Out-of-Sample Predictability of Gold Market Volatility: The role of US Nonfarm Payroll

Afees A. Salisu*, Elie Bouri** and Rangan Gupta***

Abstract

In this study, we make a three-fold contribution to the literature on gold market analysis. First, we provide evidence for the predictive value of US Nonfarm Payroll (USNP) in the out-of-sample forecast of gold market volatility. Second, we extend our analysis to other precious metals and the US stock market index for robustness purposes. Third, we utilize mixed data frequencies based on the availability of data, thus, circumventing any bias or information loss due to the use of monthly (low frequency) USNP data and daily (high frequency) gold price data. The results show that the USNP, which reflects gain/loss in US non-farm jobs, is negatively related to gold return volatility implying that deterioration (improvement) in the economy due to job losses (gains) raises (lowers) the gold market volatility as its trading improves (deteriorates) while the reverse is the case for US stocks. The out-of-sample predictive value of USNP in the return volatility of gold is also established as the model which includes the former offers better out-of-sample forecast gains than the benchmark model which ignores it. Additional analyses involving other precious metals, namely palladium, platinum, rhodium, and silver, show the same direction of relationship as gold, albeit with higher forecast gains for silver than the others. Our findings have useful implications for financial analysts and investors.

Keywords: Gold market volatility; US Nonfarm Payroll; Out-of-sample predictability; GARCH-MIDAS

JEL Codes: E31, F47, J21, J23

*Centre for Econometric and Allied Research, University of Ibadan, Ibadan, Nigeria. Email: adebare1@yahoo.com.

** Corresponding Author. Adnan Kassar School of Business, Lebanese American University, Beirut, Lebanon. Email: elie.elbouri@lau.edu.lb

*** Department of Economics, University of Pretoria, Pretoria, 0002, South Africa. Email: rangan.gupta@up.ac.za.

1. Introduction

In examining the state of an economy, the condition of the labour market is crucial, and the information contained in the United States (US) Nonfarm Payroll (henceforth, USNP) offers a useful guide in this regard. Previous studies indicate that payrolls can have an impact on financial markets (Ederington et al., 2019; Bhatia, 2020). Available evidence suggests that when US non-farm jobs increased from March to April 1996, almost all major sectors showed gains in their employment and this correlated with a moderate growth in the US economy during the first quarter of the same year, with the real Gross Domestic Product (GDP) growing by 2 percent and real personal consumption expenditures increasing by 3.6 percent (Alguire, 1996). The foregoing suggests how crucial Nonfarm Payroll is to the US economy and financial markets, as it is capable of eliciting the sentiments of investors who pay close attention to the stream of macroeconomic news and react to the unexpected components of each release (Caruso, 2019). For example, Elder et al. (2012) examine the intensity, direction and speed of impact of US macroeconomic news announcements on the return, volatility and trading volume of important commodities such as gold, silver and copper futures, and their results suggest that the response of these commodities to economic news surprises is both swift and significant, with Nonfarm Payroll having the largest impact.

Given the foregoing, this paper examines the role of US Nonfarm Payroll in the out-of-sample predictability of gold market volatility by employing long-range data staring from 1968. We attempt to establish the nexus between the state of the US labour market, as captured by Nonfarm Payroll, and risk in the gold market. This is relevant given that gold serves as a store and source of value to guard against inflation, and the safe haven, hedging, and diversification properties of gold are considered by investors during times of economic turmoil (Lucey and Li, 2015; O'Connor et al., 2015; Fang et al., 2018; Vidal & Kristjanpoller, 2020). Events such as 'Black Monday' in 1987, the global financial crisis of 2008, the stock market crash of 2015, among others, all validate gold as a highly sought-after commodity for reserves when the economic coasts are not clear for major economies (Vidal & Kristjanpoller, 2020).

