Obtaining Accurate Gold Prices

Amit K. Sinha

Department of Economics and Finance, Bradley University, Peoria, IL 61625, USA; asinha@fsmail.bradley.edu; Tel.: +1-309-677-3582

Abstract: Gold prices have been of major interest for a lot of investors, analysts, and economists. Accordingly, a number of different modeling approaches have been used to forecast gold prices. In this manuscript, the geometric Brownian motion approach, used in the pricing of numerous types of assets, is used to forecast the prices of gold at yearly, monthly, and quarterly frequencies. This approach allows for simulating one-period-ahead prices and the associated probabilities. The expected prices obtained from the simulated prices and probabilities are found to provide reliable forecasts when compared with the observed yearly, monthly, and quarterly prices.

Keywords: geometric Brownian motion; forecasting; commodities; prices; forecasting; gold; simulation; probabilities; asset pricing

JEL Classification: C0; C6; G0; G12; G170

1. Introduction

Interest in gold objects goes back thousands of years [1–3], although, recently, gold has become a tradable investment product which can serve as an inflation hedge [4–6]. As gold prices can serve as an indicator of apparent predicaments [4], they attract a lot of attention [7] from gold miners, hedgers, investors, traders, and speculators. Accurate gold prices would help tremendously in decision making. Accordingly, obtaining accurate gold prices is of paramount importance. A number of researchers have attempted to model gold prices, such as Refs. [8–10], among others, differing in their approaches. One approach to obtain accurate gold prices is the geometric Brownian motion (GBM) approach. Refs. [11–13] used GBM to model gold prices, but this current manuscript differs from them in terms of its methodological approach to using GBM, the sample period, and the data frequency utilized to obtain accurate gold prices.

Ref. [11] applied the GBM methodology to estimate gold prices for the year 2022, with the purpose of testing if it could be applied as a forecasting model during unanticipated circumstances, such as a pandemic, specifically the COVID-19 pandemic. After having obtained gold prices from Yahoo Finance, they showed the GBM modeling process to be useful even in unanticipated and unforeseen circumstances. Ref. [12], on the other hand, applied the GBM modeling process to test its effectiveness in forecasting prices of derivative contracts of gold. They used ten years of closing prices of gold derivatives obtained from the website https://www.investing.com/ (accessed on 2 March 2024), for the period between December 2012 and December 2022. Just like Ref. [11], the mean absolute percentage error tests in Ref. [12] showed that the GBM forecasts had low rates of errors. Ref. [13] used GBM to obtain the price paths of Kijan Emas, the Malaysian gold bullion coin, for the year 2016, between the fourth of January and December thirtieth, from Bank Negara Malaysia’s website. While their results show that their methodology produced accurate forecasts for a short period of time, they also mention that forecasts could be inaccurate if intended for a window of time beyond one month. This manuscript adds to the literature of modeling gold prices and obtaining accurate gold prices using the GBM procedure. A twenty-year-long rolling window is hereby used to obtain the components necessary for the modeling process.
to implement the GBM model. The model is simulated a hundred thousand times, and, from these simulations, the expected prices are obtained from the addition of the product of simulated prices and the multiplication with their associated probability. This results in accurate and reliable gold prices. The process is replicated at yearly, quarterly, and monthly frequencies.

GBM has been applied to forecast asset prices. Ref. [14] is widely credited as the first work modeling stochastic asset prices using Brownian motion concepts, but, now, GBM has been applied to model the prices of a number of different types of assets like stocks [15–17], stock indices [18], derivatives [12,19], energy prices [20,21], agricultural products [22,23], and iron ore [24]. Ref. [25] modeled international interest rate term structures to understand the movements of currency exchange rates. Ref. [26] modeled EUR exchange rates against the USD, the SAR, and the AUD using the GBM method. Refs. [11–13] model gold prices using GBM. This article contributes to our understanding of the modeling of gold prices by providing a GBM approach that provides reliable forecasts.

\[
\frac{dP}{P} = \mu dt + \sigma dW
\]  

