**Empirical analysis of flight-to-quality from stocks into gold in US during economic recessions (1969-2020)**

Isabel Martínez de Rituerto

**Word count:** 5000 words

**Program:** R

**Abstract**

**Background:** Within financial markets, gold has long been believed to as a store for value during financial crises. In fact, gold is widely viewed as a hedge against volatility in investment portfolios and a financial safe haven in times of market turmoil. However, despite anecdotal evidence and the financial media suggesting that gold serves as a safe haven in financial markets during times of economic instability, empirical evidence supporting this claim is inconsistent.

**Methods:** This project seeks to evaluate the behaviour of gold prices within the world’s largest economy, the United States, during periods of economic recession. Historical gold and US stock market prices are analyzed to gain insights into the behaviour of investors with respect to the precious metal during periods of financial crisis. The National Bureau of Economics’ (NBER) definition of a recessionary period is used to select 8 periods of recession within the USA: 1969, 1973, 1980, 1981, 1990, 2001, 2007, and 2020. Descriptive and econometric analyses are then used to identify potential relationships between investment returns from gold and S&P500 during recessionary periods in the US spanning a period of 50 years.

**Results:** …

**Conclusion:** Understanding the relationship between financial assets and the economic climate is of significance to investors, policy makers and governments alike. The novelty of this project lies in the analysis of trends in gold price over the whole of the 20th century, covering a total of 8 periods of economic recession. As a result, while my findings indicate that gold does act as both a hedge and a safe haven in the US stock market at times, this did not hold across all of the periods studied. This suggests that the stocks‒gold relationship is not a linear one and explains why many previous studies have published conflicting results. Further analyses should seek to explore whether this variability holds in other markets, including other European markets, as well as large emerging markets, such as the BRIC countries, and in developing economies to further elucidate the role of gold, and shed further light on investor herd behaviour.

Gold is widely used as a hedge against volatility in investment portfolios and a financial safe haven in times of market turmoil. This suggests that gold is inherently negatively correlated with market performance in times of crisis, and uncorrelated otherwise. This projects uses daily stock markets returns and gold prices to construct a linear regression model to evaluate the relationship between gold and US market returns in the short-run. The recent COVID-19 pandemic and the 2008 financial crisis are used to evaluate how gold prices have behaved during recent episodes of financial market instability and whether these changes coincide with its popular denomination as a safe haven and its use as a diversification instrument.

**1. Introduction**

The recent COVID-19 crisis has created an opportunity to analyze relationships within financial markets during economic crises.

Gold is widely believed to act as a hedge for inflation and as a store for value during times of financial crisis, when returns from stock indices tend to turn negative. This suggests that the price of gold is negatively correlated with stock market performance, acting as a hedge against volatility and a financial safe haven in times of market turmoil.

In the economic model of supply and demand, prices increase in two situations: when demand rises or when supply falls. Similar to other commodities, including oil and copper, the amount of gold available for mining is limited. However, unlike the former, gold is not consumed: it is either stored in reserves or used for jewellery, wherefrom it can be recycled to be re-introduced into supply.

The idea that gold is a valuable commodity has persisted over the centuries, from its use as a monetary currency in ancient Rome to its role backing the US dollar during 1940s and 1970s as a result of the Bretton Woods agreement. Unlike other commodities, such as oil, cotton, and copper, which are widely consumed as essential goods, demand for gold can be split into different categories: jewellery, technology, investment and central bank reserves (Figure 1). As of 2021, the latter categories accounted for approximately a quarter of the global gold demand.

[Gold use, intrinsic scarcity]

[Gold reserves] – banks and governments vouch for the value of gold

[Supply and demand model for gold] – draw for elastic and inelastic. Oil consumption is known to exhibit highly inelastic behaviour ()

Baur and Lucey (2010) define a safe haven “as an asset that is uncorrelated or negatively correlated with another asset or portfolio in times of market stress or turmoil” and a hedge as “a diversifier is defined as an asset that is positively (but not perfectly correlated) with another asset or portfolio on average”. Combined with economic theory, which suggests that prices rise with an increase in demand, we would expect the price of gold to rise during period of market turmoil and thereby be negatively correlated with market performance.

The National Bureau of Economic Research (NBER) defines a recession as “between a peak of economic activity and its subsequent trough, or lowest point”. By contrast, an expansion is defined as the time “between trough and peak”.

The S&P500 is one of the largest stock market indices in the world. It is primarily comprised of [industries] and has a value of [x]. It is the 2nd largest (?) stock market index in the USA, after [x].

Figure 1 depicts the historical relationship between the S&P500 and gold prices from 1968 to 2020.

Gráfico, Histograma

Descripción generada automáticamente

**Figure 1.** Historical prices of gold and S&P 500 (1968-2020)

Prices in both markets appear to have steadily increased over time, and are clearly higher at the end of the sample than in the beginning. The prices clearly overlap four times during this period: 1973, 1991, 2008, and 2013, coinciding with historical financial crises, including the 1973 oil crisis, the run-up to the dotcom bubble in the 1990s, the 2008 financial crisis, and the European debt crisis.

These overlaps at more or less consistent intervals throughout history suggest that the US stock market and gold prices move in opposite directions during periods of economic instability. This suggests that a negative correlation exists between these two assets, and one that may only hold during the short-run, namely during times of market turmoil, be and not in the long-run. This observation coincides with Baur and Lucey’s definition of a safe haven.

However, taking a closer look at the prices of both assets over time also seems to suggest that the relationship between gold and S&P500 during financial crises is not a static. As an example, Figure 2provides an expanded view of S&P500 and gold prices during the two most recent periods of economic recession: the 2008 financial crisis (2007-2009; Figure 2a) and the COVID-19 pandemic (2020; Figure 2b).

**Gráfico, Histograma

Descripción generada automáticamente**Gráfico, Histograma

Descripción generada automáticamente

**Figure 2.** Prices of gold (USD/oz) and S&P500 (USD) during: (a) 2008 financial crisis and (b) recent COVID-19 pandemic.

**Figure 2A** shows the relationship between gold and the S&P during the end of the 2008 financial crisis. There appears to initially be a positive correlation between the two variables, both losing 26% and 17% in value during October 2008, respectively (**Table 1**). However, this correlation appears to be negative from November 2008 onwards as the price of gold appears to rise while the S&P 500 continues to fall.

**Figure 2B** shows the relationship between gold and the S&P during the beginning of the ongoing COVID-19 pandemic. Prior to the pandemic, the prices of gold and the S&P 500 were both rising. However, both the S&P 500 and gold prices fell sharply in late February and early March, losing 33% and 5% of their value in one month, respectively (**Table 1**).

These two figures alone highlight the differences, suggesting that the relationship between gold and the S&P500 index varies during recessions. This could be attributed to the different causes of each recessions, as well as many other factors not included in this model.

This project will use data on daily gold and S&P 500 returns to evaluate the relationship between the price of gold and the US stock market during eight episodes of market turmoil in the 20th century, including the 2008 financial crisis and the recent COVID-19 pandemic. My aim is to determine whether the attribute of gold as a safe haven holds in financial crises in the short-run.