The USNP, a measure of the total non-farm employment for a given month in the US, covers workers but excludes proprietors and private employees as well as unincorporated self-employed individuals, unpaid volunteers, and farm employees, which accounts for about 80 percent of the workers who contribute to the US GDP and provides an insight into economic realities as it represents the number of jobs gained or lost in the US economy[†]. Increases in USNP could mean that businesses are booming and hiring, suggesting an overall improvement in the economy. Accordingly, we rely on money demand theory to hypothesize that when the market is characterized by a 'risk-on'[‡] sentiment, a positive shock to USNP would be expected to impact gold market volatility negatively. This is because lower (higher) incomes via job losses (gains) would lead to less (more) trading in the conventional financial market, while the reverse is true for gold being a strategic commodity with hedging and safe haven potential. For example, during economic downtowns, of which UNSP serves as a prominent barometer, investors naturally seek alternative ways of securing their investment, for which the gold market comes in handy (Kayal & Maheswaran, 2021). Hence, we expect gold and stock markets to respond oppositely to USNP information. Put differently, our hypothesis is that significant job losses (gains) impact future cashflows negatively (positively) which consequently lowers (raises) stock trading while the reverse is true for gold given its hedging and safe haven properties. Therefore, we expect to see a negative relationship between USNP and gold return volatility.

We offer the following contributions to the academic literature. Firstly, the study appears to be the first to examine the predictive value of USNP for gold market volatility. Indeed, there is a plethora of studies in the literature that consider the volatility of precious metal prices in the face of a number of economic fundamentals and models (see, for example, Christie-David et al., 2000; Cai et al., 2001; Batten & Lucey, 2010; Elder et al., 2012; Trück & Liang, 2012; Fang et al., 2018; Vidal & Kristjanpoller, 2020). With the exception of Trück & Liang (2012) and Vidal & Kristjanpoller (2020) who attempt to specifically forecast gold market volatility using various statistical models (although without considering USNP), these studies, including Christie-David et al. (2000), Elder et al. (2012) among others, use gold price data rather than volatility despite the latter often being considered in investment decisions as it is a good measure of risk, the magnitude of which takes prominence in the valuation and diversification of assets in line with standard theories of finance such as the capital asset pricing model and arbitrage pricing theory, among others.

Secondly, we extend our analysis to cover the out-of-sample predictive power of USNP for gold market volatility as this gives a higher level of reliability and confidence in the results than

[†] https://fred.stlouisfed.org/series/PAYEMS
[‡] See https://www.babypips.com/forexpedia/risk-sentiment#

relying on in-sample predictability only (Rapach & Zhou, 2013). This consideration provides useful information to investors in terms of how movements in USNP can affect the future volatility of the gold market, making them better informed in making investment decisions.

Finally, we employ the GARCH-MIDAS approach of Engel, Ghysels, & Sohn (2013) which allows for the use of series in their available 'natural' form rather than restricting our analysis to a uniform frequency, when the variables of interest (the predicted and predictor series) are available at different frequencies. Since gold return is at a higher frequency than USNP, the GARCH-MIDAS model is considered a veritable choice as it can solve the problem of mismatched frequencies of daily returns and monthly macroeconomic variables and model long-term volatility (Fang et al., 2018).

Our predictability results show that, true to our hypothesis, USNP not only predicts the volatility of gold returns but also has a negative relationship with it. By implication, this shows that as USNP levels fall (rise), the volume of transactions being made in the gold market significantly improves (deteriorates) which consequently raises (lowers) its volatility. More explicitly, as activity increases in the labour market, indicating an increased level of economic activity in the real economy, investment in the financial sector, especially the stock market, correspondingly increases leading to a decrease in volume of transactions in the gold market. The forecast evaluation results suggest that the proposed model which includes USNP outperforms the benchmark which ignores it, and this result holds across all the forecast horizons considered. Some additional results are rendered for other precious metals for robustness purposes, and the outcomes validate the negative association between USNP and the volatility of precious metals.

The remainder of the paper is organized as follows. Section 2 describes the methodology and Section 3 presents the data and preliminary analyses. Section 4 presents and discusses the results. Section 5 concludes.

2. Methodology

To estimate the in-sample predictive power of USNP for the volatility of gold returns, we employ the GARCH-MIDAS framework. In the previous section, we articulate the merit of this method which permits mixed data frequencies, removing the inherent restriction of having to use the same (low) frequency for the variables of interest. Customarily, when variables are available at varying frequencies, in our case USNP (low, i.e. monthly) and gold returns (high, i.e. daily), the analysis is restricted to the low frequency often leading to information loss or biased outcomes. By using the GARCH-MIDAS framework we accommodate the high and low frequencies of the two data series in order to ensure that greater variability and more robust information is captured in the estimation process with greater potential for improved forecast outcomes.

