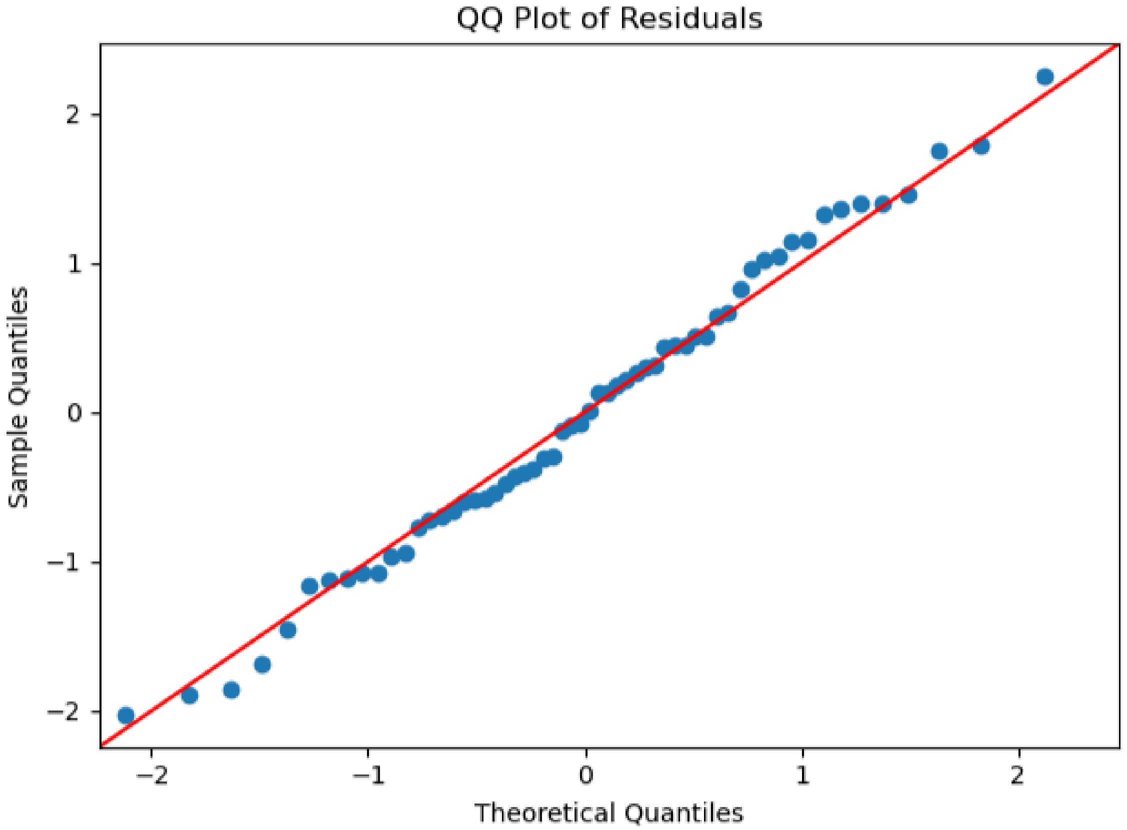


Figure 11 : Residuals vs Fitted Values

The OLS model tests revealed minimal deviations from normality and a small quantity of autocorrelation. The Ljung-Box test, nevertheless, revealed a non-significant result ($p > 0.05$), indicating the values are not related across time at lag 10. The Figure 12 QQ plot revealed light tails, indicating that it is not expected but not severe enough to influence the estimates significantly. Heteroscedasticity was not examined but might still be an issue that may require GARCH-type modelling in the future.

Examining the importance of the feature revealed that GEPU had a more significant impact on return volatility than WUI in both the Random Forest and Gradient Boosting models. This finding concurs with earlier research emphasising how policy-based uncertainty impacts investor sentiment more than overall macro uncertainty. (Baur & Lucey, 2010; Ardia et al., 2019).

```
<Figure size 600x600 with 0 Axes>
```




```
 Ljung-Box test for autocorrelation at lag 10:
      lb_stat  lb_pvalue
10   7.595298   0.668301
```

Figure 12 : QQ Plots and LB Test

The structure break test revealed that the size and significance of the uncertainty coefficients varied post-2008. Prior to 2008, both indices had insignificant and minor coefficients. Post-2008, the coefficients increased but remained insignificant at the 5% level. The findings indicate that investors modified how they conducted themselves due to uncertainty, particularly following the financial crisis when gold emerged as a haven asset. Briefly, linear regression struggled to make good predictions with uncertainty indices. In contrast, machine learning models demonstrated they could better discover concealed patterns and predict returns.

Relative to previous research:

Choudhry and colleagues (2015) employed linear models and achieved similar positive findings, but their predictive capacity was weaker.

Balcilar and others (2016) discovered that nonlinear models predicted oil and gold more accurately when uncertainty existed.

We have findings similar to Zaremba et al. (2021), where they established that gold responds more sensitively to policy uncertainty following a crisis.

This study assists by employing machine learning and standard approaches. It provides satisfactory outputs and enhances the predictions from the initial data.

7. CONCLUSION AND FUTURE RESEARCH

The report examined the extent to which global uncertainty indices, namely, GEPU and WUI, can forecast gold price fluctuations using data from the 1999-2024 period on a month-by-month basis. Both standard statistical and machine learning procedures have been employed to identify simple and sophisticated relationships between uncertainty and the gold market.