In section 2, information on the data used in my analysis is provided. Section 3 presents an overview of the historical relationship between gold and US stock market prices. Section 4 evaluates the statistical correlation between daily gold and stock market returns in two periods of market turmoil. The results of this correlation analysis are then used to construct a linear regression model in section 5. This project concludes with section 6, in which I summarize my findings. An appendix at the end provides statistical details in the form of graphs and tables not included in the main text.

**2. Literature review**

**2.1 Flight-to-quality**

Flight-to-quality is a phenomenon observed in financial markets, whereby investors experience a pronounced and generally rapid increase in risk aversion. The mechanism of flight-to-quality involves divesting from one asset and re-investing it into another to reduce exposure to risk during periods of market volatility. This is often observed in a shift of investments from the stock market to bonds.

In theory, at the level of the investor, the period of time between the transfer of money from a risky asset to a safer asset can vary, from taking place overnight, to days, weeks, or months depending on the investor’s risk appetite – at this level, there is an inherent trade-off of holding onto cash and waiting for an opportunity to re-invest versus reinvesting directly into another asset. However, at the aggregate level, we can expect investors to prefer

Lagged variables - According to the Capital Market theory, the stock price always reflect all available information in the market.

2-3 studies

**2.2 Gold in financial markets**

In addition to bonds, gold has traditionally been used as a safe haven during times of market volatility.

2-3 studies

**2.3 Relationship between gold and stock market**

2-3 studies

What methods did they use?

What data did they use?

**2.4 Relationship between gold and stock markets in USA**

Stock markets: S&P500, Dow Jones, Nasdaq

1 study

What methods did they use?

What data did they use?

Random walk, frequencies. According to the theory of random walk.

To find the market reaction to a specific event (i.e. Delta Airlines acquiring Northwest Airlines), researchers often include the day before the merger (-1,0) to capture the fact that such news sometimes leaks and the market starts reacting early.

Daily or high-frequency data may be affected by drifts and noise that could mask the dependence relationship and complicate modelling of the marginal distributions through nonstationary variances, long memory or sudden jumps. Our results were insensitive to the choice of a daily or weekly frequency

2.5 Summary and research hypothesis

Conflicting results

Arguments in relation to the ownership of gold:

• gold provides an inflation hedge

• gold serves as a currency hedge

• gold is an attractive alternative to assets with low real returns

• gold a safe haven in times of stress

• gold should be held because we are returning to a de facto world gold standard

• gold is “underowned”

Based on the literature, I will be testing two hypotheses:

**H1:** if safe haven, a (lagged?) negative correlation between stock and gold returns during a financial crisis

**H2:** if hedge against volatility, no correlation between stock and gold returns during a financial crisis

**3. Data**

**3.1. US recessions**

This project will analyse data within recessionary periods only. Here, eight recessionary periods were selected for the analysis of gold and stock price behaviour, as defined by the National Bureau of Economic Research (NBER). A graphic representation of these periods is shown in **Figure x**.

Gráfico, Gráfico de líneas, Histograma

Descripción generada automáticamente

**\***

**\***

**\***

**\***

**\***

**\***

**\***

**Figure x. Business cycle dating.** Representation of recessionary periods (gray). Here, the unemployment rate in the US is plotted to show the relationship between rising and falling rates of unemployment during and between economic

As shown in Figure x, it is clear that the most recent recession, corresponding to the COVID-19 pandemic starting in 2020, was the shortest (43 days). The NBER addresses this anomaly by stating that:

“*In the case of the February 2020 peak in economic activity, the committee concluded that the subsequent drop in activity had been so great and so widely diffused throughout the economy that, even if it proved to be quite brief, the downturn should be classified as a recession*.”

Despite this anomaly, all periods are sufficiently long for the purposes of linear statistical modeling as a minimum of 30 data points is needed to ensure normality when using certain statistical models (e.g. regression analysis).

**3.2. Gold price**

The data series used for gold are daily observations on the price of gold, quoted in US dollars per troy ounce. The gold price is the Gold Fixing Price 3:00 P.M. (London time) quote by the Federal Reserve Economic Data database. Prices from 1 December 1969 toMarch 31 2020, a total of 1835 observations for gold prices.

**3.3. S&P500 price**

The data series used for the S&P500 stock index are daily observations on the price of the index, quoted in US dollars. The stock price is the closing price of the S&P 500 quoted by the Yahoo! Finance database. Prices from 1 December 1969 toMarch 31 2020, a total of 1835 observations for the S&P500 stock index.

**3.4. Summary of data**

A summary of the number of data points corresponding to each recessionary evaluated, including daily prices and daily, weekly, and monthly returns, is given in Table x.

**Table x.** Data of the gold and S&P 500 subsamples

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Year** | **Start and end dates** | **Missing data points** | | **Daily prices** | **Daily returns** | **Weekly returns** | **Monthly returns** |
| Gold | S&P500 |
| **2020** | 3 March 2020  1 April 2020 | 1 | 2 | 44 | 43 | 8 | 2 |
| **2007** | 3 December 2007  1 June 2009 | 30 | 30 | 391 | 390 | 72 | 18 |
| **2001** | 1 March 2001  1 November 2001 | 9 | 13 | 176 | 175 | 32 | 8 |
| **1990** | 2 July 1990  1 March 1991 | 9 | 12 | 175 | 174 | 32 | 8 |
| **1981** | 1 July 1981  1 November 1982 | 22 | 20 | 350 | 349 | 64 | 16 |
| **1980** | 2 January 1980  1 July 1980 | 8 | 7 | 132 | 131 | 24 | 6 |
| **1973** | 1 November 1973  3 March 1975 | 19 | 24 | 349 | 348 | 64 | 16 |
| **1969** | 1 December 1969  2 November 1970 | 13 | 12 | 242 | 241 | 44 | 11 |

**4. Research design**

**4.1. Data pre-processing**

**4.1.1. Missing data**

As expected, due to the continuous nature of my data and the large time span of the datasets included in my analysis (n = 1853 daily data points), some data points were missing. Missing data can be addressed either by excluding or interpolating the missing data points. Here, there is an inherent trade-off between accuracy (as obtained by excluding missing data) and having enough data for statistical analysis (impossible when too much data is missing).

In the case of my dataset, the rate of missing data was on average low (~5%, i.e. 5 missing values for every 100 data points), which granted me some flexibility if I were to exclude data. However, because my analysis used different frequencies (e.g. daily, weekly, and monthly returns), although I had a large amount of data for the former (e.g. daily, with well above 30 observations), when the frequency of data was reduced to weekly and monthly data, every data point was crucial.

For this reason I chose to interpolate any missing data points using the averaging method:

…

Most often, one data point would be missing between two days. In this case, the average of the values above and below the missing data point was used to obtain a value. In cases with more than missing data point in consecutive days, the same method was used:

This is worth mentioning because at least 5% of the data points in my dataset are therefore not accurate, which could skew the results. However, because of the low rate of missing data, my interpolation of missing data points is unlikely to affect my results.

