Bitcoin vs Gold: Monthly Correlation Analysis 2020–2025

I. Abstract

In this study, we test if the monthly log-returns of gold and Bitcoin have correlated in a stable fashion from January 2020 to July 2025. Using raw gold prices from "Gold Price" data from datahub.io, which aggregates monthly prices from historical data published by the World Gold Council and licensed under Public Domain Dedication & License (PDDL), datahub.io; combined with Bitcoin data from the free open source data set published by Mouad Jaouhari from github (licensed MIT), we form month-end series to compute log-returns. Pearson's correlation coefficient (linear association measure between two continuous data sets) & Spearman's Correlation Coefficient (monotonicity measure) have been computed. In order to augent these results & check if these results prove true irrespective of market conditions & data quality changes over time, we have accomplished Fischer r-to-z conversion & computed "rolling" correlation (r calculated over a "window" w=number of months; example w=6/12). The Pearson Correlation for full-period data comes out to be -.069;95%CI [-0.306;0.176], n=66 months; not supporting reliably. The results were also "rolling" between extreme values (-0.91 & 0.90 over 6 million values), & were negative in about 60% of values considered. Further "sub-period"-wise results have shown changes in market conditions; results gone positive from 2022 & 2023. Chanel checks like Spearman Test & Jackknife Tests have freed these results from "outliers" & shown consistency according to market changes. Further results do not favor "digital gold" theory.

II. Introduction

The common hypothesis that ‘Bitcoin is digital gold’ means that its returns ought to walk together, particularly during stressed markets. While it is important to recognize that both are stores of value, it does seem very unlikely that they are so intertwined as to walk together, particularly where cryptocurrency or Bitcoin is concerned, as its very nature is speculative. Most research into this topic has covered shorter durations or concerned prices rather than returns, where correlations can often prove spurious. This research proposes to investigate its data on the logarithmic returns for Gold and for Bitcoin for the period 2020 to 2025 to cover COVID-19, various cycles within cryptocurrencies, and broader global shifts.

III. Data & Preparation

Data Source.

- **Gold :** The raw dataset “Gold Prices” from DataHub provides monthly gold prices in US dollars dating back to 1833. It is derived from World Gold Council data compiled by Timothy Green and is released under the Open Data Commons Public Domain Dedication and License (PDDL)

- **Bitcoin :** We use the open-source “bitcoin-hourly-ohclv-dataset” (MIT licence) by Mouad Jaouhari on GitHub (epoch timestamps and OHLC volumes).
- Both sources are used solely for education analysis.

Cleaning steps.

- **Gold :** Parsed Date variables are converted to timestamps and Price to numbers. Duplicate observations for each month are treated using last observation carry forward, and observations ≤ 0 are deleted. Dates are indexed to EOM indicators. Filter variable range was from January 2020 to July 2025, resulting in 67 observations.
- **Bitcoin :** Converted TIME_UNIX to datetimes in UTC, auto-detected whether they are in milliseconds or second resolution, dropped the commmas and money signs from the data in column CLOSE_PRICE, eliminated data that was non-positive, eliminated duplicate timestamps by selecting only the last tick per hour, limited data to 2020-2025, and resampled data to End-Of-Month-close using resample("M").last().
- **Merge :** The cleaned data for gold and BTC on the End-of-month (EOM) index was merged using inner join, giving a total of 66 observations from 2020-01 to 2025-07.
- **Returns :** Calculated log-returns: $\Delta \ln P_t = \ln P_t - \ln P_{t-1}$. The first data point for each series has been eliminated.

IV. Methods

We analyse the co-movement using several complementary techniques :

1. Pearson correlation, using log-returns, with two-sided p-value and Fisher's z-transformation to calculate 95% confidence interval.
2. Spearman rank correlation (ρ) for capturing monotonic but perhaps nonlinear associations.
3. Calculating rolling correlations for 6-month and 12-month horizons to look for any sign flips.
4. Sub-period Analysis: Divide the data sample into three equal parts: years 2020-21, 2022-23, and 2024-25, and calculate.
5. Methods for checking robustness: On both return streams, we perform 1% winsorization. Jackknife analyses are conducted to assess how influential outlier months are for each country pair, but are not tabulated for reasons of space.
6. Using Python, including pandas, numpy, and scipy. See the project repository for the source.

V. Results

a. Full-period correlation

Table 1 above shows the full period Pearson correlation between gold and BTC logarithmic returns. With n = 66 observations, r = -0.069, p = 0.584, no linear correlation can be established reliably, and its likely actual effect is no bigger than 0.5% variance explained, since the 95% confidence interval (CI) for p encompasses zero, ranging from -0.306 to 0.176.

