**Gold prices and volatility**

Does the price of gold increase in volatile markets?

Can volatility be measured?

How sensitive is the price of gold to volatility in the markets?

US real GDP

Share of Investment in GDP

Share of Government expenditure in GDP

Share of Consumption in GDP

Population growth

Uncertainty – war, policy, politics

Stock market indices (S&P500, Dow Jones)

Random walk:

Measure of volatility: Variance from trend, or standard deviation.

Stock market: Leading, procyclical variable for real GDP

Efficient market theory:

Changes in prices can be used to predict future movements in real GDP

But volatile – big movements in single days, based on speculation

Robert Shiller

Paul Samuelson

Risk = variability.

**The theory of investment value**

Application of free market economics to finance

<https://www.forbes.com/sites/jerrybowyer/2013/04/28/the-economy-has-nothing-to-do-with-the-stock-markets-right>

* covariance matrix
* residuals

**Random walk:**

“The null hypothesis of the test is that the time series is non-stationary.”

Running the example, we can see that the test statistic value was 0.341605. This is larger than all of the critical values at the 1%, 5%, and 10% confidence levels. Therefore, we can say that the time series does appear to be non-stationary with a low likelihood of the result being a statistical fluke.

<https://machinelearningmastery.com/gentle-introduction-random-walk-times-series-forecasting-python/>

**Predicting a Random Walk**

A random walk is unpredictable; it cannot reasonably be predicted.

Given the way that the random walk is constructed, we can expect that the best prediction we could make would be to use the observation at the previous time step as what will happen in the next time step.

Simply because we know that the next time step will be a function of the prior time step.

This is often called the naive forecast, or a persistence model.

We can implement this in Python by first splitting the dataset into train and test sets, then using the persistence model to predict the outcome using a rolling forecast method. Once all predictions are collected for the test set, the mean squared error is calculated.

Running the example estimates the mean squared error of the model as 1.

This too is expected, given that we know that the variation from one time step to the next is always going to be 1, either in the positive or negative direction, and the square of this expected error is 1 (1^2 = 1).

Another error that beginners to the random walk make is to assume that if the range of error (variance) is known, then we can make predictions using a random walk generation type process.

That is, if we know the error is either -1 or 1, then why not make predictions by adding a randomly selected -1 or 1 to the previous value.

We can demonstrate this random prediction method in Python below.

Running the example, we can see that indeed the algorithm results in a worse performance than the persistence method, with a mean squared error of 1.765.

**Persistence, or the naive forecast, is the best prediction we can make for a random walk time series.**

The price of gold can be influenced by a number of macroeconomic variables.[[28]](https://en.wikipedia.org/wiki/Gold_as_an_investment#cite_note-eprints.hud.ac.uk-28) Such variables include the price of oil, the use of [quantitative easing](https://en.wikipedia.org/wiki/Quantitative_easing), currency exchange rate movements and returns on equity markets.[[28]](https://en.wikipedia.org/wiki/Gold_as_an_investment#cite_note-eprints.hud.ac.uk-28)

Gold price return behavior.

In financial markets, gold has many faces: as a ‘safe haven’, a ‘store of value’, hedging and derivative instrument, and risk and portfolio diversification.

**Gold prices and volatility**

**ASE project April 25, 2020**

**Word count:**

**1. Introduction**

The aim of this study is explore how the price of gold reacts to: (i) market volatility and (ii) changes in macroeconomic indicators, namely inflation. That sounds good. Use **daily data** if you can get it. Compare the period **before the pandemic and the pandemic**. Look at the **correlation** between the **change in the log of the gold price** and the **change in the** **log of the stock market index**. Start by graphing the scatter between them. Get back to me if you have other questions

I look at the variation in the price of gold over time and how this is correlated with volatility in markets:

* volatility in US stock markets
  + NASDAQ, Dow, S&P500
  + equity returns
* volatility in gold markets
  + change in price over time
  + log growth

