Barış Mert Türegün

31165

09.01.20216

Analyzing Bitcoin and Gold Market Dynamics

Motivation

In financial discussions, Bitcoin is frequently compared to gold, usually framed through the popular digital gold narrative. However, it is widely accepted that Bitcoin carries much higher volatility and risk than gold does. Because of these differences, the goal of this project is not to argue that Bitcoin and gold should act exactly the same way. Instead, the study looks to see if Bitcoin shows any significant similarities to gold over longer periods and during different types of market shifts.

The research focuses on how these two assets behave regarding their price movements and risk profiles. This includes looking at long-term trends, patterns in volatility, and whether the relationship between them stays stable over time. Beyond just talking about whether they work as a store of value, the comparison also investigates if Bitcoin and gold move together in a reliable way, or if those similarities are just short-term events that disappear. The overall purpose is to provide a clear evaluation based on actual data. To do this, I use exploratory time-series analysis and a machine learning phase specifically focused on how their volatility behaves.

Data Source and Data Collection

The data for this project was gathered from public financial markets via Yahoo Finance, using the yfinance library in Python. I chose this method because it is transparent and allows the work to be reproduced easily, as the same code can be used to pull the same time series again.

I collected two specific daily time series:

- Bitcoin (BTC): Ticker BTC-USD
- Gold (proxy): Ticker GC=F (COMEX Gold Futures)

For both assets, I used the daily closing price as the main variable. This is the most used choice in financial analysis for calculating daily returns and assessing risk. While the raw data includes the full OHLCV (Open, High, Low, Close, and Volume), I focused specifically on the Close prices to maintain consistency and to filter out noise from intraday trading patterns.

The data retrieval was handled programmatically in Python by setting a specific timeframe, covering the period from 2016-01-04 to 2025-12-12. A common issue with comparing these assets is that Bitcoin trades every day, including weekends, while gold futures follow exchange schedules and close for holidays and weekends. To make sure the comparison was fair, I aligned both datasets by date and kept only the days where both assets had active trading data. Any rows that had missing values after this alignment were removed.

Once the cleaning and alignment were finished, the final dataset for Phase 2 consisted of 2501 observations with two primary variables: BTC_Close and Gold_Close. This combined dataset was used for the exploratory analysis, such as looking at returns and rolling correlations, as well as the machine learning work in Phase 3.

Data Analysis: Techniques and Stages

The analysis was organized into a structured time-series workflow, which was carried out in two main parts: exploratory analysis (Phase 2) and machine learning modeling (Phase 3).

Data Preparation

Once the daily closing prices for Bitcoin and gold were aligned by date, I calculated several specific features to help with the analysis:

- Daily simple returns and daily log returns, with log returns serving as the primary measure for the study.
- Rolling 30-day volatility, calculated as the standard deviation of log returns, to show short-term risk levels.
- Rolling 30-day correlation between the log returns of both assets to see how stable their relationship is over time.
- Lagged returns ($t-1$) for both Bitcoin and gold to account for short-term momentum effects within the machine learning models.

Phase 2 – Exploratory Data Analysis (EDA)

The second phase involved a descriptive time series analysis to compare how stable the assets are and how they move together. This part of the project included:

- Price level plots and log-scale price plots, which allowed for a comparison of long-term behavior despite the huge difference in price scales between the two assets.
- Return calculations used to check the actual differences in the size of daily price swings.
- 30 day rolling volatility plots to help point out specific risk differences and major volatility spikes.
- 30 day rolling correlation plots to investigate if Bitcoin and gold actually move together in a consistent way or if it only happens temporarily.

Phase 3 – Machine Learning

In the third phase, I applied machine learning methods to the dataset using a time-based split, training the models on earlier years and testing them on more recent data.

- Task A (daily return prediction): I used a baseline (predicting 0), Linear Regression, and Random Forest models to try and predict Bitcoin's daily log returns. I evaluated their performance using MAE, RMSE, and R^2.
- Task B (volatility prediction): This was the primary focus, where I predicted the next day rolling volatility for Bitcoin using volatility and return features from both assets. The models included a persistence baseline (where next volatility is roughly the same as current volatility), Linear Regression, and Random Forest. I used the same performance metrics as Task A and looked at Random Forest feature importance to help interpret the results.

