**Behaviour of gold prices under market volatility**

**ASE project April 25, 2020**

**Word count: 3000**

**1. Introduction**

In this analysis, I aim 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, dow jones, Nasdaq).

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. Here, I carry out a time series analysis to determine how the price of gold changes with market volatility. In my analysis I use daily data to evaluate how the price of gold changes over time and whether this is related to the level of volatility in the stock market.

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

**2. Background**

* 1. **The definition of gold**

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.

**2.5 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**

Increased volatility, increased price.

**Volatility**  
Market volatility is high when the rate of buying and selling is high. In the gold market, this translate to large volumes of gold being traded in a short period of time. The statistic that is widely used to evaluate volatility is the standard deviation: …, that is, the difference between the actual value and the average value, or the dispersion around the mean.

Standard deviations of log returns

**3. Data**

**3.1 Daily price of gold**

The price of gold is updated two times per day in the London Bullion Market: once at the open of trading in the UK market (10:30 A.M.; London time) and again at the close (3 P.M.). As a result, there are two types of daily data available for gold from the Federal Reserve Economic Data (FRED) database: Gold Fixing Price 10:30 A.M. and Gold Fixing Price 3:00 P.M. To decide which dataset to use in my analysis, I downloaded both datasets from FRED and plotted them in EViews (Figure 1).









**Figure 1.** Top: Gold Fixing Price 10:30 A.M. Bottom: Gold Fixing Price 3:00 P.M. Source: FRED.

The datasets were visually very similar. To determine whether there were any significant differences between these datasets, I conducted a t-test in EViews. First, I subtracted the data for Gold Fixing Price 3:00 P.M. from Gold Fixing Price 10:30 A.M. to calculate the differences between prices. Then, I ran a two-tailed t-test to determine whether this difference was significant (H0: µ = 0 and H1: µ ≠ 0) (Table 1).

**Table 1.** Hypothesis test for difference between Gold Fixing Prices datasets.

|  |  |  |  |
| --- | --- | --- | --- |
| Hypothesis Testing for DIFFERENCE\_GOLD\_DATASETS | | | |
| Date: 05/02/20 Time: 17:03 | | |  |
| Sample: 4/01/1968 5/01/2020 | | |  |
| Included observations: 13081 | | |  |
| Test of Hypothesis: Mean = 0.000000 | | | |
|  |  |  |  |
|  |  |  |  |
| Sample Mean = 0.164105 | | |  |
| Sample Std. Dev. = 4.465556 | | |  |
|  |  |  |  |
| Method | | Value | Probability |
| t-statistic | | 4.203082 | 0.0000\* |
|  |  |  |  |

\*P<0.05

As shown in Table 1, at a 95% significance level (P<0.05), the datasets were not significantly different. Since price did not vary significantly between the datasets, I chose to use the data from the Gold Fixing Price 10:30 A.M. dataset for my analysis because it was slightly more complete that Gold Fixing Price 3:00 P.M., with 87 more observations (13169 vs. 13082 observations, respectively) (Table 1).

**Table 2.** Characteristics ofthedata used in my analysis.

|  |  |
| --- | --- |
| **Name of dataset** | Gold Fixing Price 3:00 P.M. (London time) |
| **Source** | FRED (based on London Bullion Market) |
| **Units** | U.S. Dollars (per Troy Ounce) |
| **Frequency** | Daily |
| **Traits** | Not Seasonally Adjusted |

**3.2. Transformation of price**

As shown in Figure 1, the raw price data is highly skewed to the right. This appears to be due to the fact that, until recently (~2005), the price of gold was rarely above 400 USD. To remove this trend, I took the logs (natural log) of the prices using the following equation in EViews:

[1]

The transformation does not have an great impact on the visual interpretation, as seen in Figure 2, however, it allows me to get rid of the negative values in the data for further analyses down the line. Furthermore, it can also be used to calculate the change in the log of the gold price (Figure 4).





**Figure 2.** Log values of the gold price.

**Log growth of price:**

Analysis: …





**Figure 3.** Log growth of price.

The change in the price of gold has a very positive kurtosis, indicating that data is unlikely to

**3.3. Distribution**



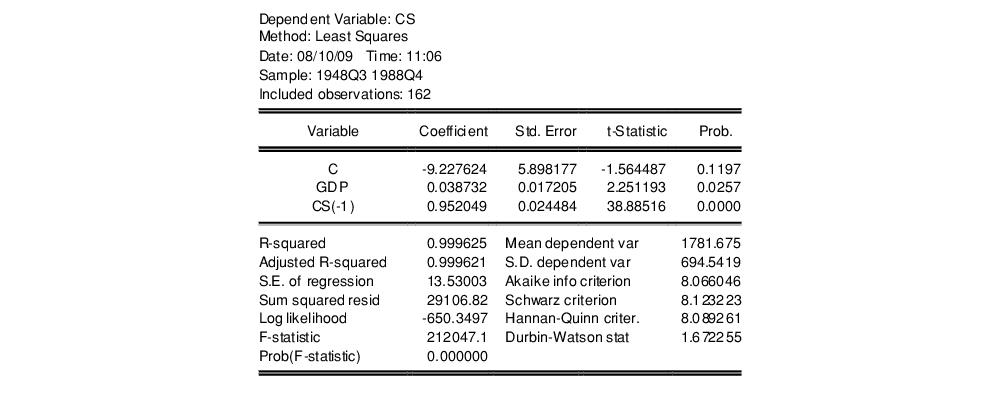
Although the average volatility for the period 1968 to 2020 is low (4.465556), on any given day the price has changed by either +80 (max) or -63 (min). This indicates that there should be a significant skew in the data. The decade between 1990 and 2000 was particularly stable, compared to periods of high volatility around 1980 and since 2008. Large spikes can be observed in 1980, 2008, 2012, and 2020.

|  |  |
| --- | --- |
|  | DIFFERENCE\_GOLD\_DATASETS |
| Mean | 0.164105 |
| Median | 0.000000 |
| Maximum | 80.00000 |
| Minimum | -63.85000 |
| Std. Dev. | 4.465556 |
| Skewness | 0.623096 |
| Kurtosis | 32.37216 |
|  |  |
| Jarque-Bera | 471066.7 |
| Probability | 0.000000 |
|  |  |
| Sum | 2146.663 |
| Sum Sq. Dev. | 260830.8 |
|  |  |
| Observations | 13081 |

**3.4. Autocorrelation**

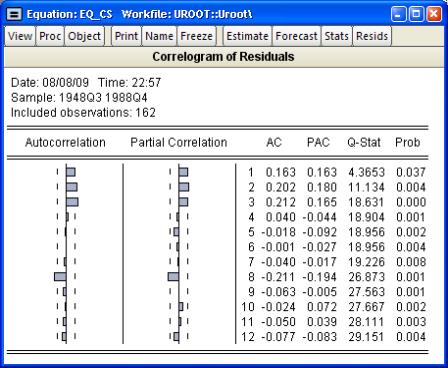
<http://www.eviews.com/help/helpintro.html#page/content/timeser-Testing_for_Serial_Correlation.html#ww192193>

As an example of the application of serial correlation testing procedures, consider the following results from estimating a simple consumption function by ordinary least squares using data in the workfile “Uroot.WF1”:



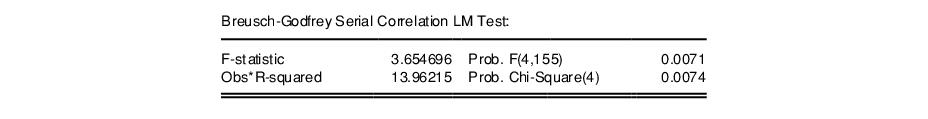
A quick glance at the results reveals that the coefficients are statistically significant and the fit is very tight. However, if the error term is serially correlated, the estimated OLS standard errors are invalid and the estimated coefficients will be biased and inconsistent due to the presence of a lagged dependent variable on the right-hand side. The Durbin-Watson statistic is not appropriate as a test for serial correlation in this case, since there is a lagged dependent variable on the right-hand side of the equation.

Selecting **View/Residual Diagnostics/Correlogram-Q-statistics** for the first 12 lags from this equation produces the following view:



The correlogram has spikes at lags up to three and at lag eight. The *Q*-statistics are significant at all lags, indicating significant serial correlation in the residuals.

Selecting**View/Residual Diagnostics/Serial Correlation LM Test…** and entering a lag of 4 yields the following result (top portion only):



The test rejects the hypothesis of no serial correlation up to order four. The *Q*-statistic and the LM test both indicate that the residuals are serially correlated and the equation should be re-specified before using it for hypothesis tests and forecasting.

Last updated: Thu, 04 Apr 2019 22:48:34 PST

[Back to top](http://www.eviews.com/help/content/timeser-Testing_for_Serial_Correlation.html)

**Regression 1**

Dependent: Gold

Independent: Stock (Dow, Nasdaq, S&P500), Oil

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Dependent Variable: LG\_GOLD | | | |  |
| Method: Least Squares | | |  |  |
| Date: 05/02/20 Time: 20:57 | | | |  |
| Sample (adjusted): 1/11/2019 4/09/2020 | | | |  |
| Included observations: 290 after adjustments | | | | |
|  |  |  |  |  |
|  |  |  |  |  |
| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|  |  |  |  |  |
|  |  |  |  |  |
| C | 0.076463 | 0.055996 | 1.365521 | 0.1732 |
| LG\_DOW | 0.671351 | 0.221409 | 3.032170 | **0.0027** |
| LG\_NASDAQ | -0.264898 | 0.162575 | -1.629392 | 0.1043 |
| LG\_OIL | -0.034919 | 0.011328 | -3.082605 | **0.0023** |
| LG\_SP500 | -0.318220 | 0.336801 | -0.944831 | 0.3455 |
|  |  |  |  |  |
|  |  |  |  |  |
| R-squared | 0.132617 | Mean dependent var | | 0.069933 |
| Adjusted R-squared | 0.120443 | S.D. dependent var | | 1.010856 |
| S.E. of regression | 0.948029 | Akaike info criterion | | 2.748227 |
| Sum squared resid | 256.1461 | Schwarz criterion | | 2.811501 |
| Log likelihood | -393.4929 | Hannan-Quinn criter. | | 2.773578 |
| F-statistic | 10.89363 | Durbin-Watson stat | | 1.858224 |
| Prob(F-statistic) | 0.000000 |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |

Only Dow and Oil reject hypothesis that coefficient is zero at 95% significance level. This is despite Dow, Nasdaq and S&P500 having an almost perfect correlation:

|  |  |  |
| --- | --- | --- |
|  | LG\_DOW | LG\_GOLD |
| LG\_DOW | 1 | 0.2235156085663471 |
| LG\_GOLD | 0.2235156085663471 | 1 |
| LG\_NASDAQ | **0.9579947321615215** | 0.147967602403121 |
| LG\_OIL | 0.1938986456821791 | -0.1451471045183852 |
| LG\_SP500 | **0.9901433883268196** | 0.1944104330998091 |

**Regression 2**

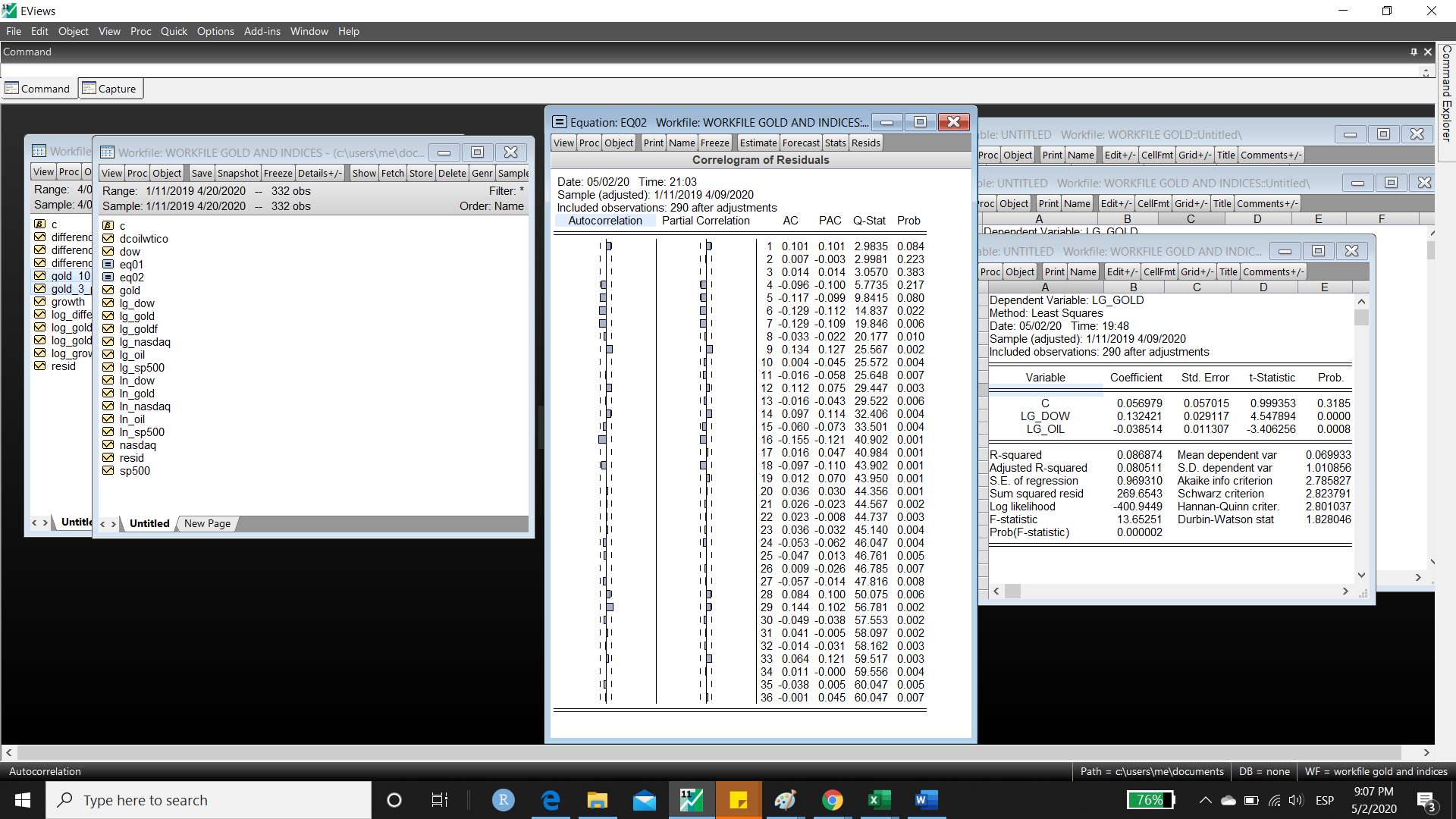
Dependent: Gold

Independent: Dow only, Oil

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Dependent Variable: LG\_GOLD | | | |  |
| Method: Least Squares | | |  |  |
| Date: 05/02/20 Time: 19:48 | | | |  |
| Sample (adjusted): 1/11/2019 4/09/2020 | | | |  |
| Included observations: 290 after adjustments | | | | |
|  |  |  |  |  |
|  |  |  |  |  |
| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|  |  |  |  |  |
|  |  |  |  |  |
| C | 0.056979 | 0.057015 | 0.999353 | 0.3185 |
| LG\_DOW | 0.132421 | 0.029117 | 4.547894 | 0.0000 |
| LG\_OIL | -0.038514 | 0.011307 | -3.406256 | 0.0008 |
|  |  |  |  |  |
|  |  |  |  |  |
| R-squared | 0.086874 | Mean dependent var | | 0.069933 |
| Adjusted R-squared | 0.080511 | S.D. dependent var | | 1.010856 |
| S.E. of regression | 0.969310 | Akaike info criterion | | 2.785827 |
| Sum squared resid | 269.6543 | Schwarz criterion | | 2.823791 |
| Log likelihood | -400.9449 | Hannan-Quinn criter. | | 2.801037 |
| F-statistic | 13.65251 | Durbin-Watson stat | | 1.828046 |
| Prob(F-statistic) | 0.000002 |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |

**Correlogram**

Correlogram (Correlogram-Q-statistics) for the first 36 lags from this equation:



The correlogram has spikes at 7 lags (autocorrelation: 5, 6, 7, 9, 12, 16, 29). The *Q*-statistics are significant (p<0.05) at all lags, indicating significant serial correlation in the residuals.

**Breusch-Godfrey test (serial correlation)**

Serial Correlation LM Testwithalag of 7 yields the following result (top portion only):

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Breusch-Godfrey Serial Correlation LM Test: | | | | |
| Null hypothesis: No serial correlation at up to 7 lags | | | | |
|  |  |  |  |  |
|  |  |  |  |  |
| F-statistic | 2.710384 | Prob. F(7,280) | | 0.0098 |
| Obs\*R-squared | 18.40328 | Prob. Chi-Square(7) | | 0.0103 |
|  |  |  |  |  |
|  |  |  |  |  |

The test rejects the hypothesis of no serial correlation up to order 7. The *Q*-statistic and the LM test both indicate that the residuals are serially correlated and the equation should be re-specified before using it for hypothesis tests and forecasting.





Data does not appear to be normally distributed. This will not affect my model since the sample size is very large. If I were to conduct an analysis with <30 observations (i.e. less than 30 days’ worth of data), I would need to account for this lack of normality in the errors.

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

- hypothesis

**3.5 Model**

- Variables   
- Hypothesis  
- Regression

**3.6 Descriptive statistics**

- Mean  
- Variance and standard deviation

**3.7 Regression**

- Durbin-Watson  
- Covariance matrix

- Residuals

- hypothesis testing

- confidence intervals

- R2 and adjusted R2

- standard error

- reparameterization

* **significant differences in price 10 am 3pm**1.Hypothesis
* H0: Mean of difference between prices is 0.
* H1: Mean of difference between prices is more than 0 (one-tailed).

2. Alpha: 0.05

3. One or two-tailed test

4. Sample size:

5. Standard deviation:

σ10 A.M. = √(xi - µ)2/(N-1) = 450.4254

σ3 P.M. = √(xi - µ)2/(N-1) = 450.3227

6. Z-score:

z = (xi - µ)/σ

**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