May 9

Second weekend looking at data and working on project.

Think I made some, small, breakthrough – look at short ranges (e.g. 2019-2020), when before I was looking at the relationship from 2010 to 2020.

Found some papers I can refer to – but, still don’t known what model to use.

Formulated a research hypothesis and have some initial findings – but, will need to confirm them: check my data is correct, check my transformations; check my dates, verify for other stock market crash dates; check regression for assumptions (normality, etc.)

Figured out how to check for the significance of a correlation – I can use this to test my other findings, i.e. in my regression model – but, is my Null hypothesis right?

May 10

Realised I was transforming my data wrong – found the correct formula, dlog.

Did all my analyses again, same results.

Ran an initial regression for 5 months before pandemic – correlation changes to significant negative relationship.

May 11

Emailed Ron to ask about how to interpret my coefficients – they are returns.

SER is low for COVID-19 pandemic dates, only 1.2% (SER = 0.012).

Next: Split data into high and low S&P volatility sub-periods.

Underlying theory is diversification (Markowitz).

May 9

<http://tastytradenetwork.squarespace.com/tt/blog/gold-bugs-rejoice>

**Since the end of November 2018, gold prices have increased roughly $100/ounce, or about 10%. With gold briefly moving above $1,340 in late February, we are now back in sight of the "nosebleed" prices observed at the start of this decade (about $1,800/ounce in 2011).**

**The most surprising aspect of the recent gold rally is that the momentum has sustained itself despite the fact that equity prices have also been rallying. After all, isn't gold usually perceived as a safety asset, which is bid up when most other asset classes are experiencing "de-risking?"**

I conducted an extensive study on [gold's historical relationship with equity prices](https://www.tastytrade.com/tt/shows/market-measures/episodes/gold-stocks-positive-or-negative-correlation-04-30-2018?_sp=ecc0d7ec-c84b-4d9a-af92-0ea219d5eb41.1589035327216) **last year**, and found rather conclusively that this ongoing relationship is anything but conclusive.

As you can see in the graphic below, the correlation between gold and equity prices has varied between strongly positive and strongly negative (and even zero correlation) over the years included in the study (1990-2018):

﻿Given the above information, it's a lot easier to rationalize the current movement in gold prices vis-a-vis stock prices - historical data clearly shows that they have rallied in unison at times in the past. My findings suggest that while **whole that assertion may reflect popular market sentiment, it's not exactly correct.**

So what other information can we examine to better understand the current dynamic in the precious metals sector?

One traditional metric that futures traders often utilize is the [Gold Silver Ratio](http://tastytradenetwork.squarespace.com/tt/blog/inside-the-volatility-gold-silver-ratio). Gold and silver enjoy a strong positive correlation, but over time the relationship tends to fluctuate between strong, and stronger. Traders often use this variable relationship as an opportunity to deploy a pairs trade involving the two.

In practical terms, the Gold/Silver Ratio simply reports the amount of silver it takes to buy one ounce of gold. As such, the ratio is calculated by dividing the current price of gold by the current price of silver.

Currently the Gold/Silver Ratio stands just short of 84, which is $gold/$silver ($1320.10/$15.78 = 83.65). That means it takes 84 ounces of silver to buy a single ounce of gold.

Over the last one hundred years, the Gold/Silver Ratio has ranged between about 15 and 100. That’s a fairly big difference in the respective piles of silver required to buy an ounce of gold.

However, in the last 20 years, that range has narrowed to more like 50 and 90. That means, at 84, the Gold/Silver Ratio is still very close to its recent historical highs (and not far from all-time highs, for that matter).

What this means in a practical sense is that gold has been rallying more strongly than silver in recent history. Traders that think the relationship will mean revert (i.e. that the Gold Silver Ratio will decline) typically consider selling gold and buying silver in these instances.

Logically, this makes sense, because a decreasing price of gold (the numerator) and an increasing price of silver (the denominator) would indeed result in a lower Gold/Silver ratio.

On the other hand, it’s entirely possible some traders are piling into gold because it is exhibiting such strength - or at least continuing to hold their existing long gold positions. This outlook might be based on the hypothetical premise that the US economy could **deteriorate in upcoming months/years.**

At the end of the day, whether a trader decides to enter the precious metals market is completely dependent on his/her own existing portfolio, strategic approach, outlook and risk profile.

Whatever the case may be, a recent episode of [*Closing the Gap: Futures Edition*](https://www.tastytrade.com/tt/shows/closing-the-gap-futures-edition/episodes/gold-silver-risk-onoff-assets-02-01-2019)breaks down recent movement in the gold and silver markets, and provides an overview of the associated position structure for traders considering a Gold Silver Ratio mean reversion trade. This can be effectively be viewed as a contrarian trade (shorting gold into recent strength), while adding on the long silver part of the “pairs” trade as a hedge.

We hope you'll take the time to review the complete episode of *Closing the Gap* focusing on precious metals when your schedule allows. Traders that want to investigate other potential positions in the gold sector may also want to review this new episode of [*Market Measures*](https://www.tastytrade.com/tt/shows/market-measures/episodes/gold-options-and-call-skew-02-20-2019), which outlines volatility-based trades in one of the best-known gold ETFs (GLD).

If you have any questions about this material, or any other trading topic, don't hesitate to leave a message in the space below, or reach out directly @tastytrade (Twitter) or support@tastytrade.com (email).

We look forward to hearing from you!

**Historical gold**



**Historical S&P500**



**2019-2020**

**Raw Price Data:**



**First Difference Data:**

Difference: remove time trend (stationary).

**First Difference-Logged Data:**

Log: remove units – percentage change (linear).

**Covariance analysis**

X-Y Correlation:   
Residuals:



Residuals-X correlation:



X-Y Correlation:



Residuals:



Residuals-X correlation:



X-Y Correlation:  


Residuals:   
Residuals-X correlation:



X-Y Correlation:  


Residuals:



Residuals-X correlation:



From scatterplot, growth of S&P500 and gold in 2019-2020 appears to be positively correlated.

* **Check correlation: Correlation matrix.**

-> Quick > Group statistics > Correlations:   
sp500\_log\_growth\_1dif gold\_log\_growth\_1dif

|  |  |  |
| --- | --- | --- |
|  | SP500\_LOG\_GROWTH\_1DIF | GOLD\_LOG\_GROWTH\_1DIF |
| SP500\_LOG\_GROWTH\_1DIF | 1 | 0.4365588020304154 |
| GOLD\_LOG\_GROWTH\_1DIF | 0.4365588020304154 | 1 |

-> Correlation (r) = 0.43  
-> Positive correlation.