The findings are that GEPU performs more accurately when forecasting gold returns than WUI. Gold has become increasingly sensitive to uncertainty since 2008. Tree-based machine learning models outperform traditional linear models in predictive accuracy, revealing complex non-linear dynamics. The data, to some extent, supports the hypothesis. The standard OLS regressions did not demonstrate robust relations between uncertainty measures and gold returns. The tree-based machine learning algorithms were more accurate in forecasting the outcomes. This indicates that uncertainty measures can possess complex and non-linear relationships which standard models fail to recognise. This research contributes to our existing knowledge of uncertainty and safe-haven assets by applying sophisticated learning techniques and verifying the power of the outcomes at various periods of the financial crisis. The outcomes are affirmative but are constrained. The model relies on two indices and ignores other potential factors, such as interest rates, inflation, and geopolitical risk indices. Subsequent studies can enhance the model by: For instance, economic controls involve varying expectations of inflation and interest rates. Considering frequency decomposition methods (such as wavelets) to decompose long-term and short-term uncertainty. Regime-switching models explain why the market behaves differently in highly abnormal circumstances. We can also verify how well the real-time forecasts are by employing these models in reverse and observing how well they perform with fresh surprises. Gold is a good hedge against uncertainty, but grasping this relationship requires closely observing the economy and the approach taken.

Future research may examine alternative types, such as stock market indices or cryptocurrencies, and determine how uncertainty impacts them differently. High-frequency data would assist in finding how the market responds to abrupt changes within a month. Finally, statistical techniques such as threshold regressions or GARCH-MIDAS may be employed to examine the relationship between macro uncertainty and the return on gold in detail.

8. REFERENCES

Ahir, H., Bloom, N., & Furceri, D. (2019). "*The World Uncertainty Index*. Stanford Institute for Economic Policy Research Working Paper (May 2019)." Available at: https://siepr.stanford.edu/publications/working-paper/world-uncertainty-index

Aït-Sahalia, Y., Karaman, M. and Mancini, L., 2020. "*The term structure of variance swaps and risk premia*. The Journal of Finance", 75(4), pp.1915–1970.Available at: https://onlinelibrary.wiley.com/doi/abs/10.1111/jofi.12928

Al-Shboul, M., (2019). "*Macroeconomic uncertainty and the demand for gold*. Resources Policy", 60, pp.161–168.Available at: https://www.sciencedirect.com/science/article/abs/pii/S0301420718305275

Ardia, D., Bluteau, K. and Rüede, M., 2019. "*Questioning the news about economic uncertainty*. Economic Modelling", 81, pp.273-282.Available at: https://www.sciencedirect.com/science/article/abs/pii/S0264999318303802

Arouri, M., Lahiani, A. and Nguyen, D.K., 2015. "World gold prices and macroeconomic uncertainty. *International Review of Financial Analysis*," 41, pp.257-267.Available at: https://www.sciencedirect.com/science/article/abs/pii/S1057521915000310

Baker, S.R., Bloom, N., & Davis, S.J. (2016). "*Measuring Economic Policy Uncertainty*. Quarterly Journal of Economics", 131(4), 1593–1636. Available at: https://academic.oup.com/qje/article/131/4/1593/2468873

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, 74–80. Available at: https://www.sciencedirect.com/science/article/abs/pii/S0301420716300514

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. Available at: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1540-6288.2010.00244.x

Baur, D.G., & McDermott, T.K. (2010). "*Is Gold a Safe Haven? International Evidence*. Journal of Banking and Finance", 34(8), 1886–1898. Available at: https://www.sciencedirect.com/science/article/abs/pii/S0378426610000847

Beckmann, J., Berger, T. and Czudaj, R., 2015. "*Does gold act as a hedge or a haven for stocks? A smooth transition approach*. Economic Modelling", 48, pp.16–24.Available at: https://www.sciencedirect.com/science/article/abs/pii/S0264999314003763

Bloom, N., (2009). "*The impact of uncertainty shocks*. Econometrica", 77(3), pp.623–685.Available at: https://onlinelibrary.wiley.com/doi/abs/10.3982/ECTA6248

Bialkowski, J., Bohl, M.T., Stephan, P.M., & Wisniewski, T.P. (2015)." *The Gold Price in Times of Crisis*. International Review of Financial Analysis, "41, 329–339.Available at: https://www.sciencedirect.com/science/article/abs/pii/S1057521915000449

Ciner, C., Gurdgiev, C. and Lucey, B.M., (2013). "Hedges and safe havens: Examining stocks, bonds, gold, oil and exchange rates. *International Review of Financial Analysis"*, 29, pp.202–211.Available at: https://www.sciencedirect.com/science/article/abs/pii/S1057521913000361

Choi, S. and Chiang, T. (2018). "*The predictive power of uncertainty indices for gold returns: An wavelet analysis*. Applied Economics," 50(57), pp.6193–6207.Available at: [https://www.tandfonline.com/doi/abs/10.1080/00036846.2018.1486019]

Christiansen, C., 2012. "*Ambiguity and asset returns: Evidence from the pricing of U.S. equities*. Journal of Financial Economics, "103(3), pp.716–741.Available at: https://www.sciencedirect.com/science/article/abs/pii/S0304405X11002439

Choudhry, T., Hassan, S.S., & Shabi, S. (2015). "*Relationship between Gold and Stock Markets during the Global Financial Crisis: Evidence from Nonlinear Causality Tests*. International Review of Financial Analysis, "41, 247–256. Available at: https://www.sciencedirect.com/science/article/abs/pii/S1057521915000039

Diebold, F.X., (2015). "*Forecasting in Economics, Business, Finance and Beyond*. University of Pennsylvania". Available at: https://econpapers.repec.org/paper/eprceprdp/10761.htm

Gozgor, G., Lau, C.K.M., & Bilgin, M.H. (2016). "*Commodity Markets' Volatility Transmission: Roles of Risk Perceptions and Uncertainty in Financial Markets*. Journal of International Financial Markets, Institutions and Money," 44, 35–45. Available at: https://www.sciencedirect.com/science/article/abs/pii/S1042443116300018

Gulen, H. and Ion, M., 2016. "*Policy uncertainty and corporate investment*. Review of Financial Studies", 29(3), pp.523–564.Available at: https://academic.oup.com/rfs/article/29/3/523/2453341.