Given a daily gold return series computed as - $r_{i,t} = 100 * ln\left(P_{i,t}\right) - ln\left(P_{i-1,t}\right)$, where $P_{i,t}$ represents the closing price for day i in month t with $t = 1,\dots,T$ and $i = 1,\dots,N_t$ denoting monthly and daily frequencies, respectively, and N_t is the number of days in a given month t, we construct a GARCH-MIDAS-X model where the monthly USNP (in returns) serves as a predictor. Essentially, there are two components of the mean and conditional variance equations, while the latter is further divided into short- and long-run components to accommodate the predictor series.

$$r_{i,t} = \mu + \sqrt{\tau_t \times h_{i,t}} \times \varepsilon_{i,t}, \quad \varepsilon_{i,t} \left| \Phi_{i-1,t} \sim N\left(0,1\right), \quad \forall \; i = 1,\dots,N_t \right. \tag{1}$$

$$h_{i,t} = \left(1 - \alpha - \beta\right) + \alpha \frac{\left(r_{i-1,t} - \mu\right)^2}{\tau_i} + \beta h_{i-1,t} \tag{2}$$

$$\tau_i^{(rw)} = m^{(rw)} + \theta^{(rw)} \sum_{k=1}^{K} \phi_k \left(w_1, w_2\right) X_{i-k}^{(rw)} \tag{3}$$

Equation (1) defines the mean equation while equations (2) and (3) are the conditional variance components specified for short- and long-run components, respectively. The parameter μ is the unconditional mean of the return series as specified in equation (1); $h_{i,t}$ is the short-run component at a high frequency which, as specified in equation (2), follows the $\mathrm{GARCH}\left(1,1\right)$ process, where α and β are the ARCH and GARCH terms, respectively, conditioned to be positive and/or at least zero ($\alpha > 0$ and $\beta \geq 0$) and sum to less than unity $\left(\alpha + \beta < 1\right)$; τ_t captures the long-run component which incorporates the exogenous macroeconomic series (or realized volatility where there is no macroeconomic series), and involves repeating the monthly value throughout the days in that month. The superscript (rw) in equation (3) denotes the implementation of a rolling window framework (which allows the secular long-run component to vary daily), while m represents the long-run component intercept. The focus of our analysis is the MIDAS slope coefficient (θ) that indicates the predictive power of the incorporated exogenous predictor X_{i-k} where $\phi_k\left(w_1, w_2\right) \geq 0,\;\; k = 1,\ldots,K$ is the weighting scheme that must sum to one for the parameters of the model to be identified; and K is chosen based on the log-likelihood statistic for each pair of the predicted and predictor series in order to filter the secular component of the MIDAS weights.

For the out-of-sample forecast performance evaluation, we compare the forecasts of our proposed GARCH-MIDAS predictive model (involving USNP), i.e. GARCH-MIDAS-X, with that of the conventional GARCH-MIDAS specifications that include realized volatility (GARCH-MIDAS-RV). The out-of-sample forecast performance is evaluated for forecast horizons which correspond to short- and long-run predictability (*h* = 30, 60, 90). Given that the contending models are not nested, we employ the modified Diebold and Mariano (1995) test, as per Harvey, Leybourne, and Newbold (1997), which calculates the p-value and addresses the issue with an assumption of zero covariance at unobserved lags, to formally ascertain whether the forecast errors associated with the contending models differ significantly. The test statistic is formulated as:

$$DM\ Stat = \frac{\bar{d}}{\sqrt{V(d)/T}} \sim N(0,1) \tag{6}$$

where $\bar{d} = \frac{1}{T}\sum_{t=1}^{T} d_t$ is the mean of the loss differential $d_t \equiv l\left(\varepsilon_{xt}\right) - l\left(\varepsilon_{rvt}\right)$; $l\left(\varepsilon_{xt}\right)$ and $l\left(\varepsilon_{rvt}\right)$ are the loss functions of the forecast errors (ε_{xt} and ε_{rvt} , respectively) associated with the GARCH-MIDAS-X and GARCH-MIDAS-RV, respectively; and $V\left(d_t\right)$ is the unconditional variance of the loss differential d_t. The null hypothesis of relative equality of the forecast precision of the contending model pairs is tested by examining $E\left[d_t\right] = 0$; with statistical significance implying a statistically significant difference in the forecast precision of the contending model pairs. Based on the p-value of Harvey, Leybourne, and Newbold (1997) for the DM statistic, a statistically significant negative DM statistic implies the adoption of the GARCH-MIDAS-X model while the benchmark (GARCH-MIDAS-RV) model is chosen if the test statistic is positive and significant. However, if the test statistic is not significant (implying a non-rejection of the null hypothesis), the forecast performance of the two competing models is assumed to be identical.