Equation (1)

Typically, the GBM process is a stochastic process mathematically represented in Equation (1). The process requires two terms and a random variable. The two terms usually proxied as the average return and the standard deviation are the drift and the diffusion terms. In Equation (1), they are represented as \( \mu \) and \( \sigma \), respectively. In this equation, \( dW \) is the random Weiner variable, which is normally distributed and has a mean of zero, and a value of one as its standard deviation. Also, in this equation, \( P \) represents the prices, while \( dP \) and \( dt \) represent changes in price and time, respectively.

\[
P_t = P_{t-1} e^{((\mu - \frac{1}{2} \sigma^2) \Delta t + \sigma \epsilon \sqrt{\Delta t})}
\]  

Equation (2)

Represented in Equation (2) is the closed-form solution of Equation (1). Refs. [12,18,27–29] also used this equation. In this equation, \( P_t \) and \( P_{t-1} \) are the prices of two adjacent periods, \( \mu \) and \( \sigma \) are the drift and diffusion terms, respectively, proxied by the averaged return and the standard deviation, \( \Delta t \) is the change in time, while \( \epsilon \) is a random variable with a mean of zero and a standard deviation of 1. Refs. [27,30,31] provide a good understanding of the GBM process. Generally, it is assumed that the drift and diffusion terms would remain constant, but Ref. [30] pointed out that this assumption could be weakened. Accordingly, Ref. [18] estimated the drift and diffusion terms using a twenty-year-long rolling window, essentially obtaining a unique set of drift and diffusion terms, at the start of every forecasting period. Moreover, by randomly altering the numerical values of \( \epsilon \), many different numerical values of \( P_t \) can be obtained. From the universe of simulated prices, associated probabilities for those values can be obtained, and expected prices can be estimated by summing up products of these simulated prices and their multiplication with the associated probabilities. The expected price obtained in such a way can be compared to the observed price at monthly, quarterly, and yearly frequencies. A similar approach to forecasting stock indices was used in Refs. [18,32,33], where results indicate these forecasts to be reliable. In the current manuscript, the methodology used in Refs. [18,32,33] is applied to gold prices, with similar results, in that the expected prices can provide reliable forecasts of the actual prices.

2. Data

The manuscript’s data were obtained from DataStream. The data consist of gold prices per troy ounce, obtained using the DataStream code GOLDHAR. Initially, the data were extracted at a daily frequency from the 29 December 1978 to the 27 December 2023. The data were converted into annual, quarterly, and monthly frequencies by considering the last price of a year, quarter, and month. During this period, the lowest price of $216.63

Commodities 2024, 3

116
per ounce was observed on the 15 January 1979, while the highest price of $2078.95 was
observed on the 17 December 2023.

\[ r_{a,q,m} = \ln \left( \frac{P_t}{P_{t-1}} \right) \]  

(3)

The annual, quarterly, and monthly returns were obtained by taking the natural
logarithms of two adjacent prices at the respective frequencies by applying the formula
shown in Equation (3). In this equation, \( r \) is the return, and \( a \) indicates the yearly
frequency, \( q \) the quarterly frequency, and \( m \) the monthly frequency. Using this return
(\( \bar{r}_{a,q,m} \)) and the standard deviations (\( \sigma_{a,q,m} \)) were calculated using Equations (4) and (5).

\[ \bar{r}_{a,q,m} = \frac{\sum_{n=1}^{n} r_{a,q,m}}{n} \]  

(4)

\[ \sigma_{a,q,m} = \sqrt{\frac{\sum_{n=1}^{n} (r_{a,q,m} - \bar{r}_{a,q,m})^2}{n-1}} \]  

(5)

Table 1 provides these calculations. The average returns were found to be 4.91%, 1.23%,
and 0.41%, at yearly, quarterly, and monthly frequencies. The standard deviations were
found to be 19.51%, 8.16%, and 5.09% at the yearly, quarterly, and monthly frequencies,
respectively.