**2. Background**

* 1. **The definition of gold**

Quantify the volatility of the price of gold to evaluate it’s status as a safe haven. Rather than attempt to forecast the price of gold, I aimed to investigate the relationship between uncertainty and volatility in financial markets and the price of gold. Specifically, I focus on the period of Nov. 2018 to present, including 2019-2020, to study the changes in the price of gold a year before, several months before, during the start of and at the current state of the current COVID-19 pandemic. I used the logged values of daily prices of gold, daily stock market indices (s&p 500, dax,

Commodities are… Gold is considered a commodity (“Commodity”, FRED category; however, unlike other commodities, the supply of gold is relatively constant, with changes in its prices driven mostly by investor behaviour. While other commodities, such as oil, metals, etc., represent important inputs into production of manufactured goods and hence manufacturing economies, gold has often been used as a hedge for inflation in finance, considered a ‘safe haven’, or a ‘store of value’, for example in times of high rates of inflation, as well as a derivative instrument, and in risk and portfolio diversification.

The price of gold has a negative relationship with interest rates. For this reason gold is considered a “safe haven” when interest rates are low. Understanding how and why commodities prices fluctuate can determine your success in trading these instruments. Without this knowledge, you may be fighting a losing battle.

Based on this theory, I expect to following patterns:

* Gold prices to rise with increase inflation/Gold prices to fall with increase interest rates
* Gold prices to fall with increased equity returns/Gold prices to rise with increase market volatility
* Gold prices to rise with unemployment.

**2.2 The definition of volatility**

Here, I am defining volatility by the range of the price on a given day. For this I use the standard deviation of the price on a given day. A large standard deviation indicates high price volatility, while a small standard deviation indicates a low price volatility on a given day, calculated from the daily log price change. A period of >x days with high price volatility is denoted as a volatile period, and is historically associated with recessions.

Volatility shows how the prices fluctuate in a given period of time. The greater the volatility, the wider the range of prices. High volatility means that the price of the asset can change dramatically over a short time period of time in either direction. A lower volatility means the asset's value does not fluctuate as dramatically.Instead its value tends to change at a steady pace

**Volatility** is measured by the day-to-day percentage difference in the **price** of the **commodity**. The degree of variation, not the level of **prices**, defines a **volatile market**

**Volatility threshold?** When do prices start to change.

**2.3 Efficient markets hypothesis**

Supply and demand.   
More sellers than buyers. Price goes down.   
More buyers than sellers. Prices go up.  
Market clearance?

the price of oil

the use of [quantitative easing](https://en.wikipedia.org/wiki/Quantitative_easing) (inflation)

currency exchange rate movements

returns on equity markets.[[28]](https://en.wikipedia.org/wiki/Gold_as_an_investment#cite_note-eprints.hud.ac.uk-28)

**2.4 Random walk**

**…**

**2.5 Factors driving the gold price**

1. Quantity of money, this is the combined Monetary Base (MB) of the US, Euro Zone and Japan. These 3 MB are monthly averages, stated or translates to USD with the applicable monthly average exchange rate.
2. Real Interest Rate, measured as the annual rate of the US 1 Year treasury bill minus the US CPI inflation rate
3. USD Exchange Rate, measured as the exchange rate of the USD against a basket of currencies, consisting of the Euro, Renminbi, Yen and Rupee, weighted according to their GDP.
4. Financial uncertainty, measured as the yield spread between Junk Bunds, measured as the Yield on the HYG ETF minus the yield on long term US government bonds, measured as the yield on the TLT ETF.
5. <https://goldenopportunitiesforinvestors.wordpress.com/gold-price-regression-model/>