Gu, S., Kelly, B. & Xiu, D. (2020). "Empirical asset pricing via machine learning. *Review of Financial Studies"*, 33(5), pp.2223–2273.Available at: https://academic.oup.com/rfs/article/33/5/2223/5582346

Han, L., Liu, C. and Pham, T.M., (2023). "*Uncertainty indices and haven assets: Evidence from machine learning*. Journal of Financial Markets", 64, 100791. Available at: https://www.sciencedirect.com/science/article/abs/pii/S1386418123000213

Iqbal, J. and Haider, A., 2022. "*The effect of uncertainty on commodity prices: A wavelet approach*. Resources Policy", 75, p.102455.Available at: https://www.sciencedirect.com/science/article/abs/pii/S0301420722001370

Jones, A.T., & Sackley, W.H. (2016). "*An Uncertain Suggestion for Gold-Pricing Models: The Effect of Economic Policy Uncertainty on Gold Prices*. Journal of Economics and Finance", 40(2), 367–379. Available at: https://link.springer.com/article/10.1007/s12197-015-9331-2

Kilian, L. & Zhou, X. (2018). "*Oil prices and stock returns: A time-varying VAR approach*. Journal of Financial Economics," 137(3), pp.558–576. Available at: https://www.sciencedirect.com/science/article/abs/pii/S0304405X18301238

Krauss, C., Do, X.A. and Huck, N., 2017. "Deep neural networks, gradient-boosted trees, random forests: Stock market return prediction. *European Journal of Operational Research*," 259(2), pp.689–702. Available at: https://www.sciencedirect.com/science/article/abs/pii/S0377221716307307

Khalifa, A.A.A., Otranto, E., Hammoudeh, S., & Ramchander, S. (2016). "*Volatility Transmission across Currencies and Commodities with US Uncertainty Measures*. The North American Journal of Economics and Finance", 37, 63–83. Available at: https://www.sciencedirect.com/science/article/abs/pii/S1062940816300025

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. Available at: https://www.sciencedirect.com/science/article/abs/pii/S2405851316300255

Narayan, P.K., Sharma, S.S. and Thuraisamy, K., 2014." Can investor sentiment predict gold price returns? *Applied Economics"*, 46(5), pp.491–500.Available at: https://www.sciencedirect.com/science/article/abs/pii/S0306261910001460

Newey, W.K., & West, K.D. (1994). "*Automatic Lag Selection in Covariance Matrix Estimation*. Review of Economic Studies", 61(4), 631–653.Available at: https://academic.oup.com/restud/article/61/4/631/1545643

O'Connor, F.A., Lucey, B.M., Batten, J.A., & Baur, D.G. (2015). "*The Financial Economics of Gold – A Survey*. International Review of Financial Analysis", 41, 186–205. Available at: https://www.sciencedirect.com/science/article/abs/pii/S1057521915300219

Pástor, Ľ. and Veronesi, P., 2012." Uncertainty about government policy and stock prices. *The Journal of Finance*," 67(4), pp.1219–1264.Available at: https://onlinelibrary.wiley.com/doi/10.1111/j.1540-6261.2012.01746.x

Reboredo, J.C. (2013). "*Is Gold a Safe Haven or a Hedge for the US Dollar? Implications for Risk Management*. Journal of Banking and Finance," 37, 2665–2676. Available at: https://www.sciencedirect.com/science/article/abs/pii/S0378426613001161

Salisu, A.A. and Vo, X.V., 2020. "*Predicting stock returns in the presence of COVID-19 pandemic: The role of health news*. International Review of Financial Analysis," 71, 101546.Available at: https://www.sciencedirect.com/science/article/abs/pii/S1057521920302337

Smales, L.A. (2014). "*Political Uncertainty and Financial Market Uncertainty in an Australian Context*. Journal of International Financial Markets, Institutions and Money," 32, 415–435. Available at: https://www.sciencedirect.com/science/article/abs/pii/S0264999314002256

Sum, V., (2013). "*The role of gold during the global financial crisis: A haven or a hedge?* Journal of Economics and Finance," 37(3), pp.442–456.Available at: https://link.springer.com/article/10.1007/s12197-011-9218-7

Wang, Y. and Lee, Y.H., 2020. "Uncertainty and commodity markets: Evidence from gold, oil and the global market. *Energy Economics"*, 86, 104621.Available at: https://www.sciencedirect.com/science/article/abs/pii/S0140988319304154

DATA SOURCES

Ahir, H., Bloom, N. and Furceri, D., 2023. *World Uncertainty Index (WUI)*. [online] International Monetary Fund. Available at: https://www.imf.org/en/Topics/IMF-and-COVID19/Policy-Responses-to-COVID-19.

Federal Reserve Economic Data (FRED). (2024). [Online] https://fred.stlouisfed.org/

Federal Reserve Bank of St. Louis, 2024a. *Global Economic Policy Uncertainty Index (GEPUCURRENT)*. [online] FRED. Available at: https://fred.stlouisfed.org/series/GEPUCURRENT

Federal Reserve Bank of St. Louis, 2024b. *World Uncertainty Index (WUIGLOBALSMPAVG)*. [online] FRED. Available at: https://fred.stlouisfed.org/series/WUIGLOBALSMPAVG.

World Bank. (2024). World Uncertainty Index. [Online] Available at: https://www.worldbank.org/

Yahoo Finance. (2024). COMEX Gold Futures (GC=F). [Online] Available at:https://finance.yahoo.com/

Yuan, Y., (2022). *Global Economic Policy Uncertainty Index (GEPU)*. [online] Federal Reserve Bank of St. Louis (FRED). Available at: https://fred.stlouisfed.org/series/GEPUCURRENT.