3. Data and Preliminary Analysis

We use the daily closing prices of gold, platinum, palladium, rhodium, and silver, expressed in US dollars, as well as the daily US stock market index represented by the S&P 500 index. These daily data are extracted from DataStream of Thomson Reuters. We collect monthly data on the USNP from https://fred.stlouisfed.org/. The starting point of each data series is given in Table 1, as dictated by the availability of data, whereas the ending point of the sample period is common (30 April 2021).

Table 1: Summary Statistics

	Mean	Std. Dev	Skewness	Kurtosis	CV	Start date	End date
USNP	0.1203	0.647	-18.716	438.343	5.3762	1/4/1968	4/30/2021
Gold	0.0282	1.2274	0.1215	30.5506	43.568	1/4/1968	4/30/2021
Palladium	0.036	1.9633	-0.2613	11.0982	54.4788	1/5/1987	4/30/2021
Platinum	0.018	1.6073	-0.4755	12.614	89.4675	1/2/1976	4/30/2021
Rhodium	0.0322	1.8693	0.3846	41.4838	58.1262	7/1/1992	4/30/2021
Silver	0.0178	2.1358	-0.1036	20.365	119.953	1/3/1968	4/30/2021
US stocks	0.0539	1.6564	-0.1692	10.9302	30.7543	1/10/1985	4/30/2021

Notes: Std. Dev. is the standard deviation. CV represents coefficient of variation, computed as standard deviation/mean. All values are expressed in returns. Returns is computed as 100*log (price/ price (-1)).

Table 1 gives the summary statistics for changes in USNP, gold, US stocks and other precious metals. The table summarizes the mean, standard deviation, skewness, kurtosis and coefficient of variation (CV) for all the variables highlighted. The reported mean values represent the average returns value for each of the variables represented. It is evident from the table that, given the time period considered, all variables have positive returns. Despite cycles of economic crises, US employment levels increase on average although they are highly volatile. Of the precious metals (gold inclusive), palladium and rhodium have the highest returns while platinum and silver have the least returns. Gold shows a modest average positive return over time. However, the standard deviation results indicate that the prices of precious metals are highly volatile, reflecting instability in macroeconomic and financial environments. USNP is negatively skewed while gold is positively skewed, although both are leptokurtic. All other series except rhodium show negative skewness. The empirical literature identifies some similar characteristics of precious metal returns, such as fat tails and asymmetry (Naeem et al., 2019). This outcome justifies our choice of the GARCH-MIDAS approach, since standard GARCH models may produce biased results for skewed data (Franses and Van Dijk, 1996). Meanwhile, the CV results show that all the variables are sparsely clustered around their mean, except USNP which has a dense cluster. This shows that although employment may increase on average, levels of employment, in terms of losses and gains, change weakly over time.

Table 2 provides some preliminary analyses, including conditional heteroscedasticity, autocorrelation, and higher order autocorrelation tests at lags 5, 10 and 20, along with the autoregressive conditional heteroscedasticity (ARCH) test, Q-statistics and Q2-statistics. The results show that UNSP does not exhibit ARCH effects or higher order correlation. In fact, serial correlation is observed only at higher lags 10 and 20, probably due to its relatively low frequency. However, on the other hand, the remaining return series, especially gold, exhibit ARCH effects, serial correlation and higher order correlation at all the specified lags. Previous studies make similar findings; for example, Arouri et al. (2012) examine 4 precious metals – gold, silver, platinum and palladium – and find strong evidence for long-range dependence in the conditional returns and volatility processes. This feature is attributable to the high frequency nature of the data. Therefore, given the presence of ARCH effects, serial correlation, and the mixed frequency of the data (daily and monthly), the GARCH-MIDAS model is the most suitable approach.