**4.1.2. Detrending**

As the statistical modeling of time series variables using regression requires stationary time series, I computed daily, monthly and weekly returns from prices using the following formula:

*Return = [P(t)-P(t-1))]/p(t-1)*

The resulting values were used for all subsequent statistical analyses.

**4.2. Regression analysis**

To evaluate the relationship between gold and stock market returns over time, the following univariate linear regression models were constructed:

*Daily gold returns = α + β (S&P500 daily returns) + ut (1)*

*Weekly gold returns = α + β (S&P500 weekly returns) + ut (2)*

*Monthly gold returns = α + β (S&P500 monthly returns) + ut (3)*

where *gold returns* denotes the daily returns of gold and *sp500 returns* denotes the daily returns of the S&P 500 index.

Gold was chosen as the dependent variable because…

S&P500 was chosen as the independent variable because…

A regression was run for each of the short-run subsamples, and a summary of the regression output is provided in the Results section (**Table 4)**.

The following sections discuss the test used to evaluate the assumptions required to use regression analysis:

(1) normality (if the sample size is large enough (n > 30), we can ignore the distribution of the data and use parametric tests);

(2) linearity (the relationship between X and the mean of Y is linear);

(3) multicollinearity;

(4) homoscedasticity (variance of residual is the same for all values of the variable);

(5) independence (no auto-correlation/serial correlation).

**4.2.1. Normality**

According to the standard of normal distribution proposed by Minglong (2009), the absolute value of the skewness coefficient should be less than 3, and the absolute value of the kurtosis coefficient should be less than 10. The closer the coefficient values of the skewness and kurtosis coefficients are to 0, the more they conform to a normal distribution.

Normality tests (Anderson-Darling, Kolmogorov-Smirnov, and Cramer von Mises tests) were performed using ETX 2.0.

As shown in Table 4, there was no significant difference in the standard deviation of all observed variables. A standard deviation value in the range of 0.544-0.881 can ensure high identification. The corresponding deviation coefficient was in the range of ‒1.294-0.297, while the kurtosis coefficient was in the range of ‒0.975-3.724, proving that the sample data approximately approach the normal distribution.

- Histograms

- Normality significance

**4.2.2. Linearity**

Test:

**4.2.3. Multicollinearity**

Test:

**4.2.4. Homoscedasticity**

Test:

**4.2.5. Independence**

Test:

Serial correlation – random walk.

**- do all tests for regression**

**- run regression using daily, week, monthly**

**- compare results using different frequencies**

**- compare to robustness tests (period returns)**

**5. Results**

“To investigate site-specific SSDs for 21 indigenous species (Table 3), site-specific acute toxicity values were predicted for each site used in the natural water toxicity test before deriving the PNECs of Korean freshwaters. The SSD curves for the five natural waters are shown in Fig. 3. The results of the normality test for all SSDs were accepted at four significance levels (0.01, 0.025, 0.05, and 0.1). The species sensitivity ranks from the predicted acute toxicity values for nickel that differed slightly for each site. Microalgae were most sensitive to nickel at the four sites, except for site B. Both bivalves (*L. fortunei*) and insects (*C. kiiensis*) had a higher tolerance to nickel than vertebrates in all site waters.”

“The PNECs using BLM-based ACR were positively correlated with the Ca and DOC concentrations (*r* = 0.69 Ca, *r* = 0.39 DOC, *p* < 0.001)”

**5.1. Prices and returns**

**5.1.1. Price: Maximum and minimum values**

Returns (percentage change) are useful for statistical analysis, as a detrended and normally distributed variable. However, a lot of detail is lost using this type of data. A better understanding of the degree of change in a given time period can be gained observing the maximum and minimum values, providing insights into price volatility, which can complement subsequent statistical analysis.

Table x summarizes the maximum and minimum values for the prices of gold and S&P500 index during the recessionary periods analyzed in this project.

**Table 1.** Min. and max. values for gold and S&P500 prices during eight recessionary periods.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Period** | **Variable** | **Min.** | **Max.** | **Change (%)** |
| **COVID-19 pandemic  (2020)** | **Gold** | 1474 | 1684 | 14% |
| **S&P 500** | 2237 | 3386 | 51% |
| **The Great Recession  (2007-2008)** | **Gold** | 712 | 1011 | 42% |
| **S&P 500** | 677 | 1516 | 124% |
| **Early 2000s recession (2001)** | **Gold** | 255 | 293 | 15% |
| **S&P 500** | 965 | 1312 | 36% |
| **Early 1990s recession  (1990)** | **Gold** | 351 | 414 | 18% |
| **S&P 500** | 295 | 369 | 25% |
| **Early 1980s recession II (1981)** | **Gold** | 297 | 481 | 62% |
| **S&P 500** | 102 | 139 | 36% |
| **Early 1980s recession I (1980)** | **Gold** | 481 | 850 | 77% |
| **S&P 500** | 98 | 118 | 21% |
| **End of Bretton Woods and oil crisis  (1973)** | **Gold** | 90 | 195 | 117% |
| **S&P 500** | 62 | 107 | 73% |
| **Recession of 1969–1970 (1969)** | **Gold** | 34 | 39 | 12% |
| **S&P 500** | 69 | 93 | 35% |

**5.1.2. Returns: Period returns**

**Table x.** Returns for S&P500 index and gold for entire eight recessionary periods.

|  |  |  |
| --- | --- | --- |
| **Period** | **Variable** | **Return (%)** |
| **COVID-19 pandemic**  **(2020)** | **Gold** | +2% |
| **S&P 500** | -20% |
| **The Great Recession**  **(2007-2008)** | **Gold** | +24% |
| **S&P 500** | ‒38% |
| **Early 2000s recession**  **(2001)** | **Gold** | +5% |
| **S&P 500** | ‒15% |
| **Early 1990s recession**  **(1990)** | **Gold** | +1% |
| **S&P 500** | +3% |
| **Early 1980s recession II**  **(1981)** | **Gold** | +1% |
| **S&P 500** | +4% |
| **Early 1980s recession I**  **(1980)** | **Gold** | +18% |
| **S&P 500** | +9% |
| **End of Bretton Woods and oil crisis**  **(1973)** | **Gold** | +87% |
| **S&P 500** | ‒23% |
| **Recession of 1969–1970**  **(1969-1970)** | **Gold** | +6% |
| **S&P 500** | ‒10% |

Even in periods when both variables experience positive returns, gold returns were always higher, except for 3 successive recessionary periods (1980, 1981, and 1990).

Trends – plots of price

Detrending and volatility.

Gráfico, Gráfico de líneas

Descripción generada automáticamenteGráfico, Histograma

Descripción generada automáticamente

Gráfico, Gráfico de líneas

Descripción generada automáticamenteGráfico, Gráfico de líneas, Histograma, Gráfico de dispersión

Descripción generada automáticamente

**Figure x. Prices versus returns.** Top:Daily prices of gold (USD/oz) and S&P500 (USD) during eight recessionary periods. Bottom: Daily returns.