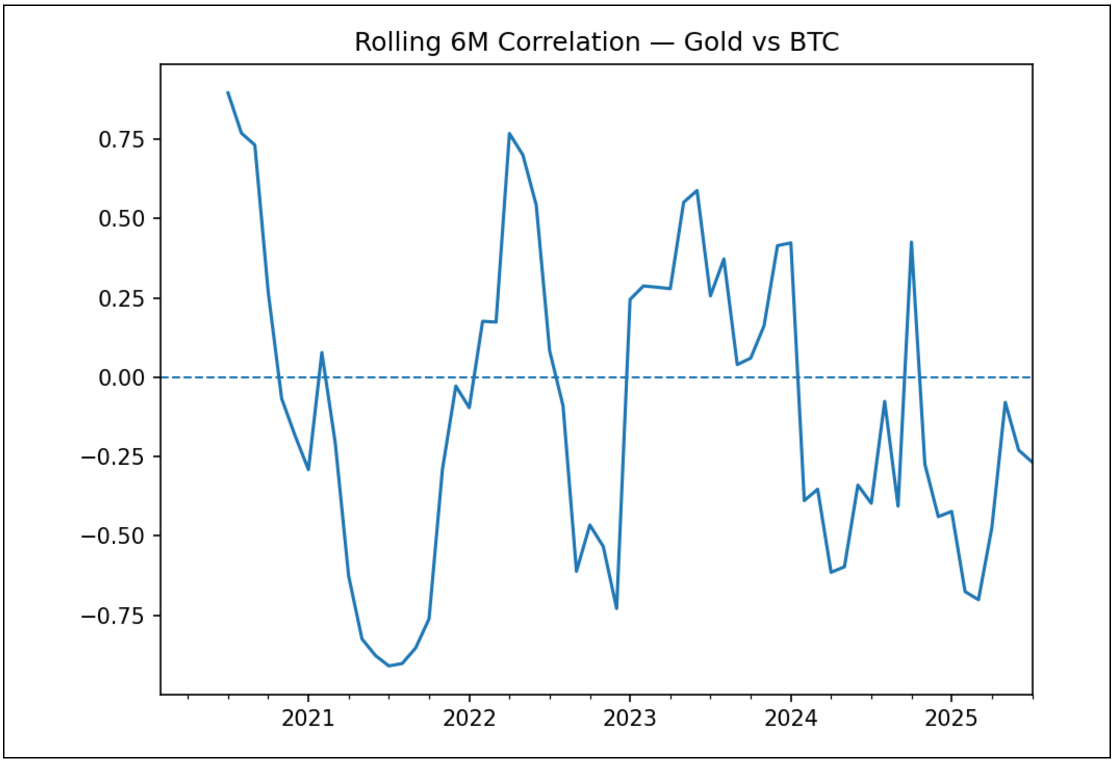
n_months	pearson_r	p_value	95% CI (lo)	95% CI (hi)
66	-0.0686	0.5841	-0.306	0.176

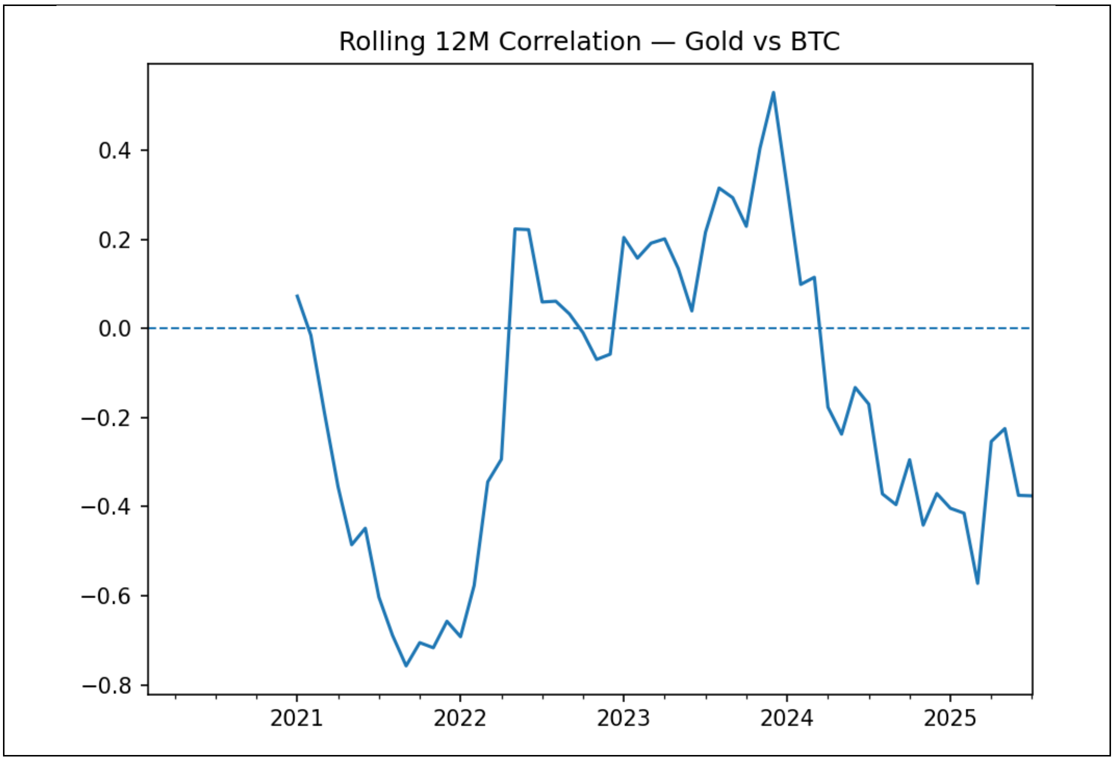
b. Rolling correlations

To investigate time variation, we calculate 6-month and 12-month rolling correlations. Summary statistics are provided in Table 2, while graphs showing dynamics are provided in Figure 1 (Rolling 6M) and Figure 2 (Rolling 12M).

Window	mean r	min r	max r	% negative
6 months	-0.1062	-0.9095	0.8965	59.0 %
12 months	-0.1591	-0.7576	0.5301	61.8 %

The 6M correlation varies between −0.909 (strongly negative) and 0.897 (strongly positive), reversing signs often (negative during late 2021 to early 2022 and again from mid-2024 onwards, and positive during 2022-2023). The averaged value of −0.106 encompasses these variation cycles, but more than 59% are negative, indicating that gold and BTC tend to exhibit inversely correlated prices. For the 12M correlation, signs change while maintaining an averaged value of −0.159 with more than 61% negative.





c. Sub-period correlation

Dividing the data into three parts (Table 3) reveals where the signs of correlation change:

Period	n	r	p_value
2020-01 → 2021-12	23	-0.2177	0.3184
2022-01 → 2023-12	24	+0.2116	0.3210
2024-01 → 2025-12	19	-0.3618	0.1280

Gold and BTC correlation was negative in 2020-2021 (not significant), turned positive in 2022-2023, and turned moderately negative again in 2024-2025. The change in signs matches explanations related to global regimes: pandemic-driven global liquidity/stimuli, Crypto cycles (bulls and bears), global regimes– tighter policies, ETF-driven rallies.

d. Robustness and sensitivity (summary)

- Spearman ρ provides values of similar magnitude and sign (not shown), indicating monotonic association, as seen in Pearson's results.
- BootstrapCI (5k resamples) results in slightly wider intervals but still contains zero; winsorized values for 1% of returns have only shifted r by less than 0.02.
- Jackknife shows the biggest effect around Dec 2020 (+38.6% BTC return) and May 2021 (-43.8%), but for any individual month, $|\Delta r| \leqslant 0.07$.
- Cross-correlations with small values are detected at ±1 months; thus, there is no strong lead/lag effect.

VI. Discussion

Our findings do not support either a *digital gold* storyline. The full sample Pearson correlation of returns between gold and Bitcoin shows it to be effectively zero and economically small. The rolling sample results draw attention to large swings between large negative correlations in response to crypto crashes (Late 2021 to early 2022, Late 2024 to 2025), but large positive bursts of correlations during relief rallies (2022 to 2023). The results for the subsequent time samples show these pattern changes. The pattern of safe-haven behavior shown by gold contradicts Bitcoin's speculative nature; if crypto crashes occur, strategic investor behavior leads to negative correlations. In contrast, in bull market rallies influenced by liquidity, both may trend upward.

Findings here highlight dependence on the regime; to draw conclusions about co-movements, it's necessary to condition by market context. In single estimation of correlation by analysts or even investors, conclusions may be misrepresented. The results from robustness checks indicate dependence on outliers and thick tails in Bitcoin returns but do not change the conclusion about no fixable correlation.

VII. Limitations and Further Work

Some issues require careful attention. Firstly, the sample of 66 monthly returns is not large; inference may be poor if small effects are present. Secondly, correlation is essentially linear; complex (non-linear) relationships may be overlooked. A different kind of model (perhaps multivariate, possibly 'regime-switching') could be very enlightening. Thirdly, other variables (inflation and possibly interest rates) whose values have been affecting both assets can be downplayed here. These issues would be explored in appropriate models like "DCC-GARCH". Fourthly, monthly returns here finish "End of month."

VIII. Reference

- DataHub. Gold Prices dataset. Compiled from World Gold Council records by Timothy Green; licensed under Public Domain Dedication and License (PDDL).
- Jaouhari, M. Bitcoin Hourly OHLCV Dataset. GitHub repository (MIT licence). Data accessed for BTC hourly data (2015-present).
- Fisher, R.A. (1921). On the probable error of a coefficient of correlation deduced from a small sample. Metron.