**2.6 Hypothesis**

**3. Data**

Daily price of gold (10 a.m.) FRED

**3.1 Historical price of gold**



**3.2 Log growth of stock market**

**3.3 Relationship between price volatility and increased gold prices**

- hypothesis

**3.4 Model**

- Variables   
- Hypothesis  
- Regression

**3.5 Descriptive statistics**

- Mean  
- Variance and standard deviation

**3.6 Regression**

- Durbin-Watson  
- Covariance matrix

- Residuals

- hypothesis testing

- confidence intervals

- R2 and adjusted R2

- standard error

- reparameterization

**4. Analysis**

**4.1 Model**

**4.2 Descriptive statistics**

**4.3 Regression**

* **Gold as a store of value**
* **Gold as an indicator of recessions**
  + **Chow break test during previous recessions**Nov 1973–Mar 1975  
    Jan 1980–July 1980  
    July 1990–Mar 1991  
    Mar 2001–Nov 2001  
    Dec 2007–June 2009  
    Mar 2020–present (Lockdown Recession)
* **Gold during COVID-19 pandemic**

**6. Summary and Conclusions**

**7. Bibliography**

You must give a bibliographic citation for any work referred to in the text. If in doubt, follow the Harvard system, used in most economics articles.

**8. Appendices**

Not counted in word count.

|  |  |
| --- | --- |
| **COVID-19 timeline** | **Market volatility** |
| **Dec 8.** first patients with covid symptoms in wuhan **Dec 31.** china alerts WHO of rise in pneumonia cases **Jan 1.** wuhan fish market shut ds own **Jan 7.** formal identification of covid-19 virus **Jan 11.** first recorded death of coronavirus  **Jan 13.** first case reported outside of china (Thailand) **Jan 23.** wuhan lockdown **Jan 29. Mar 6.** 100,000 cases (John Hopkins  **Mar 9.** italy lockdown **Mar 15.** spain lockdown **Mar 23.** uk lockdown **April 2.** 1,000,000 cases **April 14.** 2,000,000 cases **April 20.** US oil price crash, <$1. May oil futures, $-40. Brent (North Sea), WTI (US) |  |

**References**

John Hopkins, Coronavirus data. <https://coronavirus.jhu.edu/map.html>. Accessed: April 20, 2020.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  |  |  |  |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Dependent Variable: LOG\_GOLD | | | |  |
| Method: Least Squares | | |  |  |
| Date: 04/21/20 Time: 15:38 | | | |  |
| Sample (adjusted): 1/02/2018 4/20/2020 | | | |  |
| Included observations: 565 after adjustments | | | | |
|  |  |  |  |  |
|  |  |  |  |  |
| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|  |  |  |  |  |
|  |  |  |  |  |
| C | 14.11535 | 0.982718 | 14.36358 | 0.0000 |
| LOG\_DOW | -3.860544 | 0.297428 | -12.97975 | 0.0000 |
| LOG\_NASDAQ | -1.279814 | 0.236885 | -5.402669 | 0.0000 |
| LOG\_SP500 | 5.502990 | 0.518393 | 10.61549 | 0.0000 |
|  |  |  |  |  |
|  |  |  |  |  |
| R-squared | 0.549901 | Mean dependent var | | 7.215259 |
| Adjusted R-squared | 0.547494 | S.D. dependent var | | 0.092127 |
| S.E. of regression | 0.061973 | Akaike info criterion | | -2.717196 |
| Sum squared resid | 2.154574 | Schwarz criterion | | -2.686493 |
| Log likelihood | 771.6079 | Hannan-Quinn criter. | | -2.705212 |
| F-statistic | 228.4639 | Durbin-Watson stat | | 0.046066 |
| Prob(F-statistic) | 0.000000 |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |



**Market and price volatility**

**Log growth:**

lg\_gold = (ln\_gold - ln\_gold(-1))\*100

Similarly for S&P500, Dow and Nasdaq indices.

**Gold and Dow, S&P500, Nasdaq**



**Gold and Dow**  
 

**Gold and NASDAQ**



**Gold and S&P500**



**Stock market index correlation**

|  |  |  |  |
| --- | --- | --- | --- |
|  | LG\_DOW | LG\_NASDAQ | LG\_SP500 |
| LG\_DOW | 1 | 0.93577 | 0.985634 |
| LG\_NASDAQ | 0.935773 | 1 | 0.969791 |
| LG\_SP500 | 0.985634 | 0.96979 | 1 |

**Growth of S&P500, NASDAQ and Dow Jones is highly correlated (r = 0.9697777455)**



**Growth of Dow is poorly correlated with both NASDAQ (0.1184650107) and S&P500 (0.1362538714)**



**Descriptive statistics**

**Gold**



**Dow**



**NASDAQ**



**S&P500**



**Regression**

* **specify a model:**

lg\_gold c lg\_dow lg\_nasdaq lg\_sp500

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Dependent Variable: LG\_GOLD | | | |  |
| Method: Least Squares | | |  |  |
| Date: 04/21/20 Time: 16:26 | | | |  |
| Sample (adjusted): 1/03/2018 4/20/2020 | | | |  |
| Included observations: 534 after adjustments | | | | |
|  |  |  |  |  |
|  |  |  |  |  |
| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|  |  |  |  |  |
|  |  |  |  |  |
| C | 0.036862 | 0.035261 | 1.045407 | 0.2963 |
| LG\_DOW | 0.464067 | 0.146297 | 3.172080 | 0.0016 |
| LG\_NASDAQ | -0.078817 | 0.099289 | -0.793813 | 0.4277 |
| LG\_SP500 | -0.317241 | 0.222901 | -1.423242 | 0.1553 |
|  |  |  |  |  |
|  |  |  |  |  |
| R-squared | 0.056758 | Mean dependent var | | 0.031005 |
| Adjusted R-squared | 0.051419 | S.D. dependent var | | 0.835527 |
| S.E. of regression | 0.813763 | Akaike info criterion | | 2.433167 |
| Sum squared resid | 350.9712 | Schwarz criterion | | 2.465230 |
| Log likelihood | -645.6555 | Hannan-Quinn criter. | | 2.445713 |
| F-statistic | 10.63062 | Durbin-Watson stat | | 1.916042 |
| Prob(F-statistic) | 0.000001 |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |

**P-value: Significant only for Dow  
T-statistic: >2 only for Dow**

**Gold and Dow**

lg\_gold c lg\_dow

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Dependent Variable: LG\_GOLD | | | |  |
| Method: Least Squares | | |  |  |
| Date: 04/21/20 Time: 16:51 | | | |  |
| Sample (adjusted): 1/03/2018 4/20/2020 | | | |  |
| Included observations: 534 after adjustments | | | | |
|  |  |  |  |  |
|  |  |  |  |  |
| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|  |  |  |  |  |
|  |  |  |  |  |
| C | 0.031321 | 0.035548 | 0.881089 | 0.3787 |
| LG\_DOW | 0.094481 | 0.021443 | 4.406154 | 0.0000 |
|  |  |  |  |  |
|  |  |  |  |  |
| R-squared | 0.035208 | Mean dependent var | | 0.031005 |
| Adjusted R-squared | 0.033394 | S.D. dependent var | | 0.835527 |
| S.E. of regression | 0.821458 | Akaike info criterion | | 2.448266 |
| Sum squared resid | 358.9898 | Schwarz criterion | | 2.464297 |
| Log likelihood | -651.6870 | Hannan-Quinn criter. | | 2.454539 |
| F-statistic | 19.41420 | Durbin-Watson stat | | 1.882245 |
| Prob(F-statistic) | 0.000013 |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |

**P-value: <0.05 (0.00)**

**T-statistic: >2**

**Durbin-Watson: <2 = serial correlation**

Serial correlation suggests dynamic misspecification: one form of mis-specification occurs when the true model is dynamic and has been wrongly assumed to be static(ref: <https://link.springer.com/chapter/10.1007/978-94-009-7526-2_6>)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Dependent Variable: LG\_GOLD | | | |  |
| Method: Least Squares | | |  |  |
| Date: 04/21/20 Time: 16:52 | | | |  |
| Sample (adjusted): 1/03/2018 4/20/2020 | | | |  |
| Included observations: 534 after adjustments | | | | |
| **HAC standard errors & covariance** (Bartlett kernel, Newey-West fixed bandwidth = 6.0000) | | | | |
|  |  |  |  |  |
|  |  |  |  |  |
| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|  |  |  |  |  |
|  |  |  |  |  |
| C | 0.031321 | 0.036693 | 0.853590 | 0.3937 |
| LG\_DOW | 0.094481 | 0.039099 | 2.416427 | 0.0160 |
|  |  |  |  |  |
|  |  |  |  |  |
| R-squared | 0.035208 | Mean dependent var | | 0.031005 |
| Adjusted R-squared | 0.033394 | S.D. dependent var | | 0.835527 |
| S.E. of regression | 0.821458 | Akaike info criterion | | 2.448266 |
| Sum squared resid | 358.9898 | Schwarz criterion | | 2.464297 |
| Log likelihood | -651.6870 | Hannan-Quinn criter. | | 2.454539 |
| F-statistic | 19.41420 | Durbin-Watson stat | | 1.882245 |
| Prob(F-statistic) | 0.000013 | Wald F-statistic | | 5.839122 |
| Prob(Wald F-statistic) | 0.016009 |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |

**OLS to HAC:** The standard error on LE increases from 0.021443 to 0.039099  
  
**Residuals**



**Residuals:** Differences between fitted and actual values – should be random (if no serial correlation).

**Variance = constant?**

**Serial correlation?**

**Scatter**



**Lags**lg\_gold c lg\_dow lg\_dow(-1) lg\_gold(-1) @trend

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Dependent Variable: LG\_GOLD | | | |  |
| Method: Least Squares | | |  |  |
| Date: 04/21/20 Time: 17:37 | | | |  |
| Sample (adjusted): 1/04/2018 4/20/2020 | | | |  |
| Included observations: 503 after adjustments | | | | |
|  |  |  |  |  |
|  |  |  |  |  |
| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|  |  |  |  |  |
|  |  |  |  |  |
| C | -0.033723 | 0.074148 | -0.454807 | 0.6494 |
| LG\_DOW | 0.104966 | 0.022961 | 4.571505 | 0.0000 |
| LG\_DOW(-1) | 0.033295 | 0.023428 | 1.421183 | 0.1559 |
| LG\_GOLD(-1) | 0.078977 | 0.044911 | 1.758539 | 0.0793 |
| @TREND | 0.000181 | 0.000215 | 0.841461 | 0.4005 |
|  |  |  |  |  |
|  |  |  |  |  |
| R-squared | 0.050280 | Mean dependent var | | 0.022738 |
| Adjusted R-squared | 0.042652 | S.D. dependent var | | 0.849722 |
| S.E. of regression | 0.831404 | Akaike info criterion | | 2.478488 |
| Sum squared resid | 344.2336 | Schwarz criterion | | 2.520442 |
| Log likelihood | -618.3398 | Hannan-Quinn criter. | | 2.494947 |
| F-statistic | 6.591297 | Durbin-Watson stat | | 2.049356 |
| Prob(F-statistic) | 0.000036 |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |

Yt = β0 + β11Yt–1 + ut

Lags aren’t significant. Only Dow is significant (T>2 and P<0.05).

Is the lagged change in the Dow index a useful predictor of the current change in gold?

ΔGold = ‒0.033723 + 0.033295ΔDowt–1 R2 = 0.05

(0.074148) (0.023428)

*t =* 0.033295/0.023428 = 1.42 < 1.96 (2 S.D.)   
Do not reject H0: *β*1 = 0 at 5% significance level.

-Yes, the lagged change in inflation is a useful predictor of current change in inflation–but the 2 R is pretty low!

**Dow and Gold (-change of dependent variable to Dow)**

lg\_dow c lg\_gold

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Dependent Variable: LG\_DOW | | |  |  |
| Method: Least Squares | | |  |  |
| Date: 04/21/20 Time: 17:41 | | | |  |
| Sample (adjusted): 1/03/2018 4/20/2020 | | | |  |
| Included observations: 534 after adjustments | | | | |
|  |  |  |  |  |
|  |  |  |  |  |
| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|  |  |  |  |  |
|  |  |  |  |  |
| C | -0.014901 | 0.070646 | -0.210926 | 0.8330 |
| LG\_GOLD | 0.372646 | 0.084574 | 4.406154 | 0.0000 |
|  |  |  |  |  |
|  |  |  |  |  |
| R-squared | **0.035208** | Mean dependent var | | -0.003347 |
| Adjusted R-squared | 0.033394 | S.D. dependent var | | 1.659346 |
| S.E. of regression | 1.631404 | Akaike info criterion | | 3.820497 |
| Sum squared resid | 1415.907 | Schwarz criterion | | 3.836529 |
| Log likelihood | -1018.073 | Hannan-Quinn criter. | | 3.826770 |
| F-statistic | 19.41420 | Durbin-Watson stat | | **2.647290** |
| Prob(F-statistic) | 0.000013 |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |

**P-value: <0.05 (0.00)**

**T-statistic: >2**

**Durbin-Watson: <2 = serial correlation**

**Residuals**



**Scatter**



**Lags**

lg\_dow c lg\_gold lg\_gold(-1) lg\_dow(-1) @trend

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Dependent Variable: LG\_DOW | | |  |  |
| Method: Least Squares | | |  |  |
| Date: 04/21/20 Time: 17:54 | | | |  |
| Sample (adjusted): 1/04/2018 4/20/2020 | | | |  |
| Included observations: 503 after adjustments | | | | |
|  |  |  |  |  |
|  |  |  |  |  |
| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|  |  |  |  |  |
|  |  |  |  |  |
| C | 0.087498 | 0.141740 | 0.617312 | 0.5373 |
| LG\_GOLD | 0.383696 | 0.083932 | 4.571505 | 0.0000 |
| LG\_GOLD(-1) | 0.117751 | 0.085970 | 1.369671 | 0.1714 |
| LG\_DOW(-1) | -0.290211 | 0.042957 | -6.755779 | 0.0000 |
| @TREND | -0.000302 | 0.000411 | -0.736343 | 0.4619 |
|  |  |  |  |  |
|  |  |  |  |  |
| R-squared | **0.117962** | Mean dependent var | | 0.008658 |
| Adjusted R-squared | 0.110877 | S.D. dependent var | | 1.685778 |
| S.E. of regression | 1.589576 | Akaike info criterion | | 3.774702 |
| Sum squared resid | 1258.322 | Schwarz criterion | | 3.816656 |
| Log likelihood | -944.3376 | Hannan-Quinn criter. | | 3.791161 |
| F-statistic | 16.65032 | Durbin-Watson stat | | 1.990789 |
| Prob(F-statistic) | 0.000000 |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |

LG\_DOW(-1) and gold are significant (|T|>2 and P<0.05).  
DW ~2.

**Low R2**

* It is difficult to figure out all the factors which influence the value of YY.

Due to that, our model suffers from bias. We can able to reduce the bias, if we can able to get good estimate (without bias) for f.f. It means that if we can able to find a good estimate for ff, we can reduce some error, which is called reducible error. (from <https://www.quora.com/Why-is-the-error-term-normally-distributed>)

**Explanations for my low R2**

* The underlying theoretical equation might have a different functional form than the one chosen for the regression. For example, the underlying equation might be nonlinear in the variables for a linear regression
* If non-linear: would show up in the residuals: check residuals.
* The t-test and F- test are not applicable unless the error term is normally distributed. See Central Limit Theorem.
* Assumptions: errors are normally distributed. test
* Formulate Null hypothesis:
* Assumptions:
  + errors have constant variance – homoscedasticity
  + errors are not correlated over time – no autocorrelation, independent