**5.2. Regression analysis**

**5.2.1. Descriptive statistics**

5.2.1.1. Daily data

“The species sensitivity ranks from the predicted acute toxicity values for nickel that differed slightly for each site. Microalgae were most sensitive to nickel at the four sites, except for site B. Both bivalves (*L. fortunei*) and insects (*C. kiiensis*) had a higher tolerance to nickel than vertebrates in all site waters. “

**Table x.** Descriptive statistics for tests evaluating normality of daily returns data

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **2020** | | **2007** | | **2001** | | **1990** | | **1981** | | **1980** | | **1973** | | **1969** | |
|  | **Gold** | **SP500** | **Gold** | **SP500** | **Gold** | **SP500** | **Gold** | **SP500** | **Gold** | **SP500** | **Gold** | **SP500** | **Gold** | **SP500** | **Gold** | **SP500** |
| **Mean** |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| **Std. Dev.** |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| **Skewness** |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| **Kurtosis** |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| **Normality**  **(Jarque-Bera)** |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| **Serial correlation (Ljung-Box)** |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| **Heteroscesdasticity**  **(LM test)** |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

Table 1 reports descriptive statistics for gold and oil price return series computed on a continuous compounding basis. Average returns were very small relative to the standard deviations and there was no significant trend in the data. The negative values of the skewness statistic suggest a greater probability of large decreases in returns. The return distributions have fat tails, in accordance with the high values for the kurtosis statistic. In fact, the Jarque–Bera test strongly rejected the normality of the unconditional distribution for all the series. Likewise, the existence of serial correlation in the volatility of the returns series was indicated by the Ljung-Box statistic, and the autoregressive conditional heteroskedasticity-Lagrange multiplier (ARCH-LM) statistic suggested that ARCH effects were likely to be found in all the return series. The linear correlation coefficient indicated that oil and gold price returns were dependent; moreover, this dependence increased after the onset of the global financial crisis, even though there was no evidence of any structural change in the correlation coefficient throughout the sampling period.

5.2.1.2. Weekly data

**Table x.** Descriptive statistics for tests evaluating normality of weekly returns data

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **2020** | | **2007** | | **2001** | | **1990** | | **1981** | | **1980** | | **1973** | | **1969** | |
|  | **Gold** | **SP500** | **Gold** | **SP500** | **Gold** | **SP500** | **Gold** | **SP500** | **Gold** | **SP500** | **Gold** | **SP500** | **Gold** | **SP500** | **Gold** | **SP500** |
| **Mean** |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| **Std. Dev.** |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| **Skewness** |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| **Kurtosis** |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| **Normality**  **(Jarque-Bera)** |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| **Serial correlation (Ljung-Box)** |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| **Heteroscesdasticity**  **(LM test)** |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

5.2.1.3. Monthly data

**Table x.** Descriptive statistics for tests evaluating normality of monthly returns data

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **2020** | | **2007** | | **2001** | | **1990** | | **1981** | | **1980** | | **1973** | | **1969** | |
|  | **Gold** | **SP500** | **Gold** | **SP500** | **Gold** | **SP500** | **Gold** | **SP500** | **Gold** | **SP500** | **Gold** | **SP500** | **Gold** | **SP500** | **Gold** | **SP500** |
| **Mean** |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| **Std. Dev.** |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| **Skewness** |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| **Kurtosis** |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| **Normality**  **(Jarque-Bera)** |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| **Serial correlation (Ljung-Box)** |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| **Heteroscesdasticity**  **(LM test)** |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

**5.3. Regression results**

**Table x.** Regression results of sample< periods using daily, weekly, and monthly returns. Bold is used to denote significant p-values at α = 0.05.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Period | Frequency of data | *β* | | Adjusted R2 |
| Coefficient | *P* |
| 2020 | Daily | 1.3031 | **0.0002394\*** | 0.2837 |
| Weekly |  |  |  |
| Monthly |  |  |  |
| 2007 | Daily | -0.0609121 | 0.1315 | 0.003701 |
| Weekly |  |  |  |
| Monthly |  |  |  |
| 2001 | Daily | -4.467e-02 | 0.4009 | -0.001854 |
| Weekly |  |  |  |
| Monthly |  |  |  |
| 1990 | Daily | -0.2665940 | **0.00156\*** | 0.05631 |
| Weekly |  |  |  |
| Monthly |  |  |  |
| 1981 | Daily | 0.4306549 | **7.601e-05\*** | 0.04582 |
|  | Weekly |  |  |  |
|  | Monthly |  |  |  |
| 1980 | Daily | 0.148801 | 0.6933 | -0.007141 |
| Weekly |  |  |  |
| Monthly |  |  |  |
| 1973 | Daily | 0.042534 | 0.6035 | -0.002314 |
| Weekly |  |  |  |
| Monthly |  |  |  |
| 1969 | Daily | -0.0127037 | 0.5944 | -0.003249 |
| Weekly |  |  |  |
| Monthly |  |  |  |

\* Significant at 1% level (p<0.01).

As shown above in **Table 4**, there appears to be a no correlation between gold and S&P returns during the financial crisis, as expected (Table S3).

However, a breakpoint test (Quandt-Andrews unknown breakpoint) detects a statistically significant breakpoint in late October 2008 (10/20/2008) at the 5% level (p=0.0439) (Table S4), as suspected from the graphs (**Figure 2**) and correlation analysis (**Figure 4**).

After running a set of regression models, one before and one after the confirmed breakpoint, the relationship appears to be the opposite of what was initially expected (**Table 5**):

* The regression before the breakpoint (i.e. late September to late October 2008) indicates a significant negative relationship between daily gold returns on S&P returns in the first month of the financial crisis subsample, with a 1% change in the S&P index causing a   
  -0.30% change in the gold price (p=0.0212), at the 5% level (Table S5).
* The regression after the breakpoint (i.e. late October 2008 to late March 2009) indicates a significant positive relationship, with a 1% change in the S&P index causing a 0.16% change in the gold price (p=0.0576), at the 10% level (Table S6).

The “pre-breakpoint” regression model has <30 observations. However, both variables appear to be normally distributed (Figure S2).

Neither model shows evidence of serial correlation (DW=2.0-2.5) and both appear to have small standard errors of regression (2.1-2.6%) (Table S5 and S6).

This change in the relationship between gold and S&P returns mid-sample explains the lack of correlation found in the initial correlation analysis, and indicates that what initially appeared to be a positive correlation followed by a negative correlation is in fact the inverse.

5.3.1.3. Significance

As shown above in **Table 4**, there appears to be a significant positive relationship between gold and S&P returns during the COVID-19 pandemic.

In a regression of daily gold returns on S&P returns over the final month of 2019 and the first half of 2020, a 1% change in the S&P index caused a 0.23% change in the gold price (significant at p=0.00), with a small standard error of regression (1.23%) and no evidence of serial correlation (DW=2.53) (Table S6).