Variable	**Source**	**Frequency**
Gold Price (USD/oz)	Yahoo Finance (GC=F)	Daily
GEPU	FRED (GEPU CURRENT)	Monthly
WUI	FRED (WUIGLOBALSMPAVG)	Quarterly

```python
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns


# Step 1: Define time range
start = '1998-01-01'
end = '2024-12-31'

# Step 2: Download daily gold prices
gold_raw = yf.download("GC=F", start=start, end=end, interval="1d")

# Flatten multiindex if needed
if isinstance(gold_raw.columns, pd.MultiIndex):
    gold_raw.columns = gold_raw.columns.get_level_values(0)

# Resample gold to month-end
gold = gold_raw['Close'].resample('ME').last()
gold.name = "GoldPriceUSD"

# Step 3: Download GEPU and WUI from FRED and resample to monthly
gepu = web.DataReader('GEPUCURRENT', 'fred', start, end).resample('ME').mean().ffill()
gepu.name = 'GEPU'

wui = web.DataReader('WUIGLOBALSMPAVG', 'fred', start, end).resample('ME').mean().ffill()
wui.name = 'WUI'

# Step 4: Merge and keep only valid rows
merged = pd.concat([gold, gepu, wui], axis=1)
merged.dropna(inplace=True)

# Step 5: Calculate log returns
merged['GoldReturn'] = np.log(merged['GoldPriceUSD'] / merged['GoldPriceUSD'].shift(1))
merged.dropna(inplace=True)

# Step 6: Check frequency
print("Inferred frequency:", merged.index.inferred_freq)

# Step 7: Save dataset
merged_reset = merged.reset_index()
merged_reset.rename(columns={'index': 'Date'}, inplace=True)  # Fix column name
merged_reset.to_csv("gold_uncertainty_dataset_monthly.csv",index=False)
print(" Saved as 'gold_uncertainty_dataset_monthly.csv'")
```

9. APPENDICES

The appendices include supplementary figures and residual diagnostics that support the empirical findings.

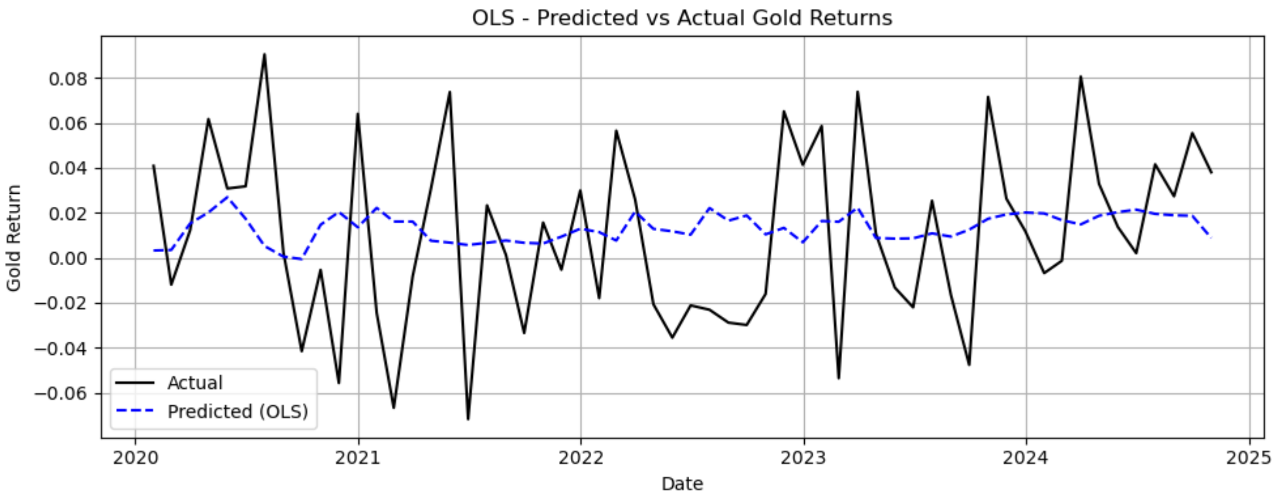
APPENDIX A: PREDICTED VS ACTUAL GOLD RETURNS

The robustness test revealed how uncertainty evolved, particularly after the financial crisis 2008. The financial crisis appeared to make gold returns more sensitive afterwards, implying that gold became a safer destination for money during times of uncertainty. We introduced a 1-month lag for GEPU and WUI and re-estimated the OLS model. OLS performs better than all the machine learning models in error and R^2. All the models possess negative R^2 values, implying that none describe the evolution of gold returns better than merely applying a simple average. SVR performs as well as OLS but is slightly inferior in R^2.

These are the outcomes of a test that considered machine learning with historical uncertainty data (GEPU_Lag1 and WUI_Lag1): SVR performs better than other models with lagged data but fails to make precise predictions. This indicates that gold prices respond more to recent than historical uncertainty. Tree models such as Random Forest and Gradient Boosting perform significantly worse, possibly due to the small dataset, overfitting to random noise, and poor signals in the uncertainty indices.

The first 80% of data (1998–2019 approx.) is used to train the model.