<table>
<thead>
<tr>
<th>Table 1. Descriptive statistics.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Date</td>
</tr>
<tr>
<td>------------------------</td>
</tr>
<tr>
<td>First date</td>
</tr>
<tr>
<td>Last date</td>
</tr>
<tr>
<td>Lowest price</td>
</tr>
<tr>
<td>Highest price</td>
</tr>
<tr>
<td><strong>Annual</strong></td>
</tr>
<tr>
<td>Mean</td>
</tr>
<tr>
<td>Standard Deviation</td>
</tr>
<tr>
<td><strong>Quarterly</strong></td>
</tr>
<tr>
<td>Mean</td>
</tr>
<tr>
<td>Standard Deviation</td>
</tr>
<tr>
<td><strong>Monthly</strong></td>
</tr>
<tr>
<td>Mean</td>
</tr>
<tr>
<td>Standard Deviation</td>
</tr>
</tbody>
</table>

This table provides some descriptive information about the data. Gold prices were initially obtained daily.
From the daily prices, annual, quarterly, and monthly frequency data were obtained by taking the last trading
price at that frequency. The changes in prices were calculated by taking the natural logarithms of adjacent prices.
The average and standard deviation of those returns at annual, quarterly, and monthly frequencies are presented
in this table.

3. Methodology

The GBM framework used in this manuscript is similar to that used in Refs. [18,27–29].
Equation (6) represents the GBM framework used to simulate gold prices, which is similar
to Equation (2). \( P_{t,\text{sim}} \) is the simulated prices, while \( P_{t-1} \) is the last period’s actual closing
price, which also serves as the beginning price for the period in which the simulation is
carried out. \( \bar{r}_t \) and \( \sigma_t \) are the average return and the standard deviation, respectively, used
to proxy the drift and diffusion terms, calculated using the previous 20 years, 80 quarters,
and 240 months for the yearly, quarterly, and monthly bases, respectively. While there are a
number of ways in which the drift and diffusion terms can be estimated, using historical
averages and standard deviation as the proxies in the GBM framework has been applied in
a number of manuscripts [12,28,32–36].

\[ P_{t,\text{sim}} = P_{t-1} \times e^{\left( \bar{r}_t - \frac{1}{2} \sigma_t^2 \right) \Delta t + \sigma_t \epsilon \sqrt{\Delta t}} \]  

(6)

While applying the GBM framework to obtain reliable forecasts for the Dow Jones
Industrial Average, S&P 500, and the NASDAQ, Ref. [33] found that historical means and
standard deviation provided reliable forecasts for those indexes. Accordingly, historical means and standard deviation are used in this manuscript. These are represented in Equations (7) and (8), respectively.

\[
\mu_{a,q,m} = \frac{\sum_n r_{a,q,m}}{n} \quad n = 20, 80, 240 \tag{7}
\]

\[
\sigma_{a,q,m} = \sqrt{\frac{\sum_n (r_{a,q,m} - \mu_{a,q,m})^2}{n-1}} \quad n = 20, 80, 240 \tag{8}
\]

A rolling window approach allows for a new drift and diffusion term for each new period. The rolling window is similar to those used in Refs. [18,32,33]. \( \epsilon \) in Equation (6) is the random Weiner term, a standard normal variable which has a mean of zero and a standard deviation of one. One hundred thousand simulated values of \( P_t \) are obtained by arbitrarily changing the values of \( \epsilon \) in each period. For each simulated gold price, the return is estimated using Equation (9). The associated probability of each simulated return is estimated by calculating the difference between the cumulative density function of two adjacent simulated returns, assuming a normal distribution with the mean and standard deviation estimated using Equations (7) and (8), respectively.

\[
r_{\text{sim},i} = \ln \left( \frac{P_{\text{sim},i}}{P_{i-1}} \right) \tag{9}
\]

Refs. [18,32,33] also used this procedure to estimate the probabilities of the simulated returns. Using the simulated gold prices obtained from Equation (6) and the probabilities estimated with Equation (9), the expected price of gold is estimated using Equation (10), a procedure also used in Refs. [18,32,33].