Table 2: Preliminary Analysis

	Arch (5)	Arch (10)	Arch (20)	Q (5)	Q (10)	Q (20)	Q2 (5)	Q2 (10)	Q2 (20)
USNP	0.42	0.21	0.12	12.20[b]	16.66[c]	17.31	2.17	2.18	2.19
Gold	146.57[a]	80.63[a]	51.53[a]	22.79[a]	32.98[a]	77.44[a]	801.97[a]	1093.60[a]	1725.50[a]
Palladium	142.09[a]	87.12[a]	46.62[a]	39.04[a]	39.90[a]	61.73[a]	1053.20[a]	1716.30[a]	2215.50[a]
Platinum	249.52[a]	159.51[a]	83.76[a]	10.87[b]	22.82[a]	44.77[a]	1927.90[a]	3515.50[a]	5014.80[a]
Rhodium	479.68[a]	268.67[a]	143.50[a]	2528.40[a]	2709.50[a]	2862.40[a]	2882.10[a]	4379.00[a]	4960.80[a]
Silver	559.08[a]	307.86[a]	160.37[a]	123.65[a]	138.78[a]	152.69[a]	4532.60[a]	6728.60[a]	8641.60[a]
US stocks	378.64[a]	218.79[a]	114.34[a]	28.73[a]	54.27[a]	121.26[a]	3135.80[a]	5189.60[a]	7403.80[a]

Notes: The Q(k) and $Q^2(k)$ statistics are obtained from the Ljung-Box test for serial correlation using the residuals and squared residuals, respectively, of the test regressions where k=5, 10, 20. ARCH (k) is the F-statistic of the ARCH-LM test used to test for conditional heteroscedasticity. 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. [a], [b] and [c] indicate statistical significance at 1%, 5% and 10%, respectively.

4. Main findings

4.1 Predictability of gold return volatility

In this section, we present the results for the predictive power of USNP for gold market volatility and the forecast performance of the GARCH-MIDAS-X model in comparison with a benchmark model which is the conventional GARCH-MIDAS model with realized volatility (GARCH-MIDAS-RV). In Table 3, our predictability results show that, true to our hypothesis, USNP not only predicts the volatility of gold returns but also has a negative relationship with it. This is implied by the negatively significant value of our result, as shown by the slope coefficient (θ). Consequently, as US employment levels rise, the volume of transactions being made in the gold market significantly reduces. More explicitly, as activity increases in the labour market, indicating an increased level of economic activity in the real economy, investment in the stock market correspondingly increases leading to a decrease in the volume of transactions in the gold market. This is entirely made possible by investors' perceptions of gold as a hedge or safe haven asset during periods of high turbulence or great uncertainty. Hence, an increase in trade in the stock market leads to a decrease in trade in the gold market. This result is amplified by recent developments in the gold market. For instance, gold surged past the $1900 mark to hit a 5-month high which caused an increase in volume of trade not seen since the 2008 financial crisis (FXSTREET, 2021). Since the outbreak of COVID-19 pandemic - a period characterized by loss of jobs, a crash in the financial market, and an economic recession - the gold price has reportedly been on the increase. All these events lend credence to our hypothesis that USNP and volatility in gold returns are negatively related. Subsequently, we present here a discussion of other important parameter estimates such as the unconditional mean stock returns (μ), the ARCH term (α), the GARCH term (β), the adjusted beta polynomial weight (ω) and the long-run constant term (m). The sum of the ARCH and GARCH terms accounts for the impact and persistence of shock to the gold market. Therefore, from our result, we find that the impact of any shock to the gold market is temporary although the shock effect may persist for a while given that the sum of the ARCH (α) and GARCH (β) terms is close to unity.

We also conduct out-of-sample forecast analysis by evaluating the relative forecast performance of the two competing GARCH-MIDAS models, that is, the GARCH-MIDAS with USNP and the conventional variant with realized volatility (which excludes the USNP predictor). Based on the modified Diebold and Mariano (1995) test, we find that our proposed model which accounts for the USNP data is consistently favoured for all forecast horizons. This result indicates that using the GARCH-MIDAX-X model may provide accurate out-of-sample predictability for volatility of gold returns, which is of particular importance to investors and forecast analysts who take a special interest in knowing the future behaviour of precious metal markets in order to make appropriate investment decisions. Naeem et al. (2019) model the volatility of precious metal markets and find that the regime-switching GARCH models outperform the single–regime GARCH specifications in predicting value-at-risk.

