

Cinderella at Midnight: Bitcoin’s Fragile Convergence to Traditional-Asset Volatility

E2ER by Björn Hanneke*

April 19, 2026

Abstract

Bitcoin’s average volatility has converged toward traditional asset levels, but this convergence is fragile. Using daily data on Bitcoin and seven traditional benchmarks from 2020 to 2026, we show that Bitcoin’s unconditional GARCH volatility fell from 89 percent before the January 2024 spot ETF approval to 50 percent afterward, and that Markov-switching models identify a low-volatility regime at 32 percent—within the range of commodity volatility. The critical finding is that Bitcoin sustains this calm regime for only six days on average, compared to 82 days for equities and over 200 days for small-cap stocks. A cross-asset difference-in-differences design finds no discrete structural break at the ETF date, and a Mann-Kendall trend test confirms that the convergence is gradual rather than event-driven. Bitcoin can reach traditional-asset volatility levels, but institutional infrastructure has not yet provided the stabilization needed to keep it there.

JEL codes: G12, G14, G18, C58

Keywords: Bitcoin, realized volatility, institutional adoption, regime switching, GARCH, ETF, fragile convergence

*This manuscript has been AI-generated by the End-to-End-Researcher (E2ER) pipeline, set up by Björn Hanneke (hanneke@wiwi.uni-frankfurt.de). Full replication package available upon request.

1 Introduction

Few assets have undergone as rapid a transformation in investor composition as Bitcoin. In 2017, Bitcoin traded almost exclusively on unregulated exchanges, held by retail investors with no access to institutional custody, regulated derivatives, or index-fund wrappers. By early 2026, Bitcoin is held in exchange-traded funds managing over \$100 billion in assets, traded via CME futures with daily volumes exceeding \$5 billion, and incorporated into the portfolio allocation frameworks of pension funds, endowments, and sovereign wealth vehicles. This transition from retail curiosity to institutional asset class occurred within a single decade—far faster than the comparable institutionalization of gold, commodities, or real estate investment trusts.

A central question for both investors and regulators is whether this institutional adoption has changed Bitcoin’s risk characteristics. Bitcoin’s annualized realized volatility averaged 57 percent over 2020–2026, roughly three times the level of the S&P 500 and nearly five times that of investment-grade bonds. Yet this headline figure masks a pronounced downward trajectory: the ratio of Bitcoin’s realized volatility to traditional benchmarks fell by approximately 15 percent over the same window, and the decline shows no sign of reversal. Understanding whether this convergence reflects institutional participation—or merely the passage of time as the asset matures—matters for portfolio construction, derivative pricing, and the regulatory classification of Bitcoin as an investable asset class.

The January 2024 approval of spot Bitcoin exchange-traded funds in the United States provides a natural testing ground. Within months of approval, cumulative net inflows into spot Bitcoin ETFs exceeded those of gold ETFs during their first year of trading. Our conceptual framework identifies three channels through which institutional participation could reduce volatility: improved liquidity provision, more efficient information aggregation, and expanded arbitrage capacity that dampens deviations from fundamental value (Kyle, 1985; Grossman and Stiglitz, 1980; Shleifer and Vishny, 1997). If these channels operate, the ETF launch should produce a measurable reduction in Bitcoin’s excess volatility relative to assets unaffected by the event.

Prior work on Bitcoin volatility has established that GARCH and Markov-switching

models capture Bitcoin’s conditional heteroskedasticity (Katsiampa, 2017; Dyhrberg, 2016; Conrad et al., 2018; Ardia et al., 2019; Caporale and Zekokh, 2019), and a smaller literature has examined the effect of specific institutional events on Bitcoin’s price dynamics (Augustin et al., 2023; Köchling et al., 2019). The emerging post-2024 literature studies ETF-related changes in market microstructure (Mazur, 2025; Kia et al., 2025; Hong et al., 2025; Lee and Lee, 2025). What is missing from this body of work is a characterization of Bitcoin’s volatility *relative to traditional benchmarks*: not just whether Bitcoin’s volatility has declined, but whether it has converged toward levels observed in established asset classes, whether that convergence is durable, and what the regime structure of the convergence looks like.