\[
P_{t,\text{exp}} = \sum_{i=1}^{100,000} (P_{t,\text{sim},i}) \ast (\text{prob}_{t,\text{sim},i}) \tag{10}
\]

The expected price in the current study was estimated for each year, quarter, and month, separately. Given the data available, twenty-five annual expected prices were obtained, through the years from 1999 to 2023. Similarly, one hundred quarterly and three hundred monthly expected prices were obtained. The actual and the expected prices were plotted against the year, the year_quarter, and the year_month, at frequencies which are yearly, quarterly, and monthly. These plots are presented in Figures 1–3, respectively. The actual prices were also plotted against the expected prices at the three frequencies, and these plots are in Figures 4–6, respectively.

The efficacy of using the GBM simulated expected price to predict the actual gold prices was tested using a simple regression, shown in Equation (11). In this equation, \( P_{t,\text{actual}} \) is the actual gold price at the end of a period, while \( P_{t,\text{exp}} \) is the GBM-simulated expected price for that period. \( \alpha \) is the intercept, while \( \beta \) is the slope coefficient. The regression error term is represented by \( \epsilon \). In Equation (11), the natural logarithms of both the actual and the expected prices are used in the regressions. This equation can be used to test whether the expected price is a good reliable forecast of the actual price. Alternatively, it could also be used as a predictive model for the actual price using the expected price as the independent variable. Values of 0 for \( \alpha \) and one for \( \beta \) would indicate the usefulness of the expected price. A similar logic was used in Refs. [18,32,33] as well as by Ref. [37] who tested the persistence of mutual fund performance between two adjacent periods. The regression represented in Equation (11) was carried out at annual, quarterly, and monthly frequencies in our study, and the results are presented in Table 2.

\[
P_{t,\text{actual}} = \alpha + \beta \ast P_{t,\text{exp}} + \epsilon \tag{11}
\]
means and standard deviation are used in this manuscript. These are represented in Equations (7) and (8), respectively.

\[ \bar{\mu}, \bar{\sigma}, \bar{\theta} = \frac{\sum_{j} \theta \_j, \theta \_j}{\sum_{j} \theta \_j, \theta \_j} = n = 20, 80, 240 \ (7) \]

\[ \sigma, \sigma, \sigma = \frac{\sum_{j} (\theta \_j, \theta \_j, \theta \_j - \bar{\mu}, \bar{\mu}, \bar{\mu})^2}{\sum_{j} (\theta \_j, \theta \_j, \theta \_j - \bar{\mu}, \bar{\mu}, \bar{\mu})^2} = n = 20, 80, 240 \ (8) \]

A rolling window approach allows for a new drift and diffusion term for each new period. The rolling window is similar to those used in Refs. [18,32,33].

\[ \epsilon \text{ in Equation (6) is the random Weiner term, a standard normal variable which has a mean of zero and a standard deviation of one. One hundred thousand simulated values of } P_t \text{ are obtained by arbitrarily changing the values of } \epsilon \text{ in each period. For each simulated gold price, the return is estimated using Equation (9). The associated probability of each simulated return is estimated by calculating the difference between the cumulative density function of two adjacent simulated returns, assuming a normal distribution with the mean and standard deviation estimated using Equations (7) and (8), respectively.} \]

\[ r_{t+1} - r_t = \ln \left( \frac{P_t}{P_{t-1}} \right) \ (9) \]

Refs. [18,32,33] also used this procedure to estimate the probabilities of the simulated returns. Using the simulated gold prices obtained from Equation (6) and the probabilities estimated with Equation (9), the expected price of gold is estimated using Equation (10), a procedure also used in Refs. [18,32,33].

\[ P_t, \theta = \left[ P_t, \theta \_t \right] \ast P_t, \theta \_t = (10) \]

The expected price in the current study was estimated for each year, quarter, and month, separately. Given the data available, twenty-five annual expected prices were obtained, through the years from 1999 to 2023. Similarly, one hundred quarterly and three hundred monthly expected prices were obtained. The actual and the expected prices were plotted against the year, the year_quarter, and the year_month, at frequencies which are yearly, quarterly, and monthly. These plots are presented in Figures 1–3, respectively. The actual prices were also plotted against the expected prices at the three frequencies, and these plots are in Figures 4–6, respectively.