5.3.1.4. Goodness-of-fit/Residual analysis

“A one-day lag was included to see how much it the predictive power of the model improved. Also, you might want to look up "Random Walk Theory" (which argues that yesterday's stock price will have no predictive power on today's stock movement (hence "Random Walk") and is why they view "technical analysis" as essentially voodoo.)”

In terms of the goodness-of-fit test of the villager participation behavior measurement model, the average variance extracted quantity of the PA1 standardized coefficient participation attitude fell short of the requirement. In other words, although some indicators in participation attitude measurement model did not satisfy the requirement, they were a close fit. Therefore, the overall degree of fitting of the measurement model was consistent with the requirements, as shown in Table 6. The chi-square value (CMIN) of the model was 599.83, the degree of freedom (DF) of the model was 295, and its chi-square degree of freedom ratio (CMIN/DF) was 2.033, indicating that the model had a good degree of fit. The SRMR was 0.051. Although it did not meet the standard of <0.05, the difference in value was very small. Thus, it was considered to meet the standard level. The root mean squared error of approximation (RMSEA) met the requirement of being less than 0.08. The parsimony normed fit index (NFI), trucker-Lewis index (TLI), comparative fit index (CFI), and incremental fit index (IFI) all satisfied the criterion of being greater than 0.9. Therefore, the model was considered to exhibit a good fit.

“However, the correlation coefficient (*r2* = 0.26, *p* < 0.001) between the PNECs using typical ACR and chronic BLM-based PNECs was lower than that of the PNECs using BLM-based ACR.”

“The correlation of the final BLM-based ACR between acute and chronic species-specific sensitivities (s) was higher (*r2* = 0.94, *p* < 0.01) than the correlation of typical FACR between EC50s and NOECs (*r2* = 0.92, *p* < 0.01). These discrepancies in reliability can be expected to increase with the number of toxicity data in various water chemistries, such as natural waters.”

**5. Discussion**

**Daily vs weekly data**

Intuition: Although gold is widely believed to be negatively correlated with stock market indices in times of market turmoil, my analysis suggests that this relationship does not hold in the US stock market.

During the 2008 financial crisis and COVID-19 pandemic, gold was found to behave counterintuitively to its status of a safe haven, either becoming positively correlated with an ailing US stock market or remaining positively correlated before and after a sustained run of negative returns in a major US stock index.

In fact, during [x], [x], and [x], gold exhibited unexpected behavior.

Although my findings seem to initially discredit the widely held belief that gold is inversely correlated or uncorrelated with the stock market, in the following I provide some intuition that mat account for these counterintuitive results.

**Overall trends**

Daily

Weekly

Monthly

Period

**Explanations**

*Economics recession vs financial crises*

Limitation: Slight difference in dates for economic recessions as defined by the NBER and periods of greatest volatility in financial markets

Future perspectives: Comparison between different dates, same dataset but using breakpoint analysis.

*A dash for cash*

In times of financial turmoil, cash is king. As markets lose value, investors rush to obtain cash by selling any available assets, such as gold, which is one of the most liquid assets after cash. This is evident in both the financial crisis in 2008 and the COVID-19 pandemic, where the price of an oz. of gold appears to drop with the stock index, as investors sell off their gold holdings in exchange for increased liquidity.

The regression model for gold returns on S&P returns during both the COVID-19 pandemic and financial crisis initial positive relationship of gold and stock during financial crises.

During the 2008 financial crisis gold was found to be negatively correlated to stock market returns initial, there is clearly a downward trend in the price of gold during the first month of my sample (**Figure 2A**)

Limitations:

Future perspectives:

*Gold as a store of value*

Here, I provide some intuition for the negative relationship of gold and stock during financial crises.

Intuition for negative relationship between gold and stock indices: As markets lose value, asset are moved to safer investment that will hold value, e.g. gold. After an initial sell-out, it is expected for gold to grow and obtain upward momentum as the market shifts into a risk-off attitude.

Limitations:

Future perspectives:

*Duration of market volatility*

Intuition for positive relationship during COVID-19 pandemic: Unlike the change in correlation observed during the 2008 financial crisis, gold appears to have remained positively correlated with markets throughout 6-month sample period, i.e. the current duration of the ongoing pandemic. This may be partly due to the nature of the market response to the COVID-19 pandemic – a sharp fall in the stock index was followed by an equally a sharp rise.

Limitations:

Future perspectives:

*Different types of flight to gold*

Investing directly in the yellow metal is not the only strategy. Depending on the severity of the financial crisis, investors can also chose to flee from stocks into gold mining shares. This paper uses a large sample of gold mining companies traded on four stock markets to identify different types of flight to gold. After extreme financial shocks, we observe that investors flee from stocks including gold mining shares to gold bullion. However, if the financial shock is less extreme, investors only flee from stocks excluding gold mining shares. This highlights that the study of gold mining companies provides important additional information in the context of flight to quality and safe haven effects.

Limitations:

Future perspectives:

**Reliability of model**

Weak overall, but does appears to give some insights into how the relationship between investment into stocks and into gold changes over time and during different periods of financial instability.

Although gold is widely believed to be negatively correlated with stock market indices in times of market turmoil, my analysis suggests that this relationship does not hold in the US stock market during the 2008 financial crisis or the current COVID-19 pandemic. In fact, gold was found to behave counterintuitively to its status of a safe haven, either becoming positively correlated with an ailing US stock market or remaining positively correlated before and after a sustained run of negative returns in a major US stock index.

Due to the ongoing nature of the COVID-19 pandemic, my analysis of the COVID-19 pandemic has been restricted to the beginning of this crisis. It will be worth monitoring the changes in the prices of gold and S&P 500 index in the coming months to see how the relationship between these two variables changes as this crisis develops further.

This analysis focused solely on the short-run relationship between the price of gold and the S&P index in times of ailing stock markets using samples comprising 6 months of data. More insights could be gained by using larger sample sizes and more variables, including analysing periods of stable markets and using volatility as a measure of correlation.

*Limitations*

* explanatory power of the model

*Future perspectives*

Wider implementation

“BLM applied to vertebrates are identical (De Schamphelaere et al., 2004; De Schamphelaere et al., 2005). Therefore, the final BLM-based ACR for zinc can be proposed, and if a reliable acute SSD for indigenous species is built, the BLM-based risk assessment for zinc can be applied to freshwater in Korea. Consequently, the BLM-based ACR determined from the ecotoxicity data of various standard toxicity test species will contribute to the application of BLM-based risk assessment for Ni, Cu, and Zn to a wider ecoregion.”

**6. Conclusion**

Due to the ongoing nature of the COVID-19 pandemic, my analysis of the COVID-19 pandemic has been restricted to the beginning of this crisis. It will be worth monitoring the changes in the prices of gold and S&P 500 index in the coming months to see how the relationship between these two variables changes as this crisis develops further.

Furthermore, with regards to the 2008 financial crisis, there is much more data available for this episode than the current COVID-19 pandemic. The results of my analysis reflect only one period during the financial crisis, which started in 2007 and had an effect on financial markets from 2007 until 2009. My analysis draws on data obtained near the end of this crisis, and it may be worth looking at other samples from this crises, whether a 6-month sample in the earlier stages of the financial crises or the full 3-year sample.

