

Gold Forecasted vs Actual Prices During COVID Market Conditions

Counterfactual Model Demo in R can be found at GitHub: <https://github.com/ErickWGSanchez>

Team Members – Role Responsibilities

Grean Ramos (Financial Analyst): Conduct research on datasets needed to be used as control variables for gold prices, documentation, and project organizer.

Richard Collie (Economic Researcher): Research needed datasets, validate/research economic information, and document research and information data.

Erick Guevara (Data Analyst): Transform and clean datasets, conduct pre linear model diagnostics, develop vector autoregressive forecasting model, conduct out-of-sample simulations and timeseries analysis, initiate; forecast variation error decomposition, impulse response analysis, and sensitivity analysis.

Problem Description

The unfortunate event of the COVID-19 Pandemic affected the entire world. No country was left untouched, and it was a time that will never be forgotten. Alongside the great number of deaths and suffering that the world experienced, the global economies also found themselves in market conditions that had never been seen before. Due to quarantine and stay-at-home orders, the manufacturing of goods slowed down tremendously, travel and shipping decreased significantly and overall confidence in the market was historically low.

Specifically in March 2020, the entire market crashed. The DOW (DJIA) had fallen roughly 26% over the course of four trading days [Mieszko]. In parallel with this index, virtually every other stock had taken a hard hit. This was due to many factors, the primary source being the high volatility and low market sentiment because of the actions taken by the government in order to mitigate the COVID-19 virus. The stock market is considered “forward-looking”, which means that current values reflect future expectations. The pandemic and quarantine created an environment with unclear expectations, thus

making markets and investors panic. The market does not like uncertainty, thus the historic crash occurred.

The goal of this project is to conduct research and develop a tip-of-the-iceberg counterfactual model that would be used to see the non-Covid results of gold prices in conjunction with forecasting to basically discover what the market conditions would have been if the pandemic did not occur. The asset of choice that will be focused is on gold prices. The team shall use methods that were taught in class and other known applications that will help to develop an accurate model which will estimate gold prices.

Data Description

The gold data shall be extracted directly from investing.com (<https://www.investing.com>) as a csv file. The prices at the end of the day will be used. A 19-year analysis of the daily historical prices will be done, starting from February 01, 2004, to November 25, 2023. From this data, prices will be analyzed and used for future forecasting with a specific focus of the timeframe of COVID quarantine. The forecasted data will be compared to the actual historical data.

The interest and inflation rate data will be extracted from R package API called quantmod since it has UpToDate and accurate historical data reaching back to a decent amount to forecast. The interest and inflation rate datasets will be monthly data that contains 184 rows of observations with 2 columns involving rate and date.

Methodology

The prices of gold are the response variable. In turn, research needs to be conducted in order to understand the predictor variables. Reasons for these specific variables/datasets is due to the fact these are major factors that influence the gold returns. The following are other datasets which will be researched further (note more may be added to the list depending on time until presentation):

- a. Interest Rate
- b. Inflation Rate
- c. Unemployment Rate

An outlined process of what results are expected is to develop a counterfactual model that will estimate what gold prices would have been based on estimation and compare it with the actual post covid

timeline to validate results. Descriptive externalities of the process of individual methods are not included.

Historical Simulations

Method involves taking historical price data for gold. Excluding data for that Covid period. To run simulations based on other factors like interest rate, inflation, and unemployment to create scenarios of the market during that period without covid.

1. **Out-of-sample Simulations** - This simulation method will be used to generate large numbers of scenarios from the vector autoregressive model that will be used to model the relationships of the control variable to gold prices and generate an estimate the trajectory from pre-covid period.
2. **Scenario Analysis** - To then analyze the simulations results with counterfactual model and current returns and develop a metric to see the difference between the trends.

Counterfactual Modeling

Construct a **vector autoregressive** model that estimates gold prices by taking in the influence from variation of other potentially impactful variables such as interest, inflation, and dollar exchange rates.

Time series analysis - will be conducted to illustrate the data over a certain timeframe. Firstly, we will test stationary, seasonality, and trends of the data to make sure the data is correctly normalized before building a forecasting model. Auto correlation function and simple data visualization will be used to test stationery, seasonality, and trends. This will be backed by performing an augmented dickey fuller test to confirm for stationary. Depending on whether data is stationary and/or has a linear combination of combine variable stationary behavior then type of model we will either use is vector autoregressive or vector error correction models to predict the data during COVID which will be the 2 years towards the end of the quarantine period of covid. Means square error and confidence intervals will be used for verifying the validation of the data by comparing the plots of the actual output of gold data value with what would've been plot.

Vector autoregressive model – This model is best for the approach of having control variables that checks gold prices before pre covid by measuring the association and developing a forecasting estimate which will tell us how COVID affected gold prices. From this model will be able to use Out sample simulations to create estimates that will be used to generate a mean trajectory to then compare analysis.

Vector error correction model – This model is used best for non-stationary and cointegrated data that involve two or more series of datasets which combine can become stationary. This works based on integrating short term pattern with long term equilibrium relationships meaning when combined and looking at the overall structure it will achieve a constant mean and variance.

Comparative Analysis/Back testing/Performance Metrics

To verify results of the estimates a series of comparing scenarios and actual historical data post covid will be used as validation.

1. **Visual Analysis** - A simple visual interpretation of results with actual historical values.
2. **Mean absolute percentage error (MAPE)** - to measure the percentage difference between observed and predicted values.
3. **Sensitivity Analysis** – Measures the robustness of the model, tests different parameters to see which has the best and most accurate results.
4. **Confidence Intervals** - Provides a range to see actual expected returns to fall a certain level.
5. **Forecast Error Decomposition** - Decompose the forecast errors into bias, variance, and other components to identify inaccuracies in the model.
6. **Residual Analysis** - Examines differences between observed and predicted to check for underestimations or overestimations returns.
7. **Impulse Response Analysis** – Measures the shock that one variable can influence another variable in this case the response of gold price from interest rate and inflation rate structural break in the period of COVID.

Expected Results

As stated above the objective of developing a counterfactual model is to see the non-COVID behavior in the gold price market index and create a baseline of what would've been if covid and mitigating policies due to COVID never occurred. This can potentially reveal interesting trends, projections, and see an estimate of the gold price not affected by COVID. A plot showcasing average estimate connecting with actual observed gold price and a comparison error analysis between the plots will be the final illustration to be able to voice the methods and validation used to output such results. Since, this is a "Making up history model", will use different methods of validation to give our model integrity such as testing the robustness, accuracy, precision, and error metrics to ultimately establish a good rapport of correctly applied model.

Works Cited

Mieszko Mazur, Man Dang, Miguel Vega,
COVID-19 and the march 2020 stock market crash. Evidence from S&P1500,
Finance Research Letters,
Volume 38, 2021, 101690, ISSN 1544-6123, <https://doi.org/10.1016/j.frl.2020.101690>.
(<https://www.sciencedirect.com/science/article/pii/S1544612320306668>)