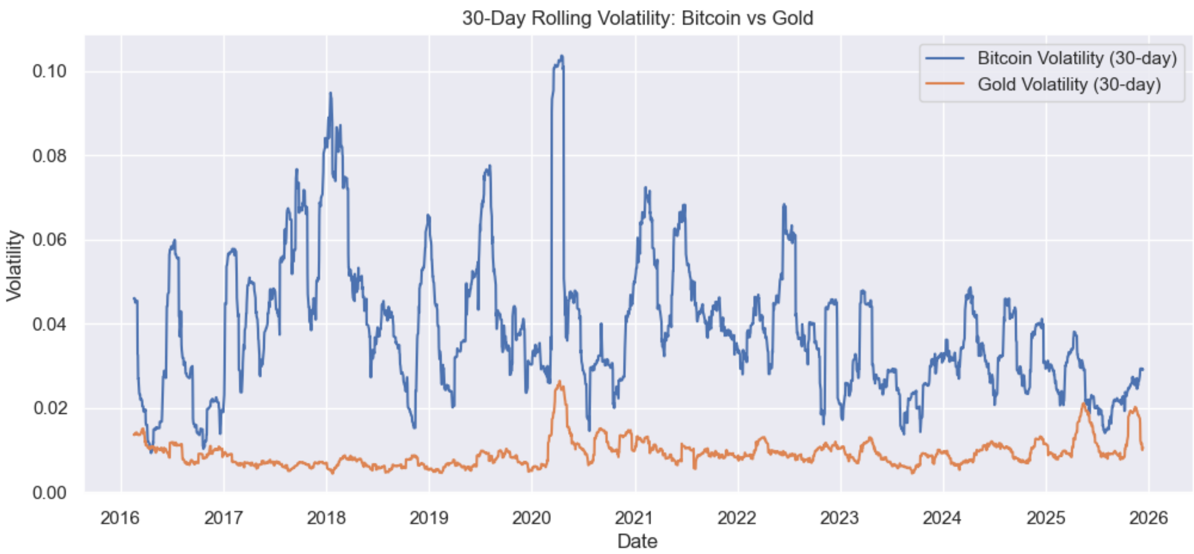
Findings

This section covers the primary results of the project and discusses why they matter when evaluating whether Bitcoin actually acts like gold as a store of value. These findings bring together the exploratory evidence from Phase 2 and the model results from Phase 3.

Bitcoin has a much higher and less stable risk profile than gold (Phase 2)

The most significant result comes from the risk comparison. Looking at the 30 day rolling volatility, it is clear that Bitcoin's volatility is consistently higher than gold's throughout the entire study. Bitcoin also goes through frequent volatility spikes, whereas gold's volatility stays relatively low and steady. This is a key point because an asset considered a "store of value" is generally expected to hold its worth with stable risk levels. This specific pattern suggests that Bitcoin still behaves more like a high risk asset rather than a stable, gold-like store of value.

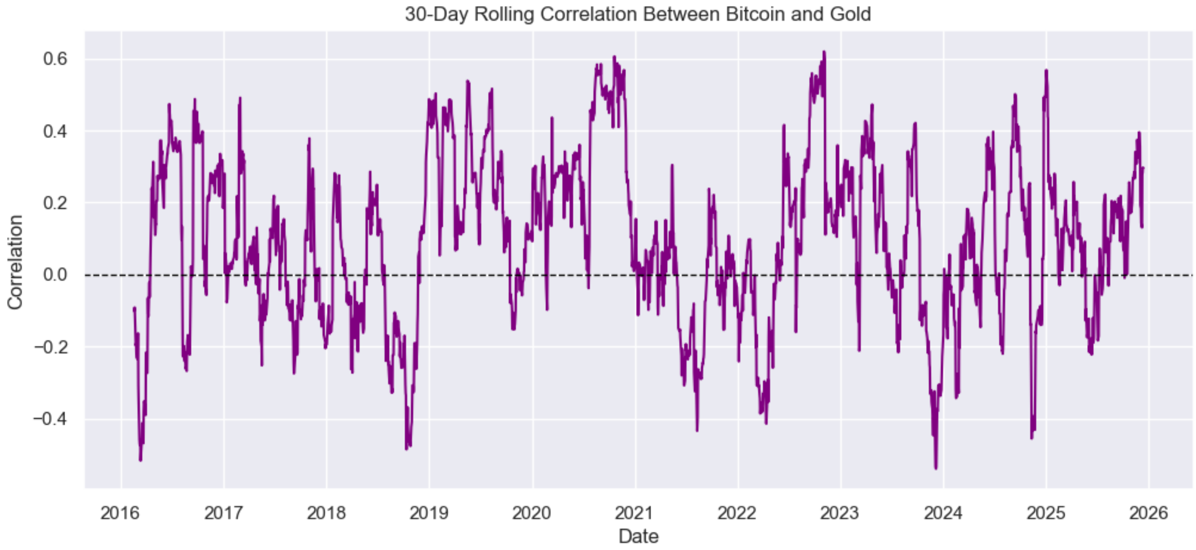
Figure 1. *30-day rolling volatility: Bitcoin vs. gold.*



The Bitcoin gold relationship is not persistent (Phase 2)

The 30 day rolling correlation between the log returns of Bitcoin and gold tends to turn around zero and often flips between positive and negative. While there are short windows where the correlation turns positive, these periods don't last, and the relationship never settles into a consistently positive level. This is an important finding because a strong "digital gold" argument would require a much more stable connection between the two assets. Instead, these results show that any similarities in their price movements are temporary and depend on the specific timeframe, rather than showing a long term structural link.

Figure 2. *30-day rolling correlation between Bitcoin and gold log returns.*



Daily return prediction (Phase 3A):

In this task, the models that are trained to predict Bitcoin’s daily log returns were unable to beat a simple baseline and actually resulted in negative R^2 values. This is an interesting outcome because it shows that daily Bitcoin returns are very noisy in this specific setup. It also suggests that gold related data does not offer reliable information for predicting Bitcoin's price movements over a short timeframe.

- Baseline : MAE 0.022636, RMSE 0.030604, R^2 -0.002456
- Linear Regression: MAE 0.022733, RMSE 0.030809, R^2 -0.015922
- Random Forest: MAE 0.022857, RMSE 0.030667, R^2 -0.006553

Volatility prediction (Phase 3B – main ML result):

The second task focused on predicting Bitcoin's next day volatility (using the 30 day rolling volatility), which connects directly to the project's goal of assessing stability. This task showed very strong performance for both the persistence baseline and the Linear Regression model, with R^2 values around 0.968. What is particularly interesting here isn’t just the high accuracy, but what it tells us: the models indicate that Bitcoin’s volatility is primarily driven by its own recent behavior.

- Baseline: MAE 0.000762, RMSE 0.001402, R^2 0.967756

- Linear Regression: MAE 0.000789, RMSE 0.001398, R^2 0.967927
- Random Forest: MAE 0.001069, RMSE 0.001892, R^2 0.941294

The feature importance from the Random Forest model strongly supports this idea. The variable BTC_vol30 dominates the model with a score of 0.986738, while the gold-related variables only contribute a tiny amount (Gold_vol30 at 0.005310, and other gold features between 0.001 and 0.003). This is a key finding because if Bitcoin truly acted like gold as a stable store of value, you would expect gold-related variables to play a much larger role in the model. Instead, the data suggests that Bitcoin's risk levels are mostly internal and not closely tied to how gold behaves.

Limitations

Gold proxy selection: I used COMEX Gold Futures (GC=F) to represent gold prices. While this is a very common proxy in financial research, it is not exactly the same as spot gold (XAU/USD). Because futures prices involve contract structures and rollover dynamics, they can sometimes show slightly different return behavior than the physical metal itself.

Different trading calendars: One major limitation was that Bitcoin trades 24/7, including weekends, while gold futures only trade on exchange days. To handle this, I aligned the datasets by matching dates and dropping any days where one asset was not trading. While this ensures a fair comparison, it also means some data was removed, which could potentially impact very short-term comparisons.

Daily frequency and noise: This study relied on daily data. Daily returns are notoriously noisy, especially for an asset like Bitcoin, which is likely why the daily return prediction models performed poorly. It is possible that the relationship between these assets would appear more clearly if analyzed at a weekly or monthly frequency instead.

Scope of the features: The analysis was limited to price based features, such as returns, volatility, and lags. I did not include broader factors that might affect both assets, like macroeconomic shifts, changes in liquidity, or general market sentiment. Adding these variables might provide a more complete picture of why their relationship changes.

Modeling limitations: For the machine learning phase, I chose relatively straightforward models to ensure the results were easy to interpret and reproduce. While more specialized time-series models might have captured deeper patterns, they were outside the scope of this specific project stage.

Future Work

Compare alternative gold representations: It would be useful to extend the analysis by testing different gold proxies, such as spot gold (XAU/USD) or gold ETFs like GLD. This would help determine if the findings are consistent across different types of gold instruments or if they are specific to futures contracts.

Use multiple time horizons: Future research could repeat this analysis using weekly or monthly returns. Moving away from daily data would likely reduce the amount of noise and make it easier to see if any real co-movement between Bitcoin and gold emerges over longer periods.

Include additional explanatory variables: Adding macroeconomic and market indicators such as inflation data, the USD index, equity market returns, or liquidity measures could reveal more. This would help show if Bitcoin behaves more like gold during specific economic shifts or under certain market regimes.

Apply time-series specific models: To get a more technical look at how these assets interact, future stages could use models specifically built for financial time series. For example, a Vector Autoregression (VAR) model could capture multivariate relationships.