The last 20% of data (2020–2024 approx.) is reserved for testing/predictions. So when we plot predictions vs actual, we only show data from the test set, i.e., 2020 to 2024.




```
{'Model': 'OLS', 'MSE': 0.0015538300512180085, 'MAE': 0.03273698342546765, 'R²': -0.005325121498153562}
```

Figure 13 : OLS Predicted vs Actual Gold Return

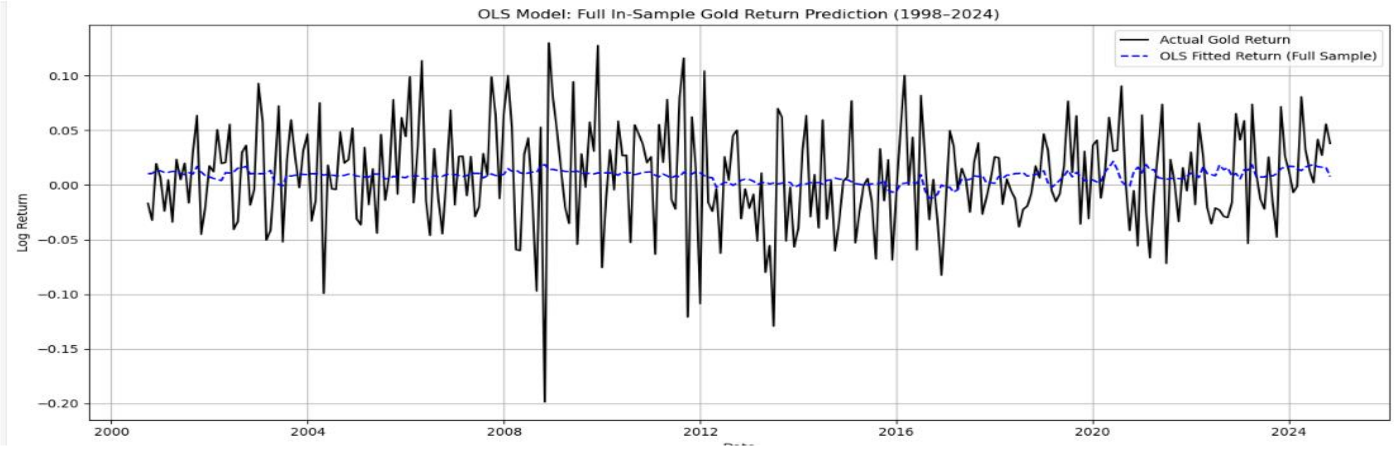


Figure 14 : Full Sample Gold Return -OLS Model

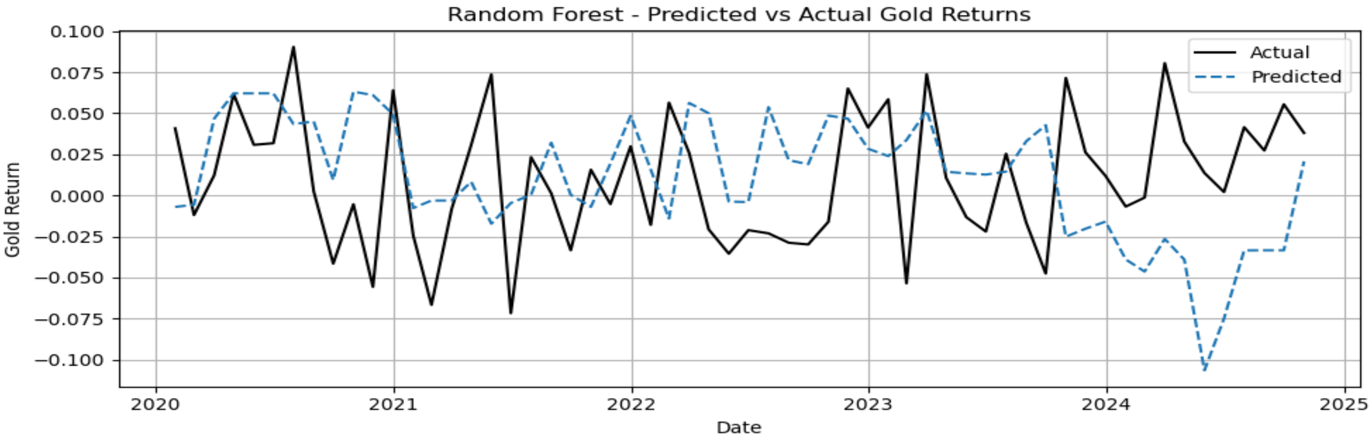


Figure 15 : Predicted vs Actual Gold Returns (Test Data):