Table 3: In-sample and out-of-sample predictability of US employment rate for gold volatility

Gold	μ	α	β	θ	w	m
In-Sample	4.89E-05	0.0206***	0.9777***	-0.0026***	14.394***	0.0001***
	(7.17E-05)	(0.0002)	(0.0001)	(0.0004)	(3.5206)	(8.61E-06)
Out-of-Sample	h=5	h=10	h=20			
	-4.9766 [0.0000]	-4.2180 [0.0000]	-4.2649 [0.000]			

Notes: μ - unconditional mean of stock price returns, α- ARCH term, β - GARCH term, θ - slope coefficient, w - the adjusted beta polynomial weight, and m - long-run constant term. The figures in parenthesis are the standard errors of the parameter estimates, while ***, ** and * indicate statistical significance at 1%, 5% and 10%, respectively. The modified Diebold and Mariano test, as per Harvey, Leybourne, and Newbold (1997), calculates the p-value and addresses the issue with the assumption of zero covariance at unobserved lags. We report both the test statistics and the corresponding p-values reported in square brackets – [ ]. If the statistic is negative and significant, the GARCH-MIDAS-X is favoured while the GARCH-MIDAS-RV is chosen if the test statistic is positive and significant. However, if the test statistic is not significant (implying a non-rejection of the null hypothesis), the forecast performance of the two competing models is assumed to be identical.

4.2 Additional results

4.2.1 Predictability of return volatilities of other precious metals

In this section, we consider the relationship between USNP and other precious metals (palladium, platinum, rhodium, and silver) to ascertain whether the results are market sensitive. The intuition here is to see whether other precious metals respond differently to changes in USNP. The results presented in Table 4 show that USNP has a negative and significant relationship with all other precious metals, with the highest magnitude for silver and rhodium. This reinforces our earlier findings for gold and shows that not only gold but other precious metals have returns volatility predictable by USNP. Specifically, an increase in USNP leads to a decrease in return volatility among the precious metals. Previous studies provide results consistent with these findings. For example, Sensoy (2013), using dynamic conditional correlations, shows that precious metals were strongly correlated with each other over the previous decade, implying that volatilities in the precious metal markets always behave alike. The study reveals that gold especially has a unidirectional volatility contagion effect on all other precious metals. As shown in Table 4, the results for other parameters such as ARCH and GARCH terms for each precious metal reveal that shocks to these markets are temporary and volatility persistence is high and mean-reverting.

The forecast evaluation results show that, of the four precious metals considered, the proposed model outperforms the benchmark for only two (rhodium and silver) for all forecast horizons, with the proposed model being favoured only at the 5-day horizon for platinum. However, the proposed model is most prominently preferred for silver, given its significance value. This is very plausible, owing to the close relationship often observed between the demands for gold and silver. Palladium, on the other hand, shows no preference for the proposed model at any of the forecast horizons.

Table 4: In-sample and out-of-sample predictive power of US employment rate for precious metal volatility

In-Sample	μ	α	β	θ	w	m
Palladium	0.0004***	0.0973***	0.8811***	-0.0225***	3.3435***	0.0004***
	(0.0001)	(0.0032)	(0.0035)	(0.0023)	(0.2554)	(2.76E-05)
Platinum	1.46E-05	0.0561***	0.9413***	-0.0087***	7.5151***	0.0005***
	(0.0001)	(0.0017)	(0.0017)	(0.0026)	(2.5822)	(7.27E-05)
Rhodium	-0.0005***	0.2243***	0.7750***	-0.6240*	2.7775***	0.0125*
	(0.0001)	(0.0022)	(0.0021)	(0.3722)	(0.1547)	(0.0074)
Silver	-0.0001	0.0721***	0.9266***	-0.9687***	2.3481***	0.0031***
	(0.0001)	(0.0021)	(0.0022)	(0.3808)	(0.6020)	(0.0010)
Out-of-Sample	h=5	h=10	h=20			
Palladium	-0.9657 [0.3342]	-0.8416 [0.4000]	-0.7494 [0.4536]			
Platinum	-1.8256 [0.0679]	-1.3516 [0.1765]	-1.0399 [0.2984]			
Rhodium	-4.9699 [0.0000]	-3.4400 [0.0005]	-2.4263 [0.0153]			
Silver	-21.3757 [0.0000]	-16.75455 [0.0000]	-14.2600 [0.0000]			

Notes: μ - unconditional mean of stock price returns, α- ARCH term, β - GARCH term, θ - slope coefficient, w - the adjusted beta polynomial weight, and m - long-run constant term. The figures in parenthesis are the standard errors of the parameter estimates, while ***, ** and * indicate statistical significance at 1%, 5% and 10%, respectively. The modified Diebold and Mariano test, as per Harvey, Leybourne, and Newbold (1997), calculates the p-value and addresses the issue with the assumption of zero covariance at unobserved lags. Thus, we report both the test statistics and the corresponding p-values reported in square brackets – [ ]. If the statistic is negative and significant, the GARCH-MIDAS-X is favoured while the GARCH-MIDAS-RV is chosen if the test statistic is positive and significant. However, if the test statistic is not significant (implying a non-rejection of the null hypothesis), the forecast performance of the two competing models is assumed to be identical.