**Figure 1.** Expected and actual gold prices at an annual frequency. The figure plots the actual vs. the expected prices at an annual frequency. The prices on the y-axis are the actual and expected prices. The expected price is obtained using the manuscript’s methodological procedure.

**Figure 2.** Expected and actual gold prices at a quarterly frequency. The figure plots the actual vs. the expected prices at a quarterly frequency. The prices on the y-axis are the actual and expected prices. The expected price is obtained using the manuscript’s methodological procedure.

**Figure 3.** Expected and actual gold prices at a monthly frequency. The figure plots the actual vs. the expected prices at a monthly frequency. The prices on the y-axis are the actual and expected prices. The expected price is obtained using the manuscript’s methodological procedure.
Figure 2. Expected and actual gold prices at a quarterly frequency. The figure plots the actual vs. the expected prices at a quarterly frequency. The prices on the y-axis are the actual and expected prices. The expected price is obtained using the manuscript’s methodological procedure.

Figure 3. Expected and actual gold prices at a monthly frequency. The figure plots the actual vs. the expected prices at a monthly frequency. The prices on the y-axis are the actual and expected prices. The expected price is obtained using the manuscript’s methodological procedure.

Figure 4. Expected vs. actual gold prices at an annual frequency. This figure plots the natural logarithms of the observed prices (y-axis) against the natural logarithms of the expected prices (x-axis) estimated using the manuscript’s methodology at an annual frequency.

Figure 5. Expected vs. actual gold prices at a quarterly frequency. This figure plots the natural logarithms of the observed prices (y-axis) against the natural logarithms of the expected prices (x-axis) estimated using the manuscript’s methodology at a quarterly frequency.
Figure 4. Expected vs. actual gold prices at an annual frequency. This figure plots the natural logarithms of the observed prices (y-axis) against the natural logarithms of the expected prices (x-axis) estimated using the manuscript’s methodology at an annual frequency.

Figure 5. Expected vs. actual gold prices at a quarterly frequency. This figure plots the natural logarithms of the observed prices (y-axis) against the natural logarithms of the expected prices (x-axis) estimated using the manuscript’s methodology at a quarterly frequency.

Figure 6. Expected vs. actual gold prices at a monthly frequency. This figure plots the natural logarithms of the observed prices (y-axis) against the natural logarithms of the expected prices (x-axis) estimated using the manuscript’s methodology at a monthly frequency.

The efficacy of using the GBM simulated expected price to predict the actual gold prices was tested using a simple regression, shown in Equation (11). In this equation, $P_{t,\text{actual}}$ is the actual gold price at the end of a period, while $P_{t,\text{exp}}$ is the GBM-simulated expected price for that period. $\alpha$ is the intercept, while $\beta$ is the slope coefficient. The regression error term is represented by $\epsilon$.

In Equation (11), the natural logarithms of both the actual and the expected prices are used in the regressions. This equation can be used to test whether the expected price is a good reliable forecast of the actual price. Alternatively, it could also be used as a predictive model for the actual price using the expected price as the independent variable. Values of 0 for $\alpha$ and one for $\beta$ would indicate the usefulness of the expected price. A similar logic was used in Refs. [18,32,33] as well as by Ref. [37] who tested the persistence of mutual fund performance between two adjacent periods. The regression represented in Equation (11) was carried out at annual, quarterly, and monthly frequencies in our study, and the results are presented in Table 2.

Table 2. Regression results.