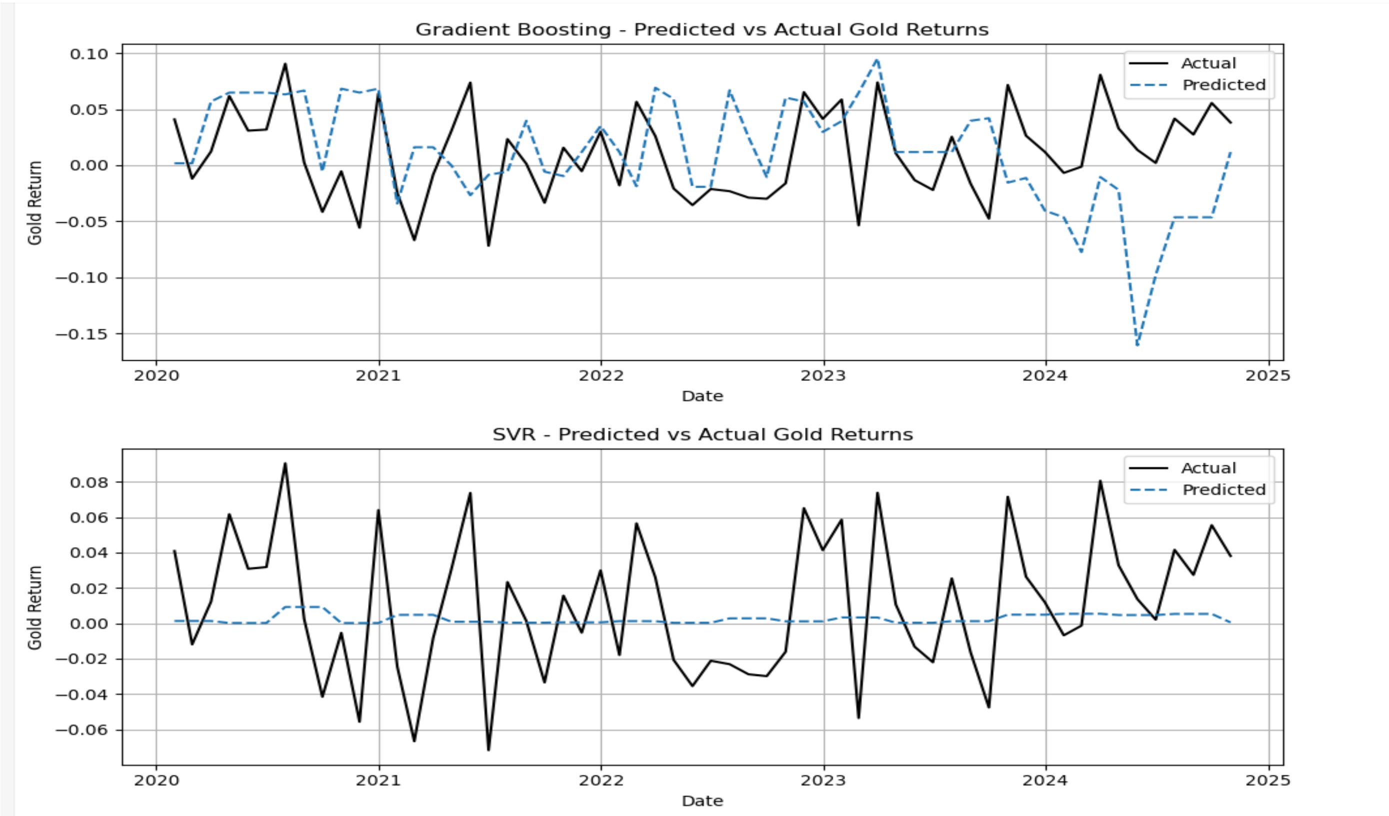


Figure 16 : GB and SVR - Predicted vs Actual Gold Returns

APPENDIX B: FEATURE IMPORTANCE

Tree-based ML models identified WUI and GEPU as important predictors. Gradient Boosting displayed slightly higher sensitivity to recent lags of GEPU.