4.2.2 Predictability of volatility of US stock returns

For robustness, we extend our analysis to include the predictive power of USNP for US stock returns (Table 5). Our results also conform to our hypothesis that an increase in UNSP leads to an increase in US stock returns volatility. The slope coefficient (θ) is positive and significant, implying that UNSP has both predictive content for the volatility of US stock returns and a positive relationship. Other parametric estimates such as ARCH and GARCH terms and adjusted beta weights are similar to those of gold returns volatility. Hence, this result validates our claim that USNP has a predictive power for gold and other precious metals as well as for US stocks. This conforms with the findings of Mensi et al. (2013) who, using a vector autoregressive GARCH (VAR-GARCH) model, show significant spillovers in terms of shock and volatility between the S&P 500 stock returns and spot commodity market returns.

Table 5: In-sample and out-of-sample predictive power of US employment rate for US stock volatility

US stock return	μ	α	β	θ	w	m
In-Sample	0.0008^{***} (0.0001)	0.1025^{***} (0.0064)	0.8875^{***} (0.0066)	0.114^{***} (0.0460)	1.0052^{***} (0.2347)	0.0001^{***} (0.0000)
Out-of-Sample	h=5	h=10	h=20			
	-0.1849 [0.8532]	-0.1473 [0.8829]	-0.1305 [0.8962]			

Note: μ - unconditional mean of stock price returns, α- ARCH term, β - GARCH term, θ - slope coefficient, w - the adjusted beta polynomial weight, and m - long-run constant term. The figures in parenthesis are the standard errors of the parameter estimates, while ***, ** and * indicate statistical significance at 1%, 5% and 10%, respectively. The modified Diebold and Mariano test as per Harvey, Leybourne, and Newbold (1997), calculates the p-value and addresses the issue with the assumption of zero covariance at unobserved lags. Thus, we report both the test statistics and the corresponding p-values reported in square brackets – [ ]. If the statistic is negative and significant, the GARCH-MIDAS-X is favoured while the GARCH-MIDAS-RV is chosen if the test statistic is positive and significant. However, if the test statistic is not significant (implying a non-rejection of the null hypothesis), the forecast performance of the two competing models is assumed to be identical.

5. Conclusion

In this study, we examine the ability of US Nonfarm Payroll (USNP) to predict the volatility of gold returns, while testing the hypothesis that significant job gains (losses) in the US impact negatively (positively) the volatility of gold returns, given the safe haven property of gold. Using the GARCH-MIDAS approach, which allows for mixed data frequency and circumvents information loss or any associated bias, we analyse monthly USNP data and daily gold returns data. Hence, we construct a GARCH-MIDAS-X model where the USNP (in returns form) serves as a predictor. The focus of our analysis is the MIDAS slope coefficient (θ) which indicates the predictive power of the incorporated exogenous predictor. For the out-of-sample forecast evaluation, we compare the forecasts of our proposed GARCH-MIDAS predictive model (involving USNP) to that of the conventional GARCH-MIDAS specifications which include realized volatility (GARCH-MIDAS-RV).

True to our a-priori assumption, the main result validates our hypothesis. The slope coefficient (θ) is negatively significant, implying an inverse relationship between the two variables (USNP and volatility of gold returns). In addition, our result is a proof of the predictive power of USNP for gold returns. Further analysis involving forecast evaluation confirms the preference of our proposed model to the benchmark.

For robustness, we extend our analysis to include other precious metals such as palladium, platinum, rhodium and silver, and the results reveal that USNP has predictive content for all precious metals considered and its relationship with precious metal returns is negative. This confirms our earlier result for gold and affirms our hypothesis that USNP has a negative relationship with gold and other precious metals. Meanwhile, the forecast evaluation results indicate that only two of the other precious metals (rhodium and silver) favour our proposed model across all three horizons. We also evaluate the predictive power of USNP for volatility of US stock returns, and the results show a positive relationship. This connotes that as US employment levels rise, the volume of trade in the stock market also rises. The forecast evaluation shows that the benchmark model is favoured over the proposed model.