Lastly, this analysis focused solely on the short-run relationship between the price of gold and the S&P index in times of ailing stock markets using samples comprising 6 months of data. More insights could be gained by using larger sample sizes and more variables, including analysing periods of stable markets and using volatility as a measure of correlation.

Try to determine why investors feel like gold is a store for value. Could bitcoin be the next gold? What other commodities or financial instruments could be targeted during flight-to-quality?

While the model was unable to explain

The times where the model was a better fit indicated times where the S&P 500 did exhibit some power on the movement of gold

Investor behaviour highlights the importance of the role of gold as a store for value in the financial market – rather than a mythical investment tool, gold provides a store for value while other investments are returning losses, mitigating potentially significant monetary losses. In parallel, withdrawing investments and holding cash exposes investors to inflation, which tends to be particularly high immediately after a financial crisis. In this context, investors should seek to invest in assets that are uncorrelated with stocks during market crisis (i.e safe haven).

However, this property appears to be short-lived

**Bibliography**

1. Baur and Lucey (2010). Is Gold a Hedge or a Safe Haven? An Analysis of Stocks, Bonds and Gold. *The Financial Review* 45 (2010) 217–229.
2. .
3. .
4. .
5. .
6. .
7. .
8. .
9. .
10. .
11. .
12. .
13. .
14. .
15. .
16. .
17. .
18. ..
19. .
20. .
21. .
22. .
23. ..
24. .
25. ..
26. .
27. .
28. .
29. .
30. .

**Appendix**

**1. Correlation analysis**

**Table S1.** Covariance analysis for gold and S&P 500 daily returns in financial crisis subsample: September 2008 to March 2009.

|  |  |  |  |
| --- | --- | --- | --- |
| Covariance Analysis: Ordinary | | |  |
| Date: 05/24/20 Time: 22:47 | | |  |
| Sample: 9/22/2008 3/20/2009 | | |  |
| Included observations: 117 | | |  |
| Balanced sample (listwise missing value deletion) | | | |
|  |  |  |  |
|  |  |  |  |
| Correlation | |  |  |
| t-Statistic | |  |  |
| Probability | |  |  |
| Observations | SP500\_DLOG | GOLD\_DLOG |  |
| SP500\_DLOG | 1.000000 |  |  |
|  | ----- |  |  |
|  | ----- |  |  |
|  | 117 |  |  |
|  |  |  |  |
| GOLD\_DLOG | -0.068167 | 1.000000 |  |
|  | -0.732719 | ----- |  |
|  | 0.4652 | ----- |  |
|  | 117 | 117 |  |
|  |  |  |  |
|  |  |  |  |

**Table S2. C**ovariance analysis of gold and S&P 500 daily returns in COVID-19 subsample: December 2019 to May 2020.

|  |  |  |  |
| --- | --- | --- | --- |
| Covariance Analysis: Ordinary | | |  |
| Date: 05/21/20 Time: 12:39 | | |  |
| Sample: 12/23/2019 5/19/2020 | | |  |
| Included observations: 89 | | |  |
| Balanced sample (listwise missing value deletion) | | | |
|  |  |  |  |
|  |  |  |  |
| Correlation | |  |  |
| t-Statistic | |  |  |
| Probability | |  |  |
| Observations | GOLD\_DLOG | SP500\_DLOG |  |
| GOLD\_DLOG | 0.000207 |  |  |
|  | 1.000000 |  |  |
|  | ----- |  |  |
|  | ----- |  |  |
|  | 89 |  |  |
|  |  |  |  |
| SP500\_DLOG | 0.000244 | 0.001084 |  |
|  | 0.515097 | 1.000000 |  |
|  | 5.605325 | ----- |  |
|  | **0.0000\*** | ----- |  |
|  | 89 | 89 |  |
|  |  |  |  |
|  |  |  |  |

\* Significant at 5% level (p<0.05).

**2. Regression output**

**2.1. Financial crisis: September 2008 to March 2009**

**Table S3.** Regression output of financial crisis subsample: September 2008 to March 2009.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Dependent Variable: GOLD\_DLOG | | | |  |
| Method: Least Squares | | |  |  |
| Date: 05/24/20 Time: 22:31 | | | |  |
| Sample: 9/22/2008 3/20/2009 | | | |  |
| Included observations: 117 | | |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|  |  |  |  |  |
|  |  |  |  |  |
| C | -0.000299 | 0.002257 | -0.132359 | 0.8949 |
| SP500\_DLOG | -0.045307 | 0.061834 | -0.732719 | 0.4652 |
|  |  |  |  |  |
|  |  |  |  |  |
| R-squared | 0.004647 | Mean dependent var | | -0.000124 |
| Adjusted R-squared | -0.004008 | S.D. dependent var | | 0.024233 |
| S.E. of regression | 0.024282 | Akaike info criterion | | -4.581241 |
| Sum squared resid | 0.067804 | Schwarz criterion | | -4.534025 |
| Log likelihood | 270.0026 | Hannan-Quinn criter. | | -4.562072 |
| F-statistic | 0.536877 | Durbin-Watson stat | | 2.117752 |
| Prob(F-statistic) | 0.465221 |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |



**Figure S1.** Plottedfitted **r**esiduals of regression analysis for financial crisis subsample: September 2008 to March 2009.

**Table S4.** Quandt-Andrews unknown breakpoint test: significant breakpoint at 10/20/2008.

|  |  |  |  |
| --- | --- | --- | --- |
| Quandt-Andrews unknown breakpoint test | | | |
| Null Hypothesis: No breakpoints within 15% trimmed data | | | |
| Varying regressors: All equation variables | | | |
| Equation Sample: 9/22/2008 3/20/2009 | | | |
| Test Sample: 10/16/2008 2/25/2009 | | | |
| Number of breaks compared: 82 | | | |
|  |  |  |  |
|  |  |  |  |
| Statistic | Value |  | Prob. |
|  |  |  |  |
|  |  |  |  |
| Maximum LR F-statistic (10/20/2008) | 5.935979 |  | **0.0439\*** |
| Maximum Wald F-statistic (10/20/2008) | 11.87196 |  | 0.0439 |
|  |  |  |  |
| Exp LR F-statistic | 1.287964 |  | 0.0969 |
| Exp Wald F-statistic | 3.406127 |  | 0.0404 |
|  |  |  |  |
| Ave LR F-statistic | 1.522421 |  | 0.1708 |
| Ave Wald F-statistic | 3.044842 |  | 0.1708 |
|  |  |  |  |
|  |  |  |  |
| Note: probabilities calculated using Hansen's (1997) method | | | |
| WARNING: estimation sample is non-continuous (probabilities | | | |
| calculated assuming a continuous sample) | | | |

\* Significant at 5% level (p<0.05).





**Figure S2.** Histograms confirming the normality of gold and S&P 500 returns in reduced financial crisis subsample (<30 observations): September to October 2008.