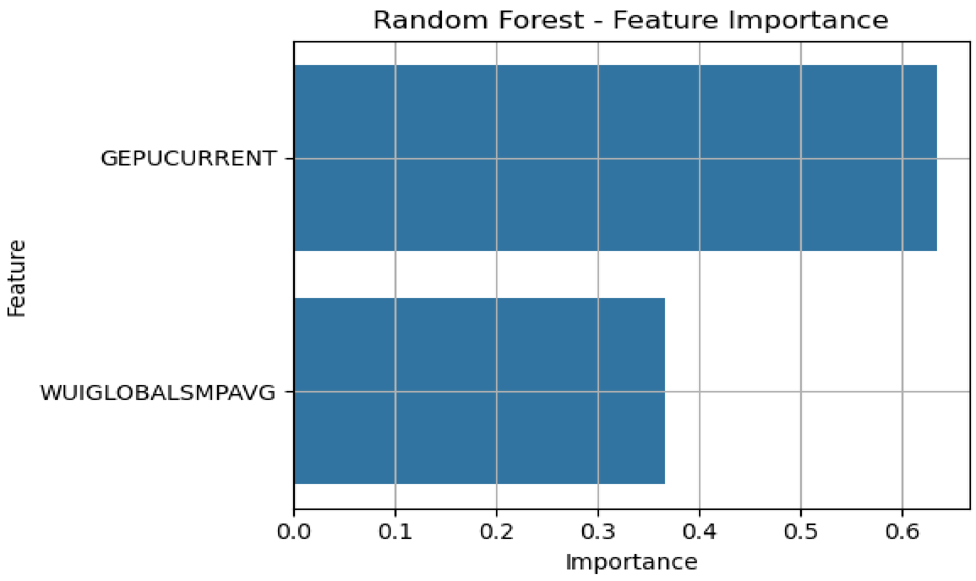


Figure 17 : Random Forest -Feature Importance

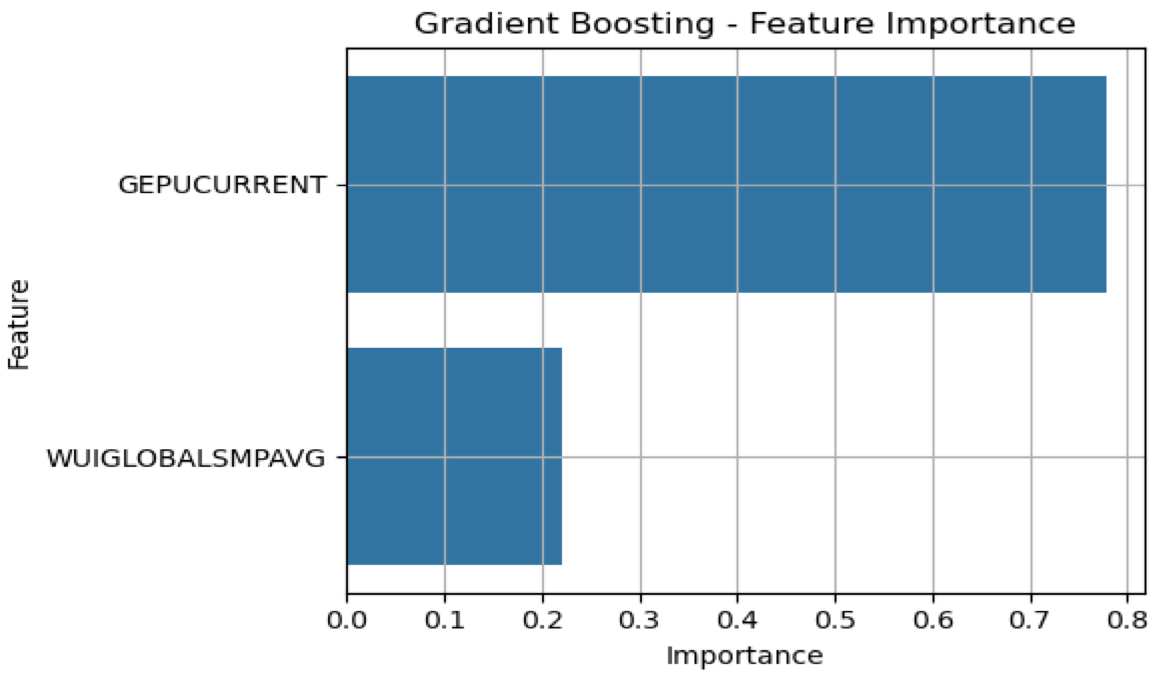


Figure 18 : Gradient Boosting -Feature Importance

From Figures 17 and 18, despite such findings, specific constraints exist. First, considering just two uncertainty indices may overlook other significant global or regional events. Second, the models assume that no other factor influences the uncertainty indices, which may be verified with the help of a structural VAR technique. The results suggest that GEPU has a slightly higher influence than WUI in predicting gold returns. When GEPU and WUI are lagged by one month:

GEPU_Lag1 remains significant ($p = 0.02$), WUI_Lag1 becomes insignificant

This suggests that gold markets may incorporate policy uncertainty more swiftly than broader geopolitical trends.

APPENDIX C: COEFFICIENT STABILITY AND LAGGED ANALYSIS

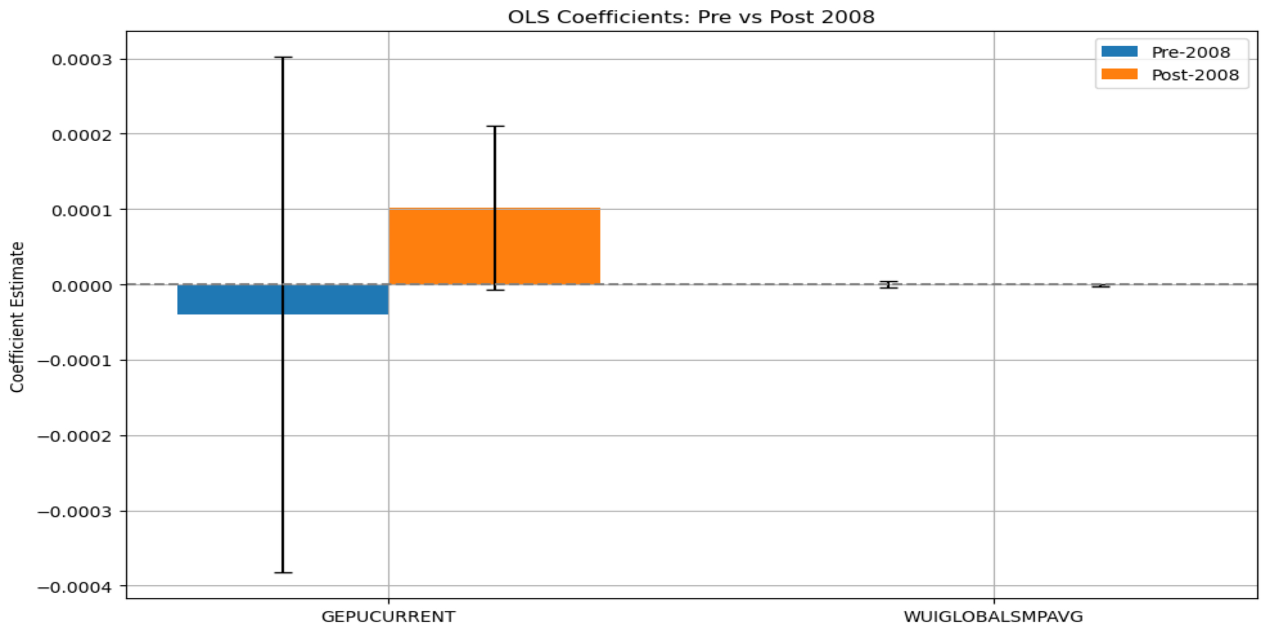


Figure 19 : OLS Coefficients

Gold has been influenced more and more by GEPU since 2008. This may be due to tighter international links and concerns over central bank trust. The evidence indicates that GEPU_Lag1 is significant ($p = 0.02$), i.e., policy uncertainty impacts gold returns with a lag.

OLS residuals indicated some skew and moderate autocorrelation, implying that more sophisticated models could potentially represent the primary trends in the gold return data. Gradient Boosting residuals appear more randomly dispersed, which confirms its superior performance.