<table>
<thead>
<tr>
<th>Frequency</th>
<th>Intercept</th>
<th>Standard Error</th>
<th>t-stat (for integer = 0)</th>
<th>t-stat (for coefficient = 0)</th>
<th>t-stat (for coefficient = 1)</th>
<th>N</th>
<th>R-square</th>
<th>Adjusted R-Square</th>
</tr>
</thead>
<tbody>
<tr>
<td>Annual</td>
<td>0.30</td>
<td>0.0243</td>
<td>12.5</td>
<td>266.08</td>
<td>14.73</td>
<td>25</td>
<td>99.97%</td>
<td>99.97%</td>
</tr>
<tr>
<td>Quarterly</td>
<td>0.08</td>
<td>0.0024</td>
<td>32.35</td>
<td>2748.16</td>
<td>37.89</td>
<td>100</td>
<td>100.00%</td>
<td>100.00%</td>
</tr>
<tr>
<td>Monthly</td>
<td>0.03</td>
<td>0.0005</td>
<td>59.99</td>
<td>14,742.02</td>
<td>70.41</td>
<td>300</td>
<td>100.00%</td>
<td>100.00%</td>
</tr>
</tbody>
</table>
Table 2. Regression results.

<table>
<thead>
<tr>
<th>Frequency</th>
<th>Annual</th>
<th>Quarterly</th>
<th>Monthly</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>0.30</td>
<td>0.08</td>
<td>0.03</td>
</tr>
<tr>
<td>Standard Error</td>
<td>(0.0243)</td>
<td>(0.0024)</td>
<td>(0.0005)</td>
</tr>
<tr>
<td>t-stat (for integer = 0)</td>
<td>12.5</td>
<td>32.35</td>
<td>59.99</td>
</tr>
<tr>
<td>Expected Price</td>
<td>0.95</td>
<td>0.99</td>
<td>0.99</td>
</tr>
<tr>
<td>Standard Error</td>
<td>(0.0036)</td>
<td>(0.00036)</td>
<td>(0.00007)</td>
</tr>
<tr>
<td>t-stat (for coefficient = 0)</td>
<td>266.08</td>
<td>2748.16</td>
<td>14,742.02</td>
</tr>
<tr>
<td>t-stat (for coefficient = 1)</td>
<td>14.73</td>
<td>37.89</td>
<td>70.41</td>
</tr>
</tbody>
</table>

The dependent variable in these regressions is the actual price of the metal, while the independent variable is the expected price. The expected price is obtained using the GBM simulation-based methodology used in the manuscript. The regressions are carried out using the natural log values for both the actual and expected prices. The t-stat for the intercept is provided for the hypothesis that it is equal to zero. The t-stats for the hypotheses that the coefficient of the independent variable is equal to zero and, separately, equal to 1 are also provided. The standard errors of the coefficients are in parenthesis. Highlighted t-stats indicate significance at the 1% level.

The usefulness of the GBM-simulated expected gold prices as predictors of the actual prices was also tested to check if the means and variances between the two were different. The two price series—the natural logarithms of the observed prices and those of the expected prices—were first tested for differences in terms of variances. The difference in variance between the two series was hypothesized to be zero, and the F-stat was estimated using Equation (12). The averages of the natural logarithms of the observed and estimated prices were also tested using the two-sample t-stat estimated using pooled variances. The formulas for the pooled variances and the t-stats are represented in Equations (13) and (14), respectively. The formulas from Equations (12)-(14) can be found in most statistical textbooks, like Ref. [38]. A t-stat test was conducted assuming equal variances after the F-stat test for variances that did not reject the hypothesis of no differences in terms of variance. In Equations (13) and (14), \( n_1 = n_2 \), and the difference between the two means, \( \mu_1 - \mu_2 \), is hypothesized to not be different from zero. Table 3 presents the results of these tests.

\[
F_{\text{stat}} = \frac{\sigma^2_{\ln(P_{\exp})}}{\sigma^2_{\ln(P_{\actual})}}
\]

\[
\sigma^2_{\text{pooled}} = \frac{\sigma^2_{\ln(P_{\exp})} + \sigma^2_{\ln(P_{\actual})}}{(n_1 - 1) + (n_2 - 1)}
\]

\[
t_{\text{stat}} = \frac{\ln(P_{\exp}) - \ln(P_{\actual}) - (\mu_1 - \mu_2)}{\sqrt{\sigma^2_{\text{pooled}} \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}
\]

Table 3. Testing the means and variances.

<table>
<thead>
<tr>
<th>Frequency</th>
<th>Annual</th>
<th>Quarterly</th>
<th>Monthly</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>Expected</td>
<td>Actual</td>
<td>Expected</td>
</tr>
<tr>
<td>Mean</td>
<td>6.78</td>
<td>6.73</td>
<td>6.77</td>
</tr>
<tr>
<td>Variance</td>
<td>0.51</td>
<td>0.46</td>
<td>0.46</td>
</tr>
</tbody>
</table>
Table 3. Cont.

<table>
<thead>
<tr>
<th></th>
<th>Annual</th>
<th>Quarterly</th>
<th>Monthly</th>
</tr>
</thead>
<tbody>
<tr>
<td>Testing Differences in Variances</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$F$-stat</td>
<td>1.11</td>
<td>1.03</td>
<td>1.01</td>
</tr>
<tr>
<td>$F$-critical (one-tail)</td>
<td>1.98</td>
<td>1.39</td>
<td>1.21</td>
</tr>
<tr>
<td>$F$-critical (two-tail)</td>
<td>2.90</td>
<td>1.68</td>
<td>1.35</td>
</tr>
<tr>
<td>$p$-value</td>
<td>0.40</td>
<td>0.45</td>
<td>0.47</td>
</tr>
<tr>
<td>Testing Differences in Means</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hypothesized Difference</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Pooled Variance</td>
<td>0.48</td>
<td>0.45</td>
<td>0.45</td>
</tr>
<tr>
<td>$t$-stat</td>
<td>0.27</td>
<td>0.14</td>
<td>0.08</td>
</tr>
<tr>
<td>$t$-critical</td>
<td>2.01</td>
<td>1.97</td>
<td>1.96</td>
</tr>
<tr>
<td>$p$-value</td>
<td>0.79</td>
<td>0.89</td>
<td>0.93</td>
</tr>
</tbody>
</table>

This table provides the results of the tests for differences in the averages and variances of the natural logarithms of the expected prices and the natural logarithms of the observed gold prices. The $F$-stat test was used to test the differences in variances. The results show that the test statistics are less than the critical values set for the test. Accordingly, the difference in the means, hypothesized to be zero, was tested under the assumption of no differences in terms of variance.

4. Results

Figures 1–3 are plots of the expected and actual prices against time. In these figures, the actual USD prices are used. Figure 1 shows the annual frequency, while Figure 2 reflects the quarterly frequency and Figure 3 the monthly frequency. Very strong correlations between the observed and expected prices are observed in each of the three figures. Numerically and as presented in the figures, the actual prices are sometimes above the expected prices, while, at other times, it is the other way around. Nevertheless, the two lines are very close and move closer to each other as we move from the annual to the monthly frequency. Figures 4–6 are plots of the natural logarithms of the observed prices against the natural logarithms of the expected prices. All three plots show upward-sloping straight lines, indicating a strong possibility of a linear relationship between the actual and expected prices. Perhaps, the expected prices obtained using the manuscript’s methodology can serve as good proxies of the actual prices or at least help in making reliable forecasts of the actual prices of gold.

Table 2 presents the results of the simple regression in Equation (11). The results are presented for all three frequencies. The values of the intercept are close to zero. At the annual frequency, the value of the intercept is equal to 0.3, but, for the quarterly and annual frequencies, it is 0.08 and 0.03, respectively. Statistically, primarily based on the $t$-stats, the hypothesis that the coefficient of the independent variable are not equal to zero cannot be rejected, primarily because the standard errors for the intercept terms are very small. The values of the coefficient for the annual, quarterly, and monthly frequencies are 0.95, 0.99, and 0.99 when rounded-off to two decimal places. The hypothesis that the coefficients are non-zero values cannot be rejected at the one percent significance level. They are also different from a value of one, given the very low values obtained for their standard errors. For all three regressions, the adjusted R-square and the R-square are very high. Given the model’s characteristics and the significance level of the $t$-stats, the expected values can be used to predict the actual gold prices.