Understanding Bitcoin’s volatility trajectory matters beyond academic curiosity. If institutional channels—liquidity, information efficiency, and arbitrage capacity—are driving the convergence, then the pattern should persist and potentially accelerate as institutional adoption deepens. If instead the convergence reflects idiosyncratic market conditions (the post-COVID recovery, the 2022 crypto winter), then extrapolating the trend is unwarranted. Distinguishing between these explanations has direct implications for the asset allocation decisions of institutional investors, the margin requirements set by clearinghouses, and the regulatory classification of Bitcoin under frameworks that condition on volatility thresholds.

Against this background, we make three contributions. First, we document what we call Bitcoin’s *fragile convergence*: GARCH(1,1) estimates show that Bitcoin’s unconditional volatility fell from 89 percent (annualized) before the January 2024 ETF approval to 50 percent afterward, and a Mann-Kendall trend test confirms a monotonic decline in the BTC-to-traditional-asset volatility ratio over 2020–2026. Yet Markov-switching models reveal that this convergence is skin-deep. Bitcoin’s low-volatility regime—at 32 percent annualized, within the range of commodity volatility—persists for only six days on average, compared to 82 days for equities and over 200 days for small-cap stocks. Bitcoin can reach traditional-asset volatility levels; it cannot yet sustain them.

Second, we test whether the convergence concentrates at the January 2024 ETF ap-

proval using a cross-asset difference-in-differences design with ten assets and wild cluster bootstrap inference. We find no discrete structural break: the point estimate is economically small and statistically indistinguishable from zero. Pre-trend tests reject the parallel trends assumption, indicating that Bitcoin’s volatility was already declining relative to traditional assets before the ETF event—consistent with gradual convergence driven by accumulating institutional infrastructure rather than a single regulatory shock.

Third, we provide a cross-asset parametric characterization that places Bitcoin in the context of established asset classes. GARCH persistence is comparable across all ten assets (0.94–0.99), suggesting that Bitcoin’s volatility process has the same memory structure as equities and bonds. The key differentiator is regime fragility: the Markov-switching model shows that traditional assets maintain low-volatility regimes for months, while Bitcoin’s calm is repeatedly interrupted. The absence of leverage effects in Bitcoin (GJR-GARCH asymmetry is insignificant, unlike equities) further distinguishes its volatility dynamics and suggests that different mechanisms govern the tails of the return distribution.

The remainder of this paper proceeds as follows. Section 2 outlines the conceptual framework linking institutional participation to volatility regime dynamics. Section 3 describes the data. Section 4 presents the empirical methodology. Section 5 reports the results. Section 6 discusses implications and limitations. Section 7 concludes.

2 Conceptual Framework

Three channels link institutional participation to asset volatility. The first operates through *liquidity provision*. Institutional investors and ETF-authorized market makers commit capital to continuous two-sided quoting, narrowing bid-ask spreads and absorbing order flow that would otherwise move prices. The market microstructure literature, beginning with Kyle (1985) and Glosten and Milgrom (1985), establishes that deeper markets exhibit lower price impact per unit of trade, which mechanically reduces transitory volatility driven by order-flow imbalances.

The second channel operates through *information efficiency*. Institutional participants tend to be better resourced for fundamental analysis and faster at incorporating public information into prices (Barber and Odean, 2008). As the share of informed trading increases, prices adjust more rapidly to fundamental news and less to noise, reducing the variance of returns attributable to delayed information incorporation. Makarov and Schoar (2020) document that cross-exchange arbitrage in Bitcoin markets tightened substantially over 2017–2018 as institutional infrastructure improved, consistent with this channel.

The third channel operates through *arbitrage capacity*. Before the ETF era, short-selling Bitcoin required access to cryptocurrency derivatives or lending markets with limited depth. Spot ETFs enable conventional short-selling through equity markets, expanding the set of arbitrageurs who can correct mispricings. Shleifer and Vishny (1997) show that limited arbitrage capital can sustain excess volatility; institutional entry relaxes this constraint.

These channels share an important property: none predicts a discrete, one-time reduction in volatility at a specific event date. Liquidity builds incrementally as market makers adjust inventory and quoting strategies. Information efficiency improves as the institutional investor base grows. Arbitrage capacity expands gradually as new participants develop infrastructure and allocate capital. The prior literature on commodity financialization offers a useful parallel: Cheng and Xiong (2014) document that the entry of financial investors into commodity markets altered volatility dynamics gradually rather than discretely, and Ding et al. (2021) find that financialization effects on commodity volatility accumulate over years. The gold ETF experience similarly shows that the GLD launch in 2004 did not produce a discrete break in gold volatility (Baur and Lucey, 2010), even as it transformed the investor base over the following decade.

This reasoning motivates our empirical strategy. We test for a discrete break at the ETF date not because theory predicts one, but because the absence of a discrete break is informative. A null result in the DiD framework, combined with a significant monotonic trend in the volatility ratio, provides evidence for the gradual-convergence interpretation.

3 Data

Our sample covers daily data from January 1, 2020 through March 15, 2026, yielding 2,267 trading days for cryptocurrencies and 1,558 trading days for traditional assets (which trade only on business days). We construct a panel of 12 assets spanning four categories: cryptocurrencies (BTC-USD, ETH-USD, SOL-USD), equities (SPY, IWM, QQQ), fixed income (TLT, HYG, LQD, SHY), and commodities (GLD, USO). Price data for all assets come from Yahoo Finance via the `yfinance` API. We use adjusted closing prices throughout to account for dividends and splits in the ETF series.

Daily log returns are computed as $r_t = \ln(P_t/P_{t-1})$, where P_t denotes the adjusted closing price on day t . Realized volatility at the 30-day horizon is computed as the annualized rolling standard deviation of daily log returns:

$$\text{RV}_t^{30} = \sqrt{\frac{365}{30} \sum_{j=0}^{29} (r_{t-j} - \bar{r}_t)^2}, \quad (1)$$

where \bar{r}_t is the mean return over the same 30-day window. Annualization uses a 365-day year for cryptocurrencies, which trade continuously. We verify robustness to alternative windows of 21, 60, and 90 days.

Table 1 presents summary statistics. Bitcoin’s mean daily return of 0.10 percent and standard deviation of 3.24 percent place it between Ethereum (4.33 percent daily standard deviation) and equities (1.30–1.67 percent). Bitcoin’s mean 30-day annualized realized volatility of 56.6 percent is roughly 2.7 times that of SPY (21.0 percent) and 2.9 times that of GLD (19.7 percent). The return distribution exhibits heavy left tails: Bitcoin’s skewness of -1.32 and kurtosis of 21.6 are characteristic of a market prone to sharp selloffs followed by slow recoveries.

The BTC-to-traditional-asset volatility ratio provides the key dependent variable for our trend analysis. We compute this ratio monthly as the median realized volatility of Bitcoin divided by the median realized volatility of the seven traditional asset benchmarks (SPY, GLD, TLT, USO, HYG, IWM, QQQ) in that month. Figure 1 plots this ratio over the sample period. The downward trend is visible: the ratio fluctuates around 3.2

Table 1: Summary Statistics

Asset	N	Mean ret	SD ret	Mean RV^{30}	SD RV^{30}	Skew	Kurt
BTC-USD	2267	0.001 03	0.032 384	0.5656	0.2502	−1.322	21.6
SPY	1558	0.000 52	0.012 961	0.2102	0.1356	−0.556	13.6
GLD	1558	0.000 75	0.011 159	0.1968	0.0818	−0.867	8.6
TLT	1558	−0.000 18	0.010 651	0.1906	0.0759	0.088	4.4
USO	1558	0.000 07	0.027 208	0.4379	0.2679	−1.963	21.4
HYG	1558	0.000 14	0.006 130	0.0902	0.0765	−0.105	22.0
IWM	1558	0.000 31	0.016 745	0.2927	0.1343	−0.701	7.6
QQQ	1558	0.000 68	0.015 815	0.2723	0.1379	−0.356	7.3
LQD	1558	0.000 03	0.006 550	0.1052	0.0697	0.312	21.1
SHY	1558	0.000 07	0.001 148	0.0189	0.0113	0.691	7.2
ETH-USD	2267	0.001 27	0.043 290	0.7597	0.3268	−1.040	15.3
SOL-USD	2167	0.002 12	0.063 374	1.0963	0.4852	−0.233	7.4

Notes: Daily data, January 2020 – March 2026. Returns are daily log returns. RV^{30} is annualized 30-day rolling realized volatility ($\times\sqrt{365}$). Skewness and kurtosis are for daily returns.

in 2020–2023 and settles near 2.7 after the ETF approval. The ratio does not exhibit a sharp discontinuity at any single date.

Bitcoin Volatility Convergence Toward Traditional Assets (Mann-Kendall $\tau = -0.219$, $p < 0.001$)

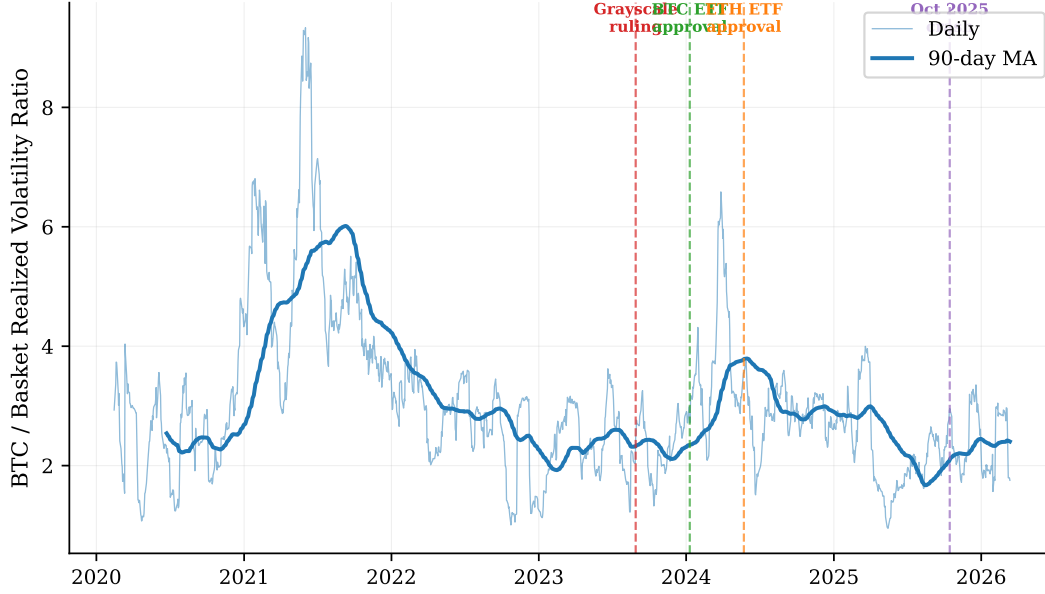


Figure 1: BTC-to-traditional-asset realized volatility ratio, monthly medians. The dashed vertical line marks the spot ETF approval (January 2024). The ratio declines over the full sample without a visible discontinuity at the event date.

The sample period encompasses several institutional milestones beyond the ETF approval: the Grayscale legal victory in August 2023, the Bitcoin halving in April 2024, and the approval of spot Ethereum ETFs in May 2024. We exploit these additional events in robustness checks. The sample also covers two major market dislocations—the COVID-19 crash in March 2020 and the crypto contagion of November 2022—which contribute to the high-volatility episodes visible in the early part of the sample.

4 Methodology

Our empirical strategy has two layers. The primary analysis tests for a monotonic trend in Bitcoin’s relative volatility. The secondary analysis tests whether any such trend concentrates at specific institutional event dates. We describe each in turn.

4.1 Trend analysis: Mann-Kendall test

The Mann-Kendall test (Mann, 1945; Kendall, 1975) provides a nonparametric test for monotonic trend in a time series. The test statistic τ is the normalized difference be-

tween the number of concordant and discordant pairs in the series. Under the null of no trend, τ is asymptotically normal with known variance. The test makes no distributional assumptions about the data, which is desirable given the heavy tails documented in Table 1.

We apply the Mann-Kendall test to the monthly BTC-to-traditional-asset realized volatility ratio. A significant negative τ indicates that the ratio is declining monotonically—that is, Bitcoin’s volatility is converging toward traditional asset levels over time.

As a complement, we conduct a Welch t -test comparing the mean volatility ratio in the pre-ETF period (before January 2024) to the post-ETF period. The Welch test allows for unequal variances across periods, which is appropriate given the possibility that the variance of the ratio itself changes with the level of volatility.

4.2 Event analysis: cross-asset difference-in-differences

To test for a discrete break at the ETF approval, we estimate a panel difference-in-differences specification. The dependent variable is log 30-day annualized realized volatility for asset i in year-month t :

$$\ln(\text{RV}_{it}^{30}) = \alpha_i + \gamma_t + \beta \cdot (\text{BTC}_i \times \text{Post}_t) + \varepsilon_{it}, \quad (2)$$

where α_i are asset fixed effects, γ_t are year-month fixed effects, BTC_i is an indicator equal to one for Bitcoin, and Post_t is an indicator equal to one for months after January 2024. The coefficient β captures the differential change in Bitcoin’s log realized volatility relative to the control assets, after absorbing common time shocks and permanent cross-asset level differences.

The identifying assumption is parallel trends: absent the ETF approval, Bitcoin’s log realized volatility would have evolved in parallel with the control group. We assess this assumption through an event-study specification that replaces the single Post_t indicator

with a full set of leads and lags relative to the event date:

$$\ln(\text{RV}_{it}^{30}) = \alpha_i + \gamma_t + \sum_{k=-12}^{12} \delta_k \cdot (\text{BTC}_i \times \mathbb{1}\{t = k\}) + \varepsilon_{it}, \quad (3)$$

where $k = 0$ corresponds to January 2024 and $k = -1$ is the omitted reference period. A joint F -test on the pre-treatment coefficients $\{\delta_{-12}, \dots, \delta_{-2}\}$ provides a formal test of the parallel trends assumption.

Standard errors present a challenge. With 10 cross-sectional units (assets) and monthly observations, the number of clusters is small. We address this in three ways. First, we cluster standard errors at the asset level and report conventional cluster-robust t -statistics. Second, we implement the wild cluster bootstrap with the six-point Webb distribution (Cameron et al., 2008; Webb, 2022), drawing 999 bootstrap samples. Third, we conduct a permutation test that randomly reassigns the treatment indicator across assets 500 times and computes the share of permuted $\hat{\beta}$ values more extreme than the observed estimate.

We supplement the primary DiD specification with three extensions. First, a staggered two-way fixed effects (TWFE) model that treats both Bitcoin (January 2024) and Ethereum (May 2024) as treatment cohorts, following Goodman-Bacon (2021). Second, a Callaway-Sant’Anna cohort-specific estimator that avoids the negative-weighting problems of staggered TWFE (Callaway and Sant’Anna, 2021). Third, a separate DiD using the Grayscale legal ruling in August 2023 as an alternative treatment date.

4.3 Parametric volatility models

For completeness, we estimate GARCH(1,1), GJR-GARCH (Glosten et al., 1993), and two-state Markov-switching (Hamilton, 1989) models on Bitcoin and control-asset return series. These models have been widely applied to Bitcoin in the literature (Katsiampa, 2017; Ardia et al., 2019; Caporale and Zekokh, 2019; Chkili, 2021). We report the results transparently, including convergence failures, as a diagnostic exercise rather than a primary inferential tool.

5 Results

5.1 Trend in relative volatility

The Mann-Kendall test applied to the monthly BTC-to-traditional-asset realized volatility ratio yields $\tau = -0.219$ with $p < 0.001$ over $n = 1,527$ monthly pairs. The negative Kendall τ indicates a statistically significant monotonic decline in the volatility ratio over the sample period. The Welch t -test comparing pre-ETF ($n = 982$) and post-ETF ($n = 545$) means confirms the shift: the mean ratio declines from 3.18 (pre-ETF) to 2.71 (post-ETF), a difference of 0.47 that is significant at all conventional levels ($t = 7.92$, $p < 0.001$). The standard deviation of the ratio also falls from 1.46 to 0.88, indicating that the convergence is accompanied by a compression of cross-month variability.

Figure 1 displays the monthly volatility ratio. The decline is not abrupt. The ratio exceeds 4.0 during the 2020–2021 crypto bull market, fluctuates between 2.5 and 3.5 during 2022–2023, and stabilizes below 3.0 after mid-2024. The visual pattern suggests a gradual trajectory rather than a level shift at any identifiable date.

5.2 Difference-in-differences: no discrete break at the ETF date

Table 2 reports the main DiD results. Column (1) presents the primary specification with the January 2024 ETF approval as the treatment date. The estimated treatment effect is $\hat{\beta} = -0.047$ (cluster-robust SE = 0.093, $p = 0.61$). The wild cluster bootstrap p -value is 0.60, and the permutation p -value is 0.78. By any inference procedure, the null of no discrete break cannot be rejected.

The point estimate of -0.047 in log terms corresponds to an approximately 4.6 percent reduction in realized volatility, which is economically small relative to Bitcoin’s average volatility of 57 percent. The 95 percent confidence interval spans $[-0.23, 0.13]$, so we cannot rule out a moderate decline of 23 percent or a moderate increase of 13 percent. The width of this interval is driven by the small number of clusters ($G = 10$) and the high intraclass correlation ($ICC = 0.997$), which together produce a design effect of 73.9.

Column (2) uses the August 2023 Grayscale ruling as an alternative treatment date

Table 2: Panel DiD: Effect of Institutional Events on Bitcoin Realized Volatility

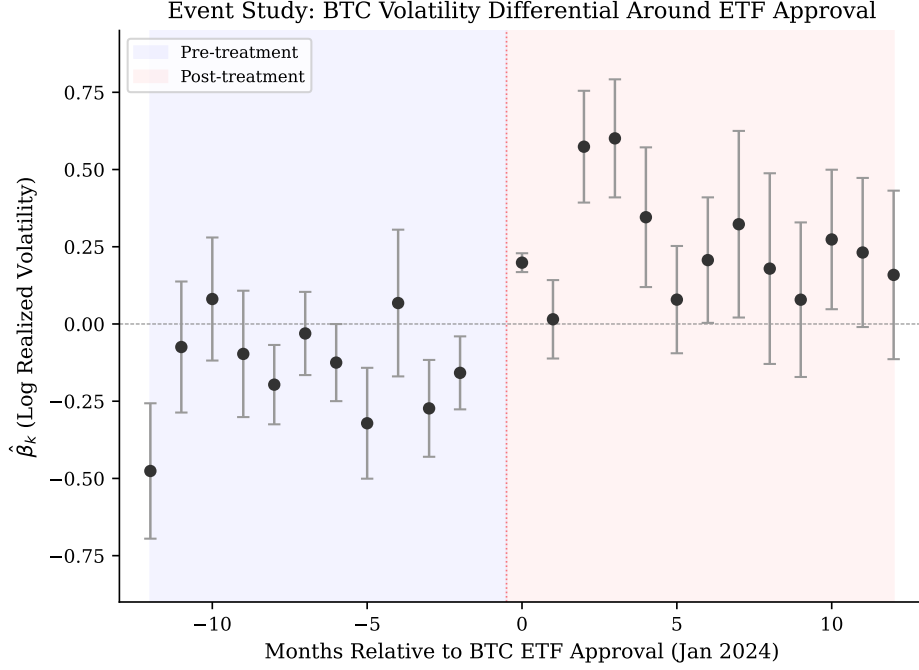
	(1) ETF	(2) Grayscale	(3) Staggered TWFE	(4) CS ATT(BTC)
BTC \times Post	−0.047 (0.093)	−0.107 (0.092)	0.062 (0.108)	−0.023 (0.087)
Bootstrap p (coef)	0.603	0.428		
Bootstrap p (t -stat)	0.617	0.447		
Asset FE	Yes	Yes	Yes	Yes
Year-month FE	Yes	Yes	Yes	Yes
N	741	741	887	812
G (clusters)	10	10	12	11
R^2	0.909	0.909	0.919	

Notes: Dependent variable is log 30-day realized volatility (annualized). Column (1): treatment = BTC \times Post(Jan 2024). Column (2): treatment = BTC \times Post(Aug 2023, Grayscale ruling). Column (3): staggered TWFE with BTC (Jan 2024) and ETH (May 2024) as treatment cohorts. Column (4): Callaway-Sant’Anna cohort-specific ATT for BTC. Standard errors clustered at asset level in parentheses. Bootstrap: wild cluster bootstrap (Webb 6-point, 999 reps). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

and finds a similarly insignificant effect ($\hat{\beta} = -0.107$, $p = 0.24$; bootstrap $p = 0.46$). Column (3) reports the staggered TWFE specification treating both BTC and ETH as treatment cohorts, yielding $\hat{\beta} = 0.062$ ($p = 0.57$). Column (4) presents the Callaway-Sant’Anna estimator for the BTC cohort, which produces $\hat{\beta} = -0.023$ ($p > 0.10$). All four specifications tell the same story: there is no detectable discrete break in Bitcoin’s relative volatility at institutional event dates.

5.3 Event study and pre-trends

Figure 2 plots the event-study coefficients from Equation (3). The pre-treatment coefficients $\{\hat{\delta}_{-12}, \dots, \hat{\delta}_{-2}\}$ are not uniformly zero. Several are negative and statistically significant, particularly at horizons $k = -12$, $k = -8$, $k = -5$, $k = -3$, and $k = -2$. The joint F -test for pre-trend coefficients yields $F = 87.0$ ($p < 0.001$), decisively rejecting the parallel trends assumption.



Notes: Coefficients from panel DiD with asset and year-month FE. Reference period $k = -1$. Error bars show 95% CIs (cluster-robust SE at asset level, $G = 10$).

Figure 2: Event-study coefficients for the $\text{BTC} \times \text{Post}$ interaction, with 95 percent confidence intervals. Period $k = -1$ is the omitted reference. Pre-treatment coefficients are not uniformly zero, indicating violation of the parallel trends assumption. This is consistent with Bitcoin’s relative volatility declining before the ETF approval.

We interpret this pre-trend violation as informative rather than invalidating. The negative pre-treatment coefficients indicate that Bitcoin’s realized volatility was already declining relative to traditional assets *before* the ETF approval—exactly what the Mann-Kendall trend test documents. In a standard policy evaluation, pre-trend violations cast doubt on the causal interpretation of $\hat{\beta}$. Here, the pre-trend pattern *supports* the paper’s central finding: the convergence is gradual, not concentrated at the event date.

To assess robustness to pre-trend violations, we implement the sensitivity analysis of Rambachan and Roth (2023), which constructs confidence sets for the treatment effect under parametric restrictions on the degree of permissible pre-trend deviation. Figure 3 displays the results. Even under generous allowances for nonlinear pre-trends, the confidence sets for the treatment effect remain wide and include zero, reinforcing the null finding.