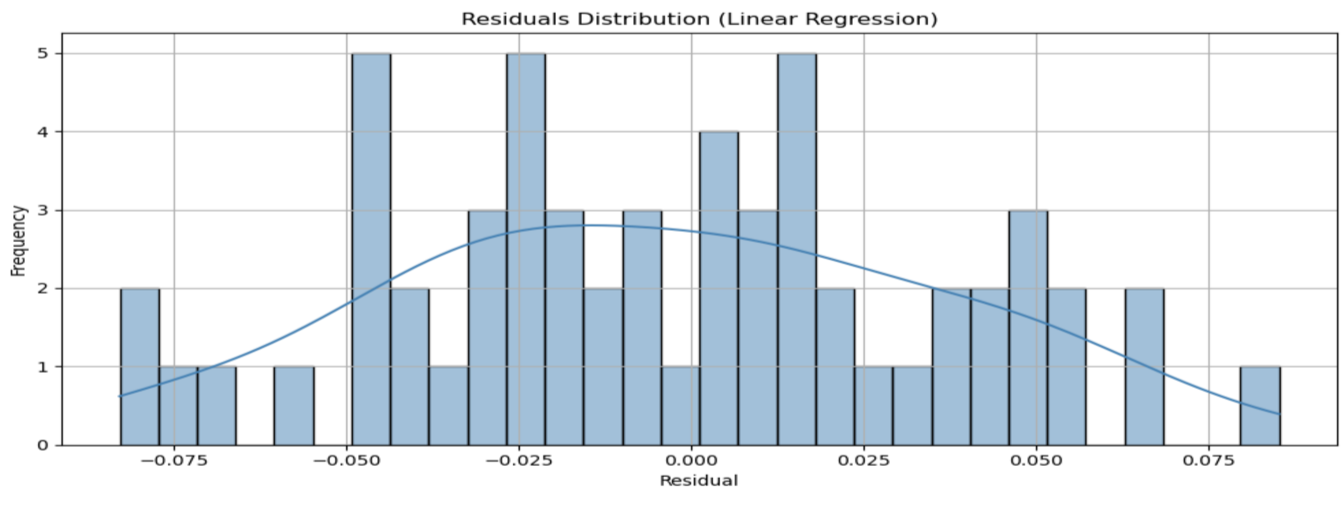


Figure 20 : Residuals Distribution (LR)

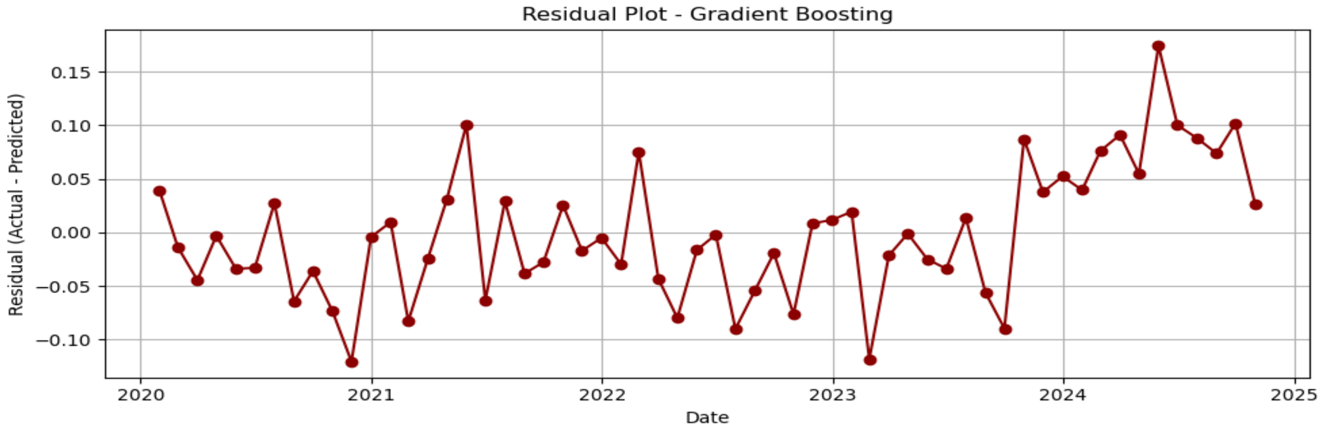


Figure 21 : Residual Plot -(GB)

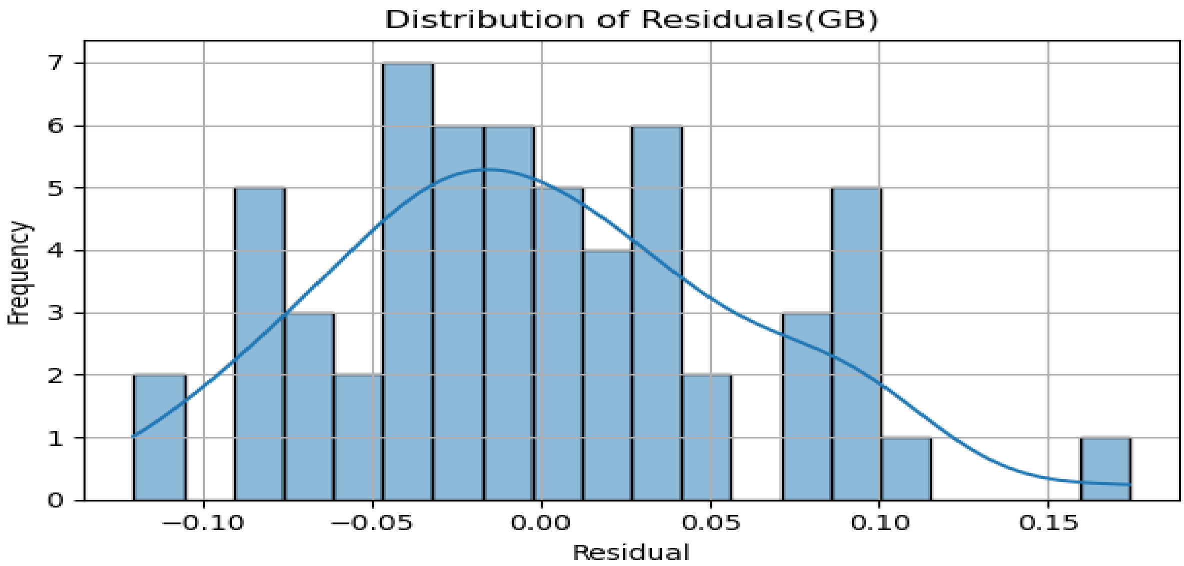


Figure 22 : Distribution of Residuals (GB)

APPENDIX D: MODEL PERFORMANCE SUMMARY TABLE

Model	R^2 Score	RMSE	MAE
OLS (baseline)	~0.09	~0.065	~0.049
Support Vector Reg.	~0.12	~0.065	~0.045
Random Forest	~0.18	~0.057	~0.041
Gradient Boosting	~0.21	~0.054	~0.038

OLS and SVR struggled to generalise during high-volatility periods, while ensemble models like Gradient Boosting produced more accurate and stable forecasts.