The results for the testing of the averages and variances of the natural logarithms of the estimated price and the observed gold prices are presented in Table 3. In our study, the tests for differences in the variances have been carried out before the tests for differences in the means. The $F$-stat, as estimated by Equation (12), shows values of 1.11, 1.03, and 1.01 at the yearly, quarterly, and monthly frequencies. The table below also presents the critical $F$-values at these frequencies, and it can be observed that the test’s statistics are less than the critical $F$-stat. The hypothesis that there is no difference in terms of the variances in the natural logarithms of the estimated prices and the natural logarithms of the observed
prices cannot be rejected. Given that the hypothesis cannot be rejected, the difference in the means of the two price series are tested assuming no differences in the variances. The test’s statistic is a t-stat, estimated, as shown in Equation (14), using the pooled variance estimated using Equation (13). The pooled variances, test statistics, critical value of the test statics, as well as the p-values are presented in Table 3. The t-stat test statistics are 0.27, 0.14, and 0.08 at the annual, quarterly, and monthly frequencies. The p-values are also higher than the normally used alpha of 5%. The hypothesis, as far as the natural logarithms of the estimated price and the natural logarithms of the actual prices are concerned, that their averages are not different from one another cannot be rejected. Thus, at least statistically speaking, there is no difference between the average and variance of the natural logarithms of the GBM-simulated gold prices and those of the actually observed gold prices at yearly, quarterly, and monthly frequencies.

5. Conclusions

Gold objects have been of prime importance for a very long time [1–3], and gold itself has gathered importance as a hedging instrument [4–6] as well. Given the growing importance of gold, gold prices are a subject of renewed interest [7] for a large number of economic entities. Accordingly, a number of researchers have developed and tested models to obtain gold prices, such as Refs. [8–10,39,40], among others. Ref. [8] applied a number of different time series models to obtain gold prices from Thailand, at a monthly frequency, between 2009 and 2021. Ref. [9] applied the ARIMA model to forecast gold prices based on monthly data obtained from the Multi Commodity Exchange of India for the period between 2003 and 2014. Ref. [10] applied a multivariate stochastic model to predict gold prices. Ref. [39] found the ARIMA approach to provide good forecasts, after comparing six different estimating procedures typically used to forecast gold prices and the prices of other metals like silver, platinum, palladium, and rhodium. Ref. [40] applied the exponential-smoothing approach to predict Malaysian monthly gold prices from data obtained from the World Bank. The GBM approach can be found among the popular methods for obtaining accurate forecasts, a method which has also been used to forecast the prices of a number of different assets like stocks, stock indices, options, energy products, agricultural products, and metals, including gold.

In this manuscript, a GBM approach was used to obtain accurate and reliable gold prices. While GBM has also been used to assess gold prices in Refs. [11–13], the methodology used in this manuscript is similar that used in Refs. [18,32,33]. Using GBM, one hundred thousand one-period-ahead prices and returns of gold were obtained at yearly, quarterly, and monthly frequencies. From the simulated returns, probabilities were obtained by taking the differences in the cumulative distribution function of two adjacent returns after the returns had been sorted from low to high. By multiplying simulated prices with the associated probabilities, and adding up the such multiplication, the expected prices were obtained. Results indicated that the expected prices can be good and reliable predictors of the actual prices. The averages and variances of the natural logarithms of the estimated prices were also found to be similar to those of the actually observed prices.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: No new data was created. Data was obtained from DataStream through a subscription from Refinitiv.

Conflicts of Interest: The author declares no conflicts of interest.
References


**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.