The findings of this study have implications for financial analysts and investors. Firstly, the government should monitor what is happening in USNP to predict gold volatility. Financial analysts and investors can build on our models and findings to become more informed about when to devest in one market and invest in the other by observing non-farm employment levels, leading to a refinement of their predictive models, with the aim of maximizing their decision making in precious metal markets. This is important, given that refining the predictability of the volatility of precious metals matters to asset pricing, portfolio analysis, option pricing, and trading strategies involving the volatility of gold and other precious metals.

An interesting area for further research consists of examining the relationship between USNP and industrial metals or non-precious metals such as copper, steel, and iron ore. Furthermore, it would be interesting to examine whether our findings prove to be generalizable to other economic releases by the US, such as the trade deficit or consumer price index. Future studies could examine the predictive power of the surprise in the USNP figure, as measured by the difference between the actual USNP figure and the median analyst forecast.

References

Alguire, M. S. (1996). Nonfarm payroll jobs in Arkansas and the nation continue to increase. *Arkansas Business and Economic Review*, *29*(1), 26.

Arouri, M. E. H., Hammoudeh, S., Lahiani, A., & Nguyen, D. K. (2012). Long memory and structural breaks in modeling the return and volatility dynamics of precious metals. *The Quarterly Review of Economics and Finance*, *52*(2), 207-218.

Batten, J. A., & Lucey, B. M. (2010). Volatility in the gold futures market. *Applied Economics Letters*, *17*(2), 187-190.

Bhatia, T. (2020). Predicting Non Farm Employment. *arXiv preprint arXiv:2009.14282*.

Cai, J., Cheung, Y. L., & Wong, M. C. (2001). What moves the gold market?. *Journal of Futures Markets: Futures, Options, and Other Derivative Products*, *21*(3), 257-278.

Caruso, A. (2019). Macroeconomic news and market reaction: Surprise indexes meet nowcasting. *International Journal of Forecasting*, *35*(4), 1725-1734.

Christie–David, R., Chaudhry, M., & Koch, T. W. (2000). Do macroeconomics news releases affect gold and silver prices?. *Journal of Economics and Business*, *52*(5), 405-421.

Diebold, F.X. & Mariano, R.S. (1995). Comparing Predictive Accuracy. *Journal of Business and Economic Statistics*, 13, 253–263.

Elder, J., Miao, H., & Ramchander, S. (2012). Impact of macroeconomic news on metal futures. *Journal of Banking & Finance*, *36*(1), 51-65.

Ederington, L., Guan, W., & Yang, L. Z. (2019). The impact of the US employment report on exchange rates. *Journal of International Money and Finance*, 90, 257-267.

Engle, R. F., Ghysels, E., & Sohn, B. (2013). Stock market volatility and macroeconomic fundamentals. *Review of Economics and Statistics*, *95*(3), 776-797.

Fang, L., Yu, H., & Xiao, W. (2018). Forecasting gold futures market volatility using macroeconomic variables in the United States. *Economic Modelling*, *72*, 249-259.

Franses, P.H., Van Dijk, D., 1996. Forecasting stock market volatility using (non□linear) Garch models. J. Forecast. 15, 229–235.

Harvey, D., Leybourne, S. & Newbold, P. (1997). Testing the equality of prediction mean squared errors. *International Journal of Forecasting*, 13 (2), 281-291.

Kayal, P., & Maheswaran, S. (2021). A study of excess volatility of gold and silver. *IIMB Management Review*.

Lucey, B. M., & Li, S. (2015). What precious metals act as safe havens, and when? Some US evidence. *Applied Economics Letters*, *22*(1), 35-45.

Mensi, W., Beljid, M., Boubaker, A., & Managi, S. (2013). Correlations and volatility spillovers across commodity and stock markets: Linking energies, food, and gold. *Economic Modelling*, *32*, 15-22.

Naeem, M., Tiwari, A. K., Mubashra, S., & Shahbaz, M. (2019). Modeling volatility of precious metals markets by using regime-switching GARCH models. *Resources Policy*, *64*, 101497.

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.

Rapach, D., & Zhou, G. (2013). Forecasting stock returns. In Handbook of economic forecasting, 2, 328-383.

Sensoy, A. (2013). Dynamic relationship between precious metals. *Resources Policy*, *38*(4), 504-511.

Trück, S., & Liang, K. (2012). Modelling and forecasting volatility in the gold market. *International Journal of Banking and Finance*, *9*(1), 48-80.

Vidal, A., & Kristjanpoller, W. (2020). Gold volatility prediction using a CNN-LSTM approach. *Expert Systems with Applications*, *157*, 113481.