**Table S5.** Regression output of pre-breakpoint financial crisis subsample: September to October 2008.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Dependent Variable: GOLD\_DLOG | | | |  |
| Method: Least Squares | | |  |  |
| Date: 05/24/20 Time: 22:36 | | | |  |
| Sample: 9/22/2008 10/20/2008 | | | |  |
| Included observations: 21 | | |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|  |  |  |  |  |
|  |  |  |  |  |
| C | -0.007735 | 0.006005 | -1.288120 | 0.2132 |
| SP500\_DLOG | -0.303586 | 0.120831 | -2.512485 | **0.0212\*** |
|  |  |  |  |  |
|  |  |  |  |  |
| R-squared | 0.249385 | Mean dependent var | | -0.004238 |
| Adjusted R-squared | 0.209879 | S.D. dependent var | | 0.030116 |
| S.E. of regression | 0.026769 | Akaike info criterion | | -4.312727 |
| Sum squared resid | 0.013615 | Schwarz criterion | | -4.213249 |
| Log likelihood | 47.28363 | Hannan-Quinn criter. | | -4.291138 |
| F-statistic | 6.312579 | Durbin-Watson stat | | 2.537897 |
| Prob(F-statistic) | 0.021175 |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |

\* Significant at 5% level (p<0.05).

**Table S6.** Regression output of post-breakpoint financial crisis subsample: October 2008 to March 2009.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Dependent Variable: GOLD\_DLOG | | | |  |
| Method: Least Squares | | |  |  |
| Date: 05/24/20 Time: 23:44 | | | |  |
| Sample: 11/20/2008 3/20/2009 | | | |  |
| Included observations: 74 | | |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|  |  |  |  |  |
|  |  |  |  |  |
| C | 0.001601 | 0.002526 | 0.633754 | 0.5282 |
| SP500\_DLOG | 0.167556 | 0.086824 | 1.929831 | **0.0576\*** |
|  |  |  |  |  |
|  |  |  |  |  |
| R-squared | 0.049182 | Mean dependent var | | 0.001580 |
| Adjusted R-squared | 0.035976 | S.D. dependent var | | 0.022133 |
| S.E. of regression | 0.021731 | Akaike info criterion | | -4.793479 |
| Sum squared resid | 0.034002 | Schwarz criterion | | -4.731207 |
| Log likelihood | 179.3587 | Hannan-Quinn criter. | | -4.768638 |
| F-statistic | 3.724248 | Durbin-Watson stat | | 2.041141 |
| Prob(F-statistic) | 0.057568 |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |

\* Significant at 10% level (p<0.10).

**2.2. COVID-19 pandemic: December 2019 to May 2020**

**Table S7.** Regression output of COVID-19 pandemic subsample.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Dependent Variable: GOLD\_DLOG | | | |  |
| Method: Least Squares | | |  |  |
| Date: 05/24/20 Time: 22:52 | | | |  |
| Sample (adjusted): 12/23/2019 5/19/2020 | | | |  |
| Included observations: 89 after adjustments | | | | |
|  |  |  |  |  |
|  |  |  |  |  |
| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|  |  |  |  |  |
|  |  |  |  |  |
| C | 0.001393 | 0.001325 | 1.051184 | 0.2961 |
| SP500\_DLOG | 0.225371 | 0.040207 | 5.605325 | **0.0000\*** |
|  |  |  |  |  |
|  |  |  |  |  |
| R-squared | 0.265325 | Mean dependent var | | 0.001022 |
| Adjusted R-squared | 0.256880 | S.D. dependent var | | 0.014485 |
| S.E. of regression | 0.012486 | Akaike info criterion | | -5.906133 |
| Sum squared resid | 0.013564 | Schwarz criterion | | -5.850209 |
| Log likelihood | 264.8229 | Hannan-Quinn criter. | | -5.883592 |
| F-statistic | 31.41967 | Durbin-Watson stat | | 1.860022 |
| Prob(F-statistic) | 0.000000 |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |

\* Significant at 5% level (p<0.05).



**Figure S3.** Fittedresiduals of COVID-19 subsample (2019-2020) regression analysis.

**Table output (R)**

**Summary statistics**

summary(X1969$`SP500 returns`)

Min. 1st Qu. Median Mean 3rd Qu. Max. NA's

-0.027682 -0.006289 -0.000804 -0.000464 0.004372 0.050224 13

> summary(X1969$`Gold returns`)

Min. 1st Qu. Median Mean 3rd Qu. Max. NA's

-0.024231 -0.001133 0.000000 0.000250 0.001198 0.011842 14

> summary(X1973$`SP500 returns`)

Min. 1st Qu. Median Mean 3rd Qu. Max. NA's

-0.036713 -0.010556 -0.001705 -0.000784 0.007623 0.045959 25

> summary(X1973$`Gold returns`)

Min. 1st Qu. Median Mean 3rd Qu. Max. NA's

-0.071207 -0.009409 0.000000 0.002173 0.012855 0.103261 20

> summary(X1980$`SP500 returns`)

Min. 1st Qu. Median Mean 3rd Qu. Max. NA's

-0.030067 -0.005152 0.001510 0.000887 0.007083 0.036373 8

> summary(X1980$`Gold returns`)

Min. 1st Qu. Median Mean 3rd Qu. Max. NA's

-0.132353 -0.020884 -0.004267 0.001501 0.025317 0.133155 9

> summary(X1981$`SP500 returns`)

Min. 1st Qu. Median Mean 3rd Qu. Max. NA's

-0.039689 -0.006282 -0.000652 0.000273 0.005745 0.047555 21

> summary(X1981$`Gold returns`)

Min. 1st Qu. Median Mean 3rd Qu. Max. NA's

-0.064373 -0.012049 -0.000613 0.000698 0.011083 0.110433 23

> summary(X1990$`SP500 returns`)

Min. 1st Qu. Median Mean 3rd Qu. Max. NA's

-0.030244 -0.007756 0.001310 0.000410 0.007858 0.037322 13

> summary(X1990$`Gold returns`)

Min. 1st Qu. Median Mean 3rd Qu. Max. NA's

-0.059553 -0.005371 0.001230 0.000597 0.008269 0.038397 10

> summary(X2001$`SP500 returns`)

Min. 1st Qu. Median Mean 3rd Qu. Max. NA's

-0.043181 -0.008970 -0.000287 -0.000269 0.008121 0.043680 14

> summary(X2001$`Gold returns`)

Min. 1st Qu. Median Mean 3rd Qu. Max. NA's

-0.034260 -0.004242 0.000000 0.000412 0.004086 0.066850 10

> summary(X2007$`SP500 returns`)

Min. 1st Qu. Median Mean 3rd Qu. Max. NA's

-0.090350 -0.012979 0.000081 -0.000828 0.010584 0.115800 31

> summary(X2007$`Gold returns`)

Min. 1st Qu. Median Mean 3rd Qu. Max. NA's

-0.076624 -0.009145 0.000586 0.000468 0.010078 0.070809 31

> summary(X2020$`SP500 returns`)

Min. 1st Qu. Median Mean 3rd Qu. Max. NA's

-5.02192 -0.93316 -0.03794 -0.01010 1.04594 5.26747 3

> summary(X2020$`Gold returns`)

Min. 1st Qu. Median Mean 3rd Qu. Max. NA's

-11.9841 -3.2705 -0.3349 -0.5385 1.0270 9.3828 2

**Normality (Shapiro-Wilk’s test)**

**Null:** “the sample distribution is normal”

**H1:** “the sample distribution is not normal”

If the test is significant, the distribution is non-normal.

data: X1969$`Gold returns`  
W = 0.81745, p-value = 1.204e-15

data: X1969$`SP500 returns`  
W = 0.96248, p-value = 1.006e-05

data: X1973$`Gold returns`  
W = 0.95767, p-value = 3.753e-08

data: X1973$`SP500 returns`  
W = 0.98982, p-value = 0.02364

data: X1980$`Gold returns`  
W = 0.9775, p-value = 0.03887

data: X1980$`SP500 returns`  
W = 0.98717, p-value = 0.3018

data: X1981$`Gold returns`  
W = 0.96561, p-value = 5.446e-07

data: X1981$`SP500 returns`  
W = 0.97469, p-value = 1.551e-05

data: X1990$`Gold returns`  
W = 0.94337, p-value = 3.471e-06

data: X1990$`SP500 returns`  
W = 0.99364, p-value = 0.6997

data: X2001$`Gold returns`  
W = 0.81305, p-value = 2.391e-13

data: X2001$`SP500 returns`  
W = 0.99183, p-value = 0.4809

data: X2007$`Gold returns`  
W = 0.97588, p-value = 9.948e-06

data: X2007$`SP500 returns`  
W = 0.94899, p-value = 7.714e-10

data: X2020$`Gold returns`  
W = 0.96834, p-value = 0.2902

data: X2020$`SP500 returns`  
W = 0.95137, p-value = 0.07818

**Autocorrelation (Durbin-Watson test)**

**Null:** “no autocorrelation”

**H1:** “data are autocorrelated”

data: lm.returns.1969

DW = 1.6442, p-value = 0.003596

data: lm.returns.1973

DW = 1.969, p-value = 0.3859

data: lm.returns.1980

DW = 2.1873, p-value = 0.8493

data: lm.returns.1981

DW = 2.5128, p-value = 1

data: lm.returns.1990

DW = 2.0749, p-value = 0.678

data: lm.returns.2001

DW = 2.075, p-value = 0.681

data: lm.returns.2007

DW = 1.966, p-value = 0.3787

data: lm.returns.2020

DW = 2.5219, p-value = 0.9559

No evidence of serial correlation, except in 1969 data.

**Results of linear model: gold~sp500**

gold returns = alpha + beta(s&p 500 returns)

**1969**

Residuals:

Min 1Q Median 3Q Max

-0.0245343 -0.0013111 -0.0002234 0.0010950 0.0116145

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 0.0002871 0.0002325 1.235 0.218

X1969$`SP500 returns` -0.0127037 0.0238235 -0.533 0.594

Residual standard error: 0.003463 on 220 degrees of freedom

(20 observations deleted due to missingness)

Multiple R-squared: 0.001291, Adjusted R-squared: -0.003249

F-statistic: 0.2843 on 1 and 220 DF, p-value: 0.5944

**1973**

Residuals:

Min 1Q Median 3Q Max

-0.072789 -0.011316 -0.002322 0.010459 0.100034

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 0.002360 0.001132 2.085 0.0379 \*

X1973$`SP500 returns` 0.042534 0.081808 0.520 0.6035

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.02011 on 315 degrees of freedom

(32 observations deleted due to missingness)

Multiple R-squared: 0.0008574, Adjusted R-squared: -0.002314

F-statistic: 0.2703 on 1 and 315 DF, p-value: 0.6035

**1980**

Residuals:

Min 1Q Median 3Q Max

-0.133166 -0.021354 -0.005268 0.024488 0.132318

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 0.001596 0.003926 0.406 0.685

X1980$`SP500 returns` 0.148801 0.376442 0.395 0.693

Residual standard error: 0.04289 on 118 degrees of freedom

(11 observations deleted due to missingness)

Multiple R-squared: 0.001322, Adjusted R-squared: -0.007141

F-statistic: 0.1562 on 1 and 118 DF, p-value: 0.6933

**1981**

Residuals:

Min 1Q Median 3Q Max

-0.075446 -0.012301 -0.002293 0.010852 0.101344

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 0.0005329 0.0011211 0.475 0.635

X1981$`SP500 returns` 0.4306549 0.1073978 4.010 7.6e-05 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.01989 on 313 degrees of freedom

(35 observations deleted due to missingness)

Multiple R-squared: 0.04886, Adjusted R-squared: 0.04582

F-statistic: 16.08 on 1 and 313 DF, p-value: 7.601e-05

**1990**

Residuals:

Min 1Q Median 3Q Max

-0.050387 -0.006284 0.000088 0.006925 0.040252

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 0.0007838 0.0009571 0.819 0.41410

X1990$`SP500 returns` -0.2665940 0.0827910 -3.220 0.00156 \*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.01203 on 156 degrees of freedom

(18 observations deleted due to missingness)

Multiple R-squared: 0.06232, Adjusted R-squared: 0.05631

F-statistic: 10.37 on 1 and 156 DF, p-value: 0.00156

**2001**

Residuals:

Min 1Q Median 3Q Max

-0.034779 -0.004191 -0.000068 0.003941 0.067580

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) -8.796e-06 7.509e-04 -0.012 0.991

X2001$`SP500 returns` -4.467e-02 5.304e-02 -0.842 0.401

Residual standard error: 0.009438 on 156 degrees of freedom

(19 observations deleted due to missingness)

Multiple R-squared: 0.004527, Adjusted R-squared: -0.001854

F-statistic: 0.7094 on 1 and 156 DF, p-value: 0.4009

**2007**

Residuals:

Min 1Q Median 3Q Max

-0.069990 -0.008694 0.000211 0.009879 0.069599

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 0.0004195 0.0010094 0.416 0.678

X2007$`SP500 returns` -0.0609121 0.0402921 -1.512 0.132

Residual standard error: 0.01879 on 345 degrees of freedom

(45 observations deleted due to missingness)

Multiple R-squared: 0.006581, Adjusted R-squared: 0.003701

F-statistic: 2.285 on 1 and 345 DF, p-value: 0.1315

**2020**

Residuals:

Min 1Q Median 3Q Max

-6.1760 -2.5963 -0.2191 2.3920 10.5005

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) -0.5579 0.6004 -0.929 0.358569

X2020$`SP500 returns` 1.3031 0.3213 4.056 0.000239 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 3.797 on 38 degrees of freedom

(4 observations deleted due to missingness)

Multiple R-squared: 0.3021, Adjusted R-squared: 0.2837

F-statistic: 16.45 on 1 and 38 DF, p-value: 0.0002394

gold = a + (b \* -0.01010)

gold = -0.5579 + (1.3031 \* -0.01010)

gold = -0.57106131

Gold