* **Check probability: Covariance analysis.**

-> View > Covariance Analysis:

|  |  |  |  |
| --- | --- | --- | --- |
| Covariance Analysis: Ordinary | | |  |
| Date: 05/10/20 Time: 00:17 | | |  |
| Sample: 2/04/2019 5/01/2020 | | |  |
| Included observations: 275 | | |  |
| Balanced sample (listwise missing value deletion) | | | |
|  |  |  |  |
|  |  |  |  |
| Covariance | |  |  |
| Correlation | |  |  |
| t-Statistic | |  |  |
| Probability | |  |  |
| Observations | SP500\_LOG\_GROWTH\_1DIF | GOLD\_LOG\_GROWTH\_1DIF |  |
| SP500\_LOG\_GROWTH\_1DIF | 10.40416 |  |  |
|  | 1.000000 |  |  |
|  | ----- |  |  |
|  | ----- |  |  |
|  | 275 |  |  |
|  |  |  |  |
| GOLD\_LOG\_GROWTH\_1DIF | 2.050325 | 2.120084 |  |
|  | **0.436559** | 1.000000 |  |
|  | **8.017483** | ----- |  |
|  | 0.0000 | ----- |  |
|  | 275 | 275 |  |
|  |  |  |  |
|  |  |  |  |

Hypothesis test:

H0: r = 0 (correlation is not significantly different from zero)  
H1: r ≠ 0 (correlation is significantly different from zero)

-> P = 0.0000  
-> P < 0.05  
-> Reject Null.   
-> Gold & SP500 are significantly correlated at r = 0.43 at 95% level.

**Regression analysis**

**Regression #1: first-difference data**

d(gold) = 0.743125 + 0.098057 dlog(sp500) + t

(0.856850) (0.016627)

R2 =0.107750

SER = 14.59138

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Dependent Variable: GOLD\_DIF | | | |  |
| Method: Least Squares | | |  |  |
| Date: 05/09/20 Time: 19:12 | | | |  |
| Sample (adjusted): 2/04/2019 5/01/2020 | | | |  |
| Included observations: 290 after adjustments | | | | |
|  |  |  |  |  |
|  |  |  |  |  |
| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|  |  |  |  |  |
|  |  |  |  |  |
| C | 0.743125 | 0.856850 | 0.867276 | 0.3865 |
| SP500\_DIF | 0.098057 | 0.016627 | 5.897412 | 0.0000 |
|  |  |  |  |  |
|  |  |  |  |  |
| R-squared | 0.107750 | Mean dependent var | | 0.772759 |
| Adjusted R-squared | 0.104652 | S.D. dependent var | | 15.42057 |
| S.E. of regression | 14.59138 | Akaike info criterion | | 8.205611 |
| Sum squared resid | 61317.59 | Schwarz criterion | | 8.230921 |
| Log likelihood | -1187.814 | Hannan-Quinn criter. | | 8.215752 |
| F-statistic | 34.77946 | Durbin-Watson stat | | 1.942152 |
| Prob(F-statistic) | 0.000000 |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |

-> E(ut) = 0 – serial correlation? check residuals

-> E(ut2) = σ2 – heteroskedasticity? check residuals

-> E(ut,ut-1) = 0 – independent errors? check residuals  
-> E(xt,ut) = 0 – plot independent variable against residuals

**Regression #2: first-difference & logs data**

d(log growth gold) = 0.001692 + 0.197068 d(log growth sp500) + t

(0.079284) (0.024580)

R2 =0.187619

SER = 1.314763

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Dependent Variable: GOLD\_LOG\_GROWTH\_1DIF | | | | |
| Method: Least Squares | | |  |  |
| Date: 05/09/20 Time: 18:44 | | | |  |
| Sample (adjusted): 2/04/2019 5/01/2020 | | | |  |
| Included observations: 275 after adjustments | | | | |
|  |  |  |  |  |
|  |  |  |  |  |
| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|  |  |  |  |  |
|  |  |  |  |  |
| C | 0.001692 | 0.079284 | 0.021340 | 0.9830 |
| SP500\_LOG\_GROWTH\_1DIF | 0.197068 | 0.024580 | 8.017483 | 0.0000 |
|  |  |  |  |  |
|  |  |  |  |  |
| R-squared | 0.190584 | Mean dependent var | | 0.004458 |
| Adjusted R-squared | 0.187619 | S.D. dependent var | | 1.458705 |
| S.E. of regression | 1.314763 | Akaike info criterion | | 3.392436 |
| Sum squared resid | 471.9084 | Schwarz criterion | | 3.418740 |
| Log likelihood | -464.4600 | Hannan-Quinn criter. | | 3.402993 |
| F-statistic | 64.28004 | Durbin-Watson stat | | 3.027360 |
| Prob(F-statistic) | 0.000000 |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |

Hypothesis test:

H0: β = 0 (coefficient is not significantly different from zero)  
H1: β ≠ 0 (coefficient is significantly different from zero)

-> P < 0.0000  
-> Reject Null.   
-> β is significantly different from zero at β = 0.19 at 95% level

SER:

-> first difference of log = percentage change  
-> SER = 1.314763, or 1.3%

-> SER must be ≤ 2.5% (i.e. 95% CI) to produce a narrow enough interval

**Regression #3: dlog(data)**

dlog(gold) = 0.000519 + 0.183624 dlog(sp500) + t

(0.000556) (0.029044)

R2 =0.121871

SER = 0.009477

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Dependent Variable: GOLD\_DLOG | | | |  |
| Method: Least Squares | | |  |  |
| Date: 05/10/20 Time: 18:16 | | | |  |
| Sample (adjusted): 2/04/2019 5/01/2020 | | | |  |
| Included observations: 290 after adjustments | | | | |
|  |  |  |  |  |
|  |  |  |  |  |
| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|  |  |  |  |  |
|  |  |  |  |  |
| C | 0.000519 | 0.000556 | 0.933022 | 0.3516 |
| SP500\_DLOG | 0.183624 | 0.029044 | 6.322179 | 0.0000 |
|  |  |  |  |  |
|  |  |  |  |  |
| R-squared | 0.121871 | Mean dependent var | | 0.000540 |
| Adjusted R-squared | 0.118822 | S.D. dependent var | | 0.010095 |
| S.E. of regression | 0.009477 | Akaike info criterion | | -6.473106 |
| Sum squared resid | 0.025864 | Schwarz criterion | | -6.447797 |
| Log likelihood | 940.6004 | Hannan-Quinn criter. | | -6.462966 |
| F-statistic | 39.96995 | Durbin-Watson stat | | 1.925919 |
| Prob(F-statistic) | 0.000000 |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |

Hypothesis test:

H0: β = 0 (coefficient is not significantly different from zero)  
H1: β ≠ 0 (coefficient is significantly different from zero)

-> P < 0.0000  
-> Reject Null.   
-> β is significantly different from zero at β = 0.18 at 95% level

SER:

-> first difference of log = percentage change  
-> SER = 0.009477, or 0.9%

-> SER must be ≤ 2.5% (i.e. 95% CI) to produce a narrow enough interval.

**Interpretation**

**S&P500 goes up 1%, gold rises by 0.18%.**

During COVID-19 crisis (12/08/2019–), gold became positively correlated with the S&P 500 composite index.

Compare to other market crises:  
– European sovereign debt  
– 2008 financial crisis

**Notes**

**Tomorrow:**

* Correlation analysis first, before regression.
* From scatterplot, check r and confirm by hypothesis testing (covariance analysis).
* *Then* you can construct a regression model – see literature.
* Check assumptions – normality, serial correlation, heteroskedasticity.
* Conclude: has the relationship between gold and the stock market changed?
* **Repeat analysis strictly for COVID-19 pandemic dates, 12/08/2019–today, AND/OR dates around the stock market 2020 crash – 2/20/2020-3/23/2020, “Worst down day since 1987” (****March 16, 2020) + FED cut rates to 0%.** Worst down day since 1987
* **Check whether my interpretation of the Null hypothesis is right**

Notes: Idea for reasoning

Gold has been traditionally viewed as a safe haven. Lucey (2010) defines a safe haven “as an asset that is uncorrelated or negatively correlated with another asset or portfolio in times of market stress or turmoil”. As such, we would expect the demand for gold to increase during a market crisis, driving up its price. Thus, as the price of the market falls, the price of gold rises – in other words, a negative relationship. This has been observed in practice, at least during a short period at the beginning of a market crash. However, in the recent COVID-19 crisis, we have seen a change in the relationship between the stock market and gold price: as the market index fell, so did the price of gold:

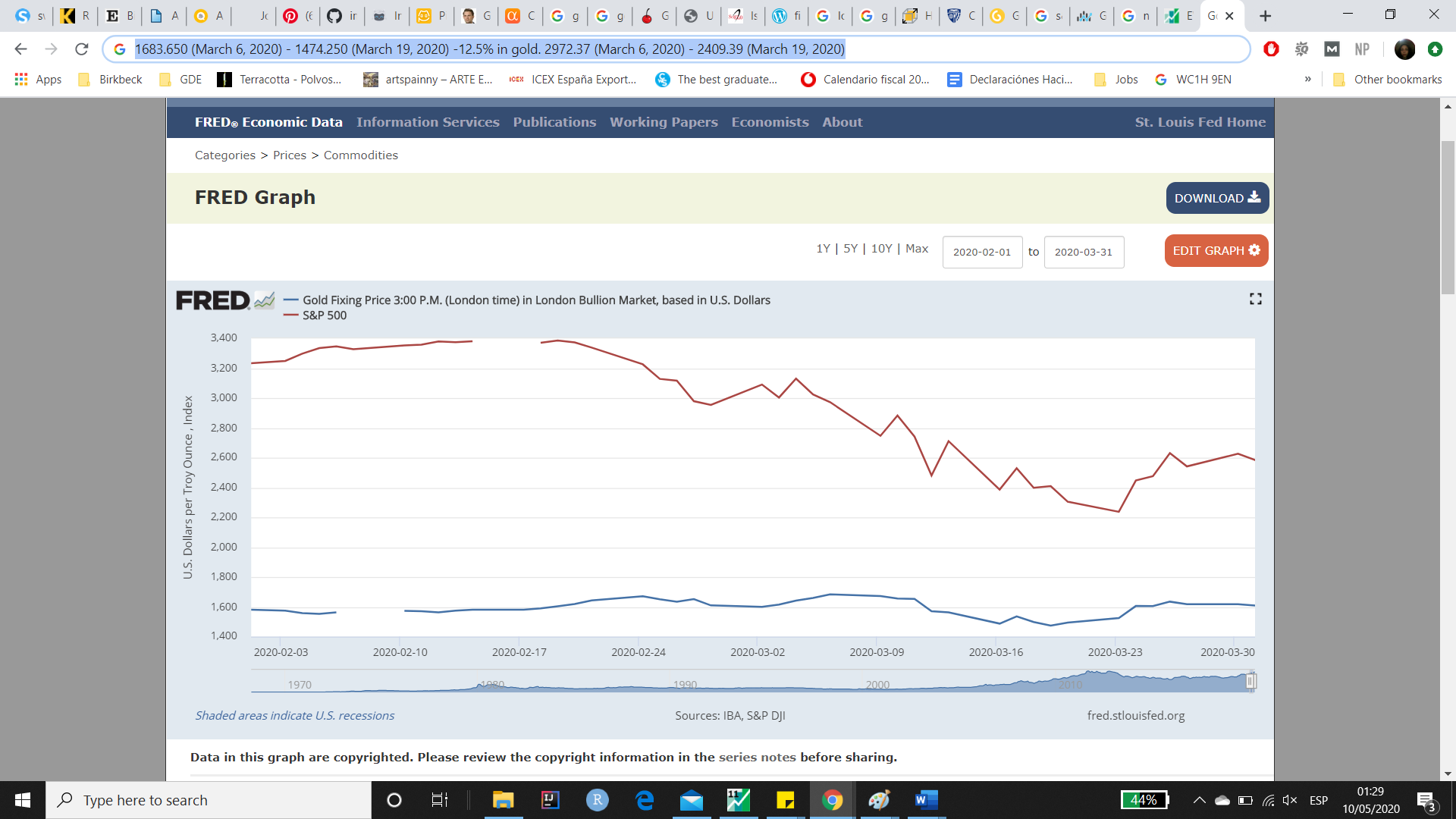
* 1683.650 (March 6, 2020) - 1474.250 (March 19, 2020)

-12.5% in gold

* 2972.37 (March 6, 2020) - 2409.39 (March 19, 2020)

-18.9% in S&P 500

Source: Federal Reserve Bank of St. Louis



Historically (1968 to today), no correlation -> **show**. Possibly correlation only during certain times, i.e. under market stress or turmoil. Re: COVID-19 crisis, European debt crisis, 2008 financial crisis, etc. Otherwise, in times of good market stability, gold price determined by other factors (interest rate, exchange rate, CPI and oil).

May 10

1. Repeat yesterday’s analysis strictly for COVID-19 pandemic dates, 12/08/2019–today, AND/OR dates around the stock market 2020 crash – 2/20/2020-3/23/2020, “Worst down day since 1987” (March 16, 2020) + FED cut rates to 0%
2. Check whether my interpretation of the Null hypothesis (covariance analysis) is right

**Normality:** Residuals don’t have to be normally distributed when k > 30. The CLT states that errors will be approximately normal when k > 30. With large enough sample sizes, the only requirement for OLS is independent values (no serial correlation) and homoskedasticity (i.e. residuals that have constant variance/are not heteroskedastic):

* remove serial correlation by detrending data.
* If heteroskedasticity is present and cannot be removed, use robust standard errors (HAC standard errors, heteroskedasticity-consistent, sandwich, Huber-Eicher-White estimates)

<https://stats.stackexchange.com/questions/29731/regression-when-the-ols-residuals-are-not-normally-distributed?fbclid=IwAR2au4u7N7HNMoL9moWlImWBjF0GnOIBUOFK5oWz-GLJAvYkl-98nWdIcWA>

-> descriptive statistics section.  
-> potential title – “Properties of OLS”.

**Model:** Specific-to-general:

* Specific: only S&P 500
* General: add variables previously proven to correlate with gold price to see when S&P 500 is an indicator (in years of market crisis) and when it is not (bull markets).

**Introduction re: gold**

“Gold is considered a commodity, but is unique in that it’s supply is more or less fixed, unlike other commodities, (oil, natural gas, etc.). As a result, any changes in its demand are not driven by its supply, and tend to be the result of changes in consumer and investor preferences. Another unique property of gold is that it used to be widely used as currency, before being used as a indicator of value for paper currency. For this reason, gold has an implicit value attached to it and is often considered a safe haven in times of market turmoil and a hedge against …”.

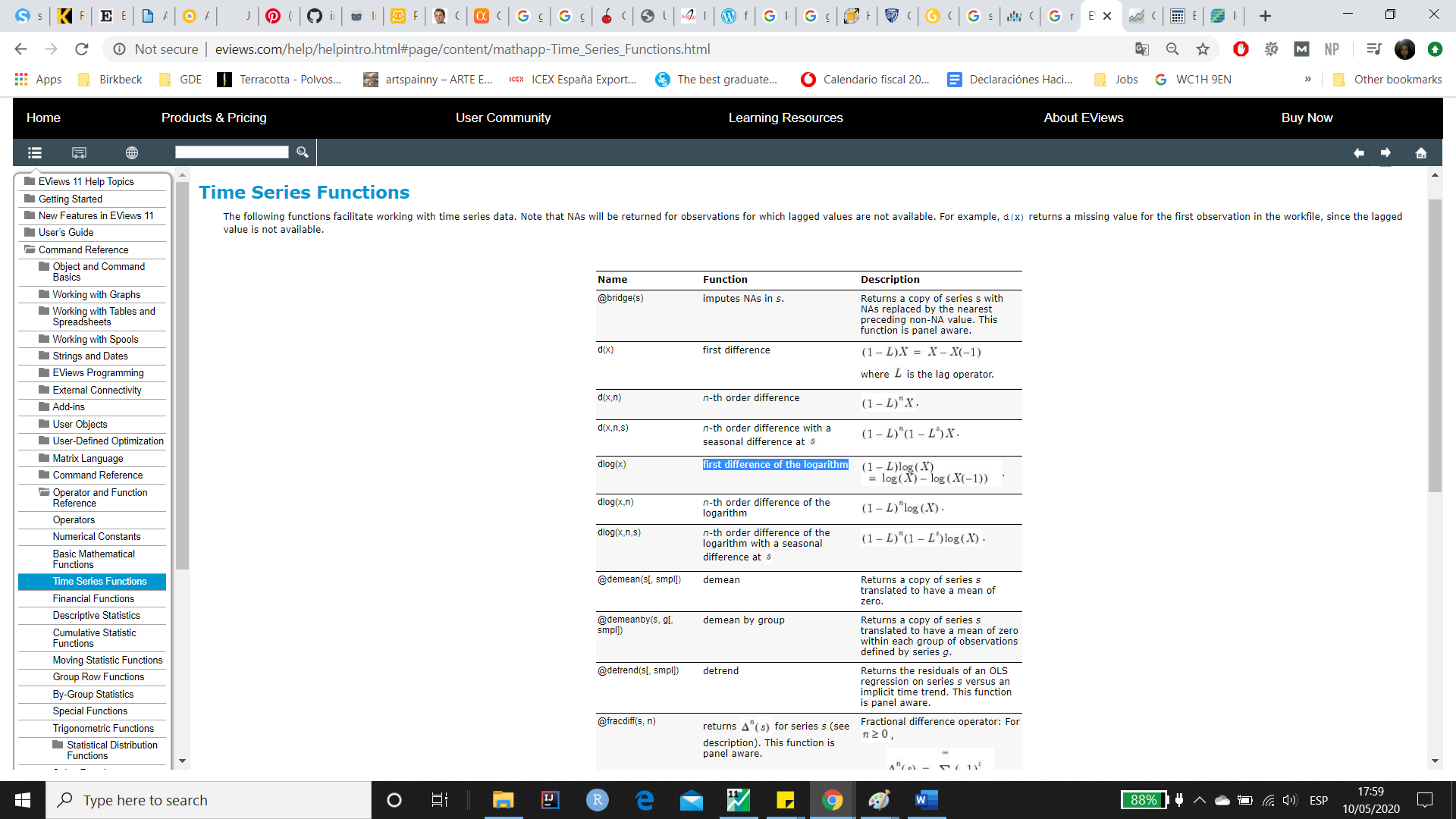
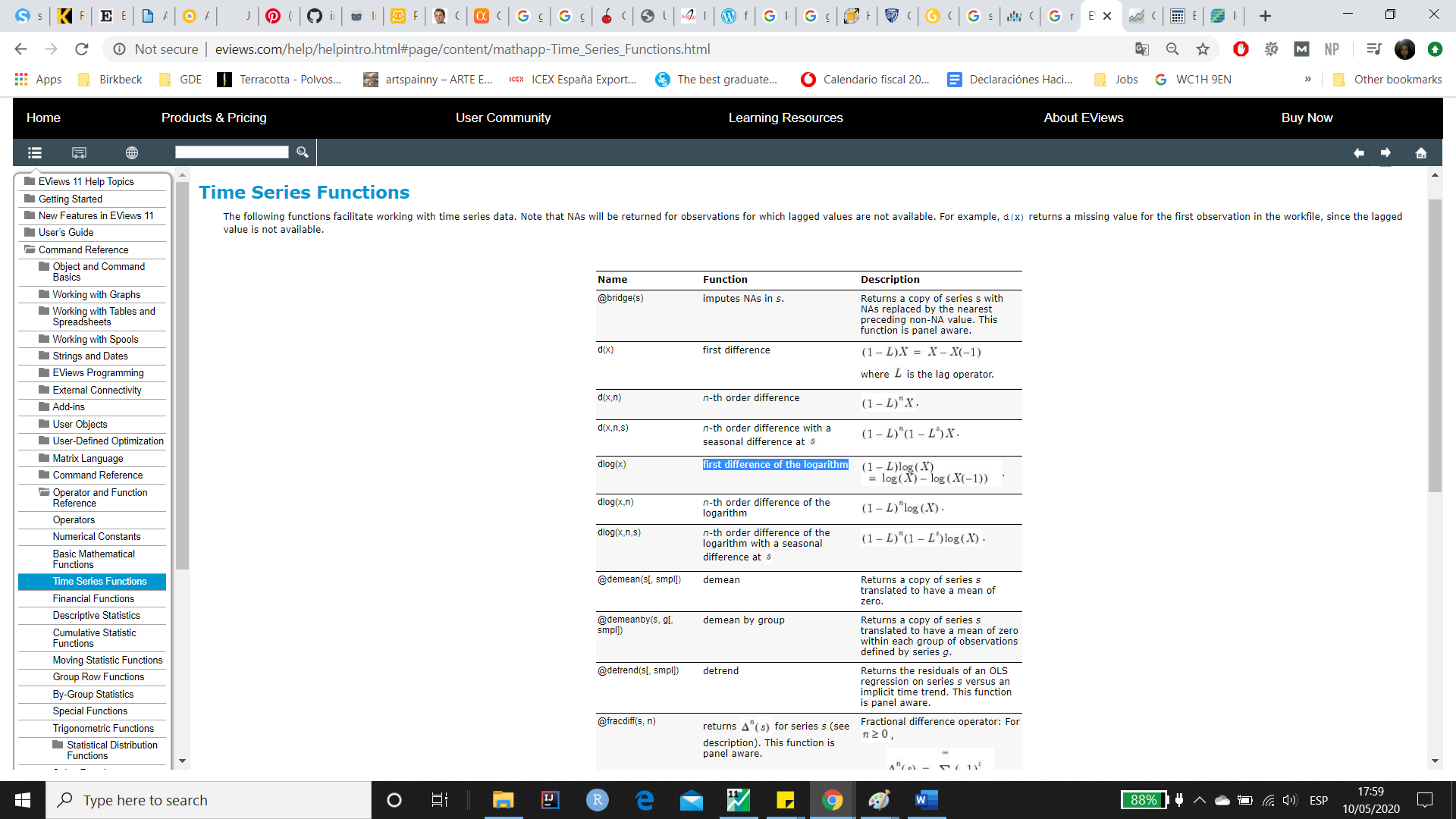
**Transformations**

**Log difference**

* Take logs to get a series with a linear trend
* Then difference to get a stationary series

<https://stats.stackexchange.com/questions/221780/are-log-difference-time-series-models-better-than-growth-rates/293854>

**Others**



**Independence:**

Corr(x,u)



|  |  |  |  |
| --- | --- | --- | --- |
| Covariance Analysis: Ordinary | | |  |
| Date: 05/10/20 Time: 18:50 | | |  |
| Sample: 2/04/2019 5/01/2020 | | |  |
| Included observations: 290 | | |  |
| Balanced sample (listwise missing value deletion) | | | |
|  |  |  |  |
|  |  |  |  |
| Covariance | |  |  |
| Correlation | |  |  |
| t-Statistic | |  |  |
| Probability | SP500\_DLOG | RESID04\_DLOG |  |
| SP500\_DLOG | 0.000367 |  |  |
|  | 1.000000 |  |  |
|  | ----- |  |  |
|  | ----- |  |  |
|  |  |  |  |
| RESID04\_DLOG | -1.28E-20 | 8.92E-05 |  |
|  | **-7.08E-17** | 1.000000 |  |
|  | **-1.20E-15** | ----- |  |
|  | 1.0000 | ----- |  |
|  |  |  |  |
|  |  |  |  |

Hypothesis test:

H0: r = 0 (correlation is not significantly different from zero)  
H1: r ≠ 0 (correlation is significantly different from zero)

-> P = 1.0000  
-> P > 0.05   
-> Do not reject Null.   
-> Independent variable & residuals are not significantly correlated at at 95% level.

**\*Before COVID-19 pandemic (5 months):**

**July to December 2019 (7/08/2019–12/06/2019)**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Dependent Variable: GOLD\_DLOG | | | |  |
| Method: Least Squares | | |  |  |
| Date: 05/10/20 Time: 20:59 | | | |  |
| Sample: 7/08/2019 12/06/2019 | | | |  |
| Included observations: 102 | | |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|  |  |  |  |  |
|  |  |  |  |  |
| C | 0.000343 | 0.000767 | 0.447157 | 0.6557 |
| SP500\_DLOG | -0.270342 | 0.091065 | -2.968662 | 0.0037 |
|  |  |  |  |  |
|  |  |  |  |  |
| R-squared | 0.080992 | Mean dependent var | | 0.000208 |
| Adjusted R-squared | 0.071802 | S.D. dependent var | | 0.008025 |
| S.E. of regression | **0.007732** | Akaike info criterion | | -6.867522 |
| Sum squared resid | 0.005978 | Schwarz criterion | | -6.816052 |
| Log likelihood | 352.2436 | Hannan-Quinn criter. | | -6.846680 |
| F-statistic | 8.812953 | Durbin-Watson stat | | 2.202558 |
| Prob(F-statistic) | 0.003744 |  |  |  |
|  |  |  |  |  |

Significantly negatively correlated: As market grows, gold falls.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Breusch-Godfrey Serial Correlation LM Test: | | | | |
| Null hypothesis: No serial correlation at up to 1 lag | | | | |
|  |  |  |  |  |
|  |  |  |  |  |
| F-statistic | 0.840912 | Prob. F(1,99) | | 0.3614 |
| Obs\*R-squared | 0.859097 | Prob. Chi-Square(1) | | 0.3540 |
|  |  |  |  |  |

**\*During COVID-19 pandemic (5 months):**

**December 2019 to May 2020 (12/09/2019–5/01/2020)**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Dependent Variable: GOLD\_DLOG | | | |  |
| Method: Least Squares | | |  |  |
| Date: 05/10/20 Time: 20:53 | | | |  |
| Sample: 12/09/2019 5/01/2020 | | | |  |
| Included observations: 89 | | |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|  |  |  |  |  |
|  |  |  |  |  |
| C | 0.001177 | 0.001305 | 0.901740 | 0.3697 |
| SP500\_DLOG | 0.228487 | 0.040012 | 5.710454 | 0.0000 |
|  |  |  |  |  |
|  |  |  |  |  |
| R-squared | 0.272632 | Mean dependent var | | 0.000822 |
| Adjusted R-squared | 0.264271 | S.D. dependent var | | 0.014334 |
| S.E. of regression | **0.012295** | Akaike info criterion | | -5.937036 |
| Sum squared resid | 0.013151 | Schwarz criterion | | -5.881112 |
| Log likelihood | 266.1981 | Hannan-Quinn criter. | | -5.914495 |
| F-statistic | 32.60928 | Durbin-Watson stat | | 1.817931 |
| Prob(F-statistic) | 0.000000 |  |  |  |
|  |  |  |  |  |

Significantly positively correlated: As market falls, gold falls.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Breusch-Godfrey Serial Correlation LM Test: | | | | |
| Null hypothesis: No serial correlation at up to 1 lag | | | | |
|  |  |  |  |  |
|  |  |  |  |  |
| F-statistic | 1.456209 | Prob. F(1,86) | | 0.2308 |
| Obs\*R-squared | 1.481914 | Prob. Chi-Square(1) | | 0.2235 |

In a regression of daily gold returns on S&P returns over the second half of 2019 a 1% change in the S&P index causes a 0.23% change in the gold price that is significant (t = 5.66), with quite a small standard error of regression, 1.23%. There is no evidence of serial correlation. That is  quite an unusual result, if you look over the full period, the effect is much smaller.

**Appendix**

General

- covariance analysis of gold-S&P correlation (significant)

- normally distributed residuals (significantly normal)

Tests

- Serial correlation:

- Test for serial correlation (Durbin-Watson)

|  |  |
| --- | --- |
| Durbin-Watson stat | 1.817931 |

- DW <2

- **Slight positive serial correlation**

Lagrange Multiplier test:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Breusch-Godfrey Serial Correlation LM Test: | | | | |
| Null hypothesis: No serial correlation at up to 2 lags | | | | |
|  |  |  |  |  |
|  |  |  |  |  |
| F-statistic | 0.698378 | Prob. F(2,90) | | **0.5001** |
| Obs\*R-squared | 1.436539 | Prob. Chi-Square(2) | | 0.4876 |
|  |  |  |  |  |
|  |  |  |  |  |

- H0: no serial correlation

- p > 0.05 (p= 0.50)

- Fail to reject Null hypothesis, therefore no significant serial correlation at 95% CI.

- Homoskedasticity:

- Test for serial correlation (Breusch-Pagan-Godfrey test):

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Heteroskedasticity Test: Breusch-Pagan-Godfrey | | | | |
| Null hypothesis: Homoskedasticity | | | |  |
|  |  |  |  |  |
|  |  |  |  |  |
| F-statistic | 0.110990 | Prob. F(1,92) | | **0.7398** |
| Obs\*R-squared | 0.113266 | Prob. Chi-Square(1) | | 0.7365 |
| Scaled explained SS | 0.118196 | Prob. Chi-Square(1) | | 0.7310 |
|  |  |  |  |  |
|  |  |  |  |  |

H0: homoscedasticity (constant variance in errors)

p > 0.05 (p= 0.50)

Fail to reject Null hypothesis, therefore no significant homoskedasticity at 95% CI.

- ARCH test:

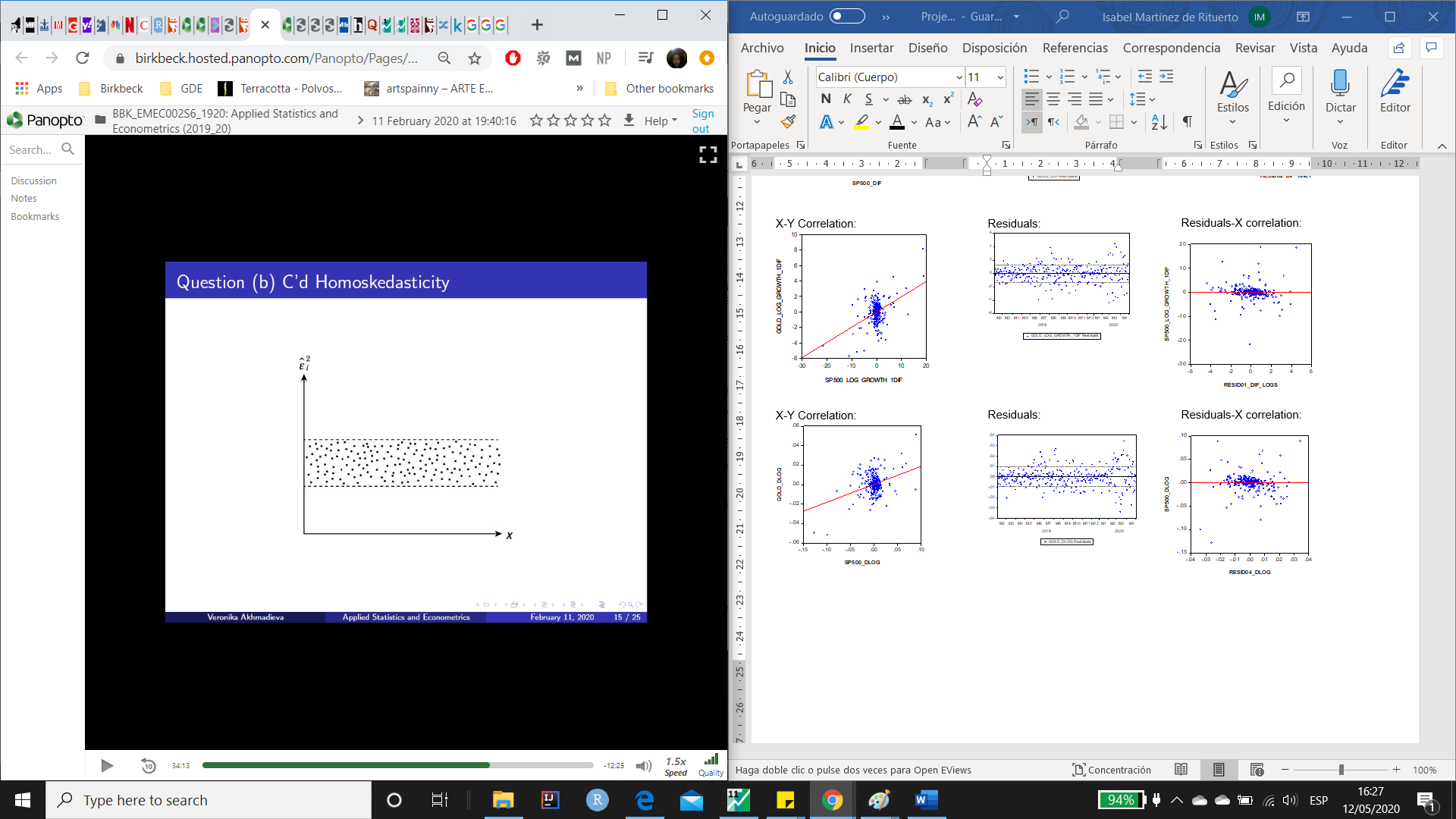
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Heteroskedasticity Test: ARCH | | | |  |
|  |  |  |  |  |
|  |  |  |  |  |
| F-statistic | 16.69910 | Prob. F(1,85) | | 0.0001 |
| Obs\*R-squared | 14.28549 | Prob. Chi-Square(1) | | 0.0002 |
|  |  |  |  |  |
|  |  |  |  |  |

H0: homoscedasticity (constant variance in errors)

p < 0.05 (p= 0.0001)

Reject Null hypothesis, therefore significant homoskedasticity at 95% CI.

- Residuals for homoskedasticity

* Solutions:
  + respecification
  + more variables: omitted variable problem
  + HAC standard errors: robust standard errors

- Exogeneity:

May 16

**Sections**

1. Abstract
2. Introduction
3. Data
4. Descriptive stats

* graphs gold
* graphs s&p500
* patterns or trends to explore
* before and after pandemic
  1. Transformation
* log graphs
* volatility
  1. Correlation matrix
* find relationship
* regression model

1. Regression
   1. Regression model

Present model equation and reason for this model.

* 1. Regression results

Present results in an equation and explain what each term means.

* 1. Intra-period stats

This table shows the estimation results for three different periods: a bull market from November 1995 until March 2000, a bear market from March 2000 until March 2003 and a bull market from March 2003 until November 2005.

1. Conclusions
2. Bibliography
3. Appendix

The residuals show some autocorrelation (negative serial correlation), most notably around late March. The raw data shows a drop in the S&P500 in the middle of the 1st quarter - in early March. Confirmed around March 16, 2020 (worst down day since 1987\*). Quandt-Andrews unknown breakpoint test suggested a break point at 3/12/2020 at the 5% level with with p = 0.0003. I split my data into subsamples to see whether this would improve the serial correlation observed in my previous sample.

**Breakpoint at 3/12/2020**

Quandt-Andrews unknown breakpoint test

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Null Hypothesis: No breakpoints within 15% trimmed data | | | | | | | |
| Varying regressors: All equation variables | | | | | | | |
| Equation Sample: **12/03/2019 5/01/2020** | | | | | | | |
| Test Sample: 12/23/2019 4/09/2020 | | | | | | | |
| Number of breaks compared: 66 | | | | | | | |
|  | | |  | |  |  | |
|  | | |  | |  |  | |
| Statistic | | | Value | |  | Prob. | |
|  | | |  | |  |  | |
|  | | |  | |  |  | |
| Maximum LR F-statistic (3/12/2020) | | | 11.35353 | |  | **0.0003** | |
| Maximum Wald F-statistic (3/12/2020) | | | 22.70706 | |  | 0.0003 | |
|  | | |  | |  |  | |
| Exp LR F-statistic | | | 2.558328 | |  | 0.0074 | |
| Exp Wald F-statistic | | | 7.558981 | |  | 0.0009 | |
|  | | |  | |  |  | |
| Ave LR F-statistic | | | 2.412474 | |  | 0.0428 | |
| Ave Wald F-statistic | | | 4.824948 | |  | 0.0428 | |
|  | | |  | |  |  | |
|  | | |  | |  |  | |
| Note: probabilities calculated using Hansen's (1997) method | | | | | | | |
| WARNING: estimation sample is non-continuous (probabilities | | | | | | | |
| calculated assuming a continuous sample) | | | | | | | |
| \*<https://www.ft.com/content/82c5c2ca-670e-11ea-800d-da70cff6e4d3>  Chow Breakpoint Test: 3/12/2020 | | | | | | |
| Null Hypothesis: No breaks at specified breakpoints | | | | | | | | |
| Varying regressors: All equation variables | | | | | | | | |
| Equation Sample: **12/03/2019 5/01/2020** | | | | | | |  | |
|  |  |  | |  | | |  | |
|  |  |  | |  | | |  | |
| F-statistic | 11.35353 |  | | Prob. F(2,89) | | | 0.0000 | |
| Log likelihood ratio | 21.13365 |  | | Prob. Chi-Square(2) | | | 0.0000 | |
| Wald Statistic | 22.70706 |  | | Prob. Chi-Square(2) | | | 0.0000 | |
|  |  |  | |  | | |  | |

Multiple breakpoint tests

|  |  |  |  |
| --- | --- | --- | --- |
| Bai-Perron tests of L+1 vs. L sequentially determined breaks | | | |
| Date: 05/13/20 Time: 23:30 | | |  |
| Sample: 12/02/2019 5/01/2020 | | |  |
| Included observations: 93 | | |  |
| Breaking variables: C SP500\_DEC2019\_DLOG | | | |
| Break test options: Trimming 0.15, Max. breaks 5, Sig. level | | | |
| 0.05 | |  |  |
|  |  |  |  |
|  |  |  |  |
| Sequential F-statistic determined breaks: | | | 1 |
|  |  |  |  |
|  |  |  |  |
|  |  | Scaled | Critical |
| Break Test | F-statistic | F-statistic | Value\*\* |
|  |  |  |  |
|  |  |  |  |
| 0 vs. 1 \* | 11.35353 | 22.70706 | 11.47 |
| 1 vs. 2 | 1.783197 | 3.566395 | 12.95 |
|  |  |  |  |
|  |  |  |  |
| \* Significant at the 0.05 level. | | |  |
| \*\* Bai-Perron (Econometric Journal, 2003) critical values. | | | |
|  |  |  |  |
| Break dates: | |  |  |
|  | Sequential | Repartition |  |
| 1 | 3/12/2020 | 3/12/2020 |  |
|  |  |  |  |
|  |  |  |  |

Split data into subsamples:

1. From 2019-2020

**12/02/2019 3/12/2020**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Dependent Variable: GOLD\_DLOG | | | |  |
| Method: Least Squares | | |  |  |
| Date: 05/14/20 Time: 15:41 | | | |  |
| Sample: 12/02/2019 3/12/2020 | | | |  |
| Included observations: 61 | | |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|  |  |  |  |  |
|  |  |  |  |  |
| C | 0.001069 | 0.001342 | 0.796919 | 0.4287 |
| SP500\_DLOG | 0.111051 | 0.057300 | 1.938063 | 0.0574 |
|  |  |  |  |  |
|  |  |  |  |  |
| R-squared | 0.059852 | Mean dependent var | | 0.000617 |
| Adjusted R-squared | 0.043917 | S.D. dependent var | | 0.010555 |
| S.E. of regression | 0.010321 | Akaike info criterion | | -6.277059 |
| Sum squared resid | 0.006285 | Schwarz criterion | | -6.207850 |
| Log likelihood | 193.4503 | Hannan-Quinn criter. | | -6.249936 |
| F-statistic | 3.756089 | Durbin-Watson stat | | 1.666770 |
| Prob(F-statistic) | 0.057405 |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |

**3/12/2020 5/01/2020**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Dependent Variable: GOLD\_DLOG | | | |  |
| Method: Least Squares | | |  |  |
| Date: 05/14/20 Time: 15:41 | | | |  |
| Sample: 3/12/2020 5/01/2020 | | | |  |
| Included observations: 34 | | |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|  |  |  |  |  |
|  |  |  |  |  |
| C | -0.000595 | 0.002524 | -0.235696 | 0.8152 |
| SP500\_DLOG | 0.321779 | 0.055083 | 5.841700 | 0.0000 |
|  |  |  |  |  |
|  |  |  |  |  |
| R-squared | 0.516071 | Mean dependent var | | -0.000480 |
| Adjusted R-squared | 0.500949 | S.D. dependent var | | 0.020836 |
| S.E. of regression | 0.014720 | Akaike info criterion | | -5.542257 |
| Sum squared resid | 0.006933 | Schwarz criterion | | -5.452471 |
| Log likelihood | 96.21837 | Hannan-Quinn criter. | | -5.511638 |
| F-statistic | 34.12546 | Durbin-Watson stat | | 1.529989 |
| Prob(F-statistic) | 0.000002 |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |

2. From 2010-2020:

**2/01/2010 2/28/2020**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Dependent Variable: GOLD\_DLOG | | | |  |
| Method: Least Squares | | |  |  |
| Date: 05/14/20 Time: 15:39 | | | |  |
| Sample (adjusted): 2/04/2019 2/28/2020 | | | |  |
| Included observations: 248 after adjustments | | | | |
|  |  |  |  |  |
|  |  |  |  |  |
| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|  |  |  |  |  |
|  |  |  |  |  |
| C | 0.000654 | 0.000475 | 1.377371 | 0.1696 |
| SP500\_DLOG | -0.169971 | 0.055841 | -3.043849 | 0.0026 |
|  |  |  |  |  |
|  |  |  |  |  |
| R-squared | 0.036296 | Mean dependent var | | 0.000589 |
| Adjusted R-squared | 0.032378 | S.D. dependent var | | 0.007595 |
| S.E. of regression | 0.007471 | Akaike info criterion | | -6.947538 |
| Sum squared resid | 0.013731 | Schwarz criterion | | -6.919204 |
| Log likelihood | 863.4947 | Hannan-Quinn criter. | | -6.936131 |
| F-statistic | 9.265015 | Durbin-Watson stat | | 1.938563 |
| Prob(F-statistic) | 0.002589 |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |

**2/28/2020 5/01/2020**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Dependent Variable: GOLD\_DLOG | | | |  |
| Method: Least Squares | | |  |  |
| Date: 05/14/20 Time: 15:39 | | | |  |
| Sample: 2/28/2020 5/01/2020 | | | |  |
| Included observations: 43 | | |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|  |  |  |  |  |
|  |  |  |  |  |
| C | 6.88E-05 | 0.002386 | 0.028830 | 0.9771 |
| SP500\_DLOG | 0.257028 | 0.052562 | 4.890026 | 0.0000 |
|  |  |  |  |  |
|  |  |  |  |  |
| R-squared | 0.368379 | Mean dependent var | | -0.000355 |
| Adjusted R-squared | 0.352974 | S.D. dependent var | | 0.019437 |
| S.E. of regression | 0.015634 | Akaike info criterion | | -5.433281 |
| Sum squared resid | 0.010022 | Schwarz criterion | | -5.351365 |
| Log likelihood | 118.8155 | Hannan-Quinn criter. | | -5.403073 |
| F-statistic | 23.91236 | Durbin-Watson stat | | 1.628652 |
| Prob(F-statistic) | 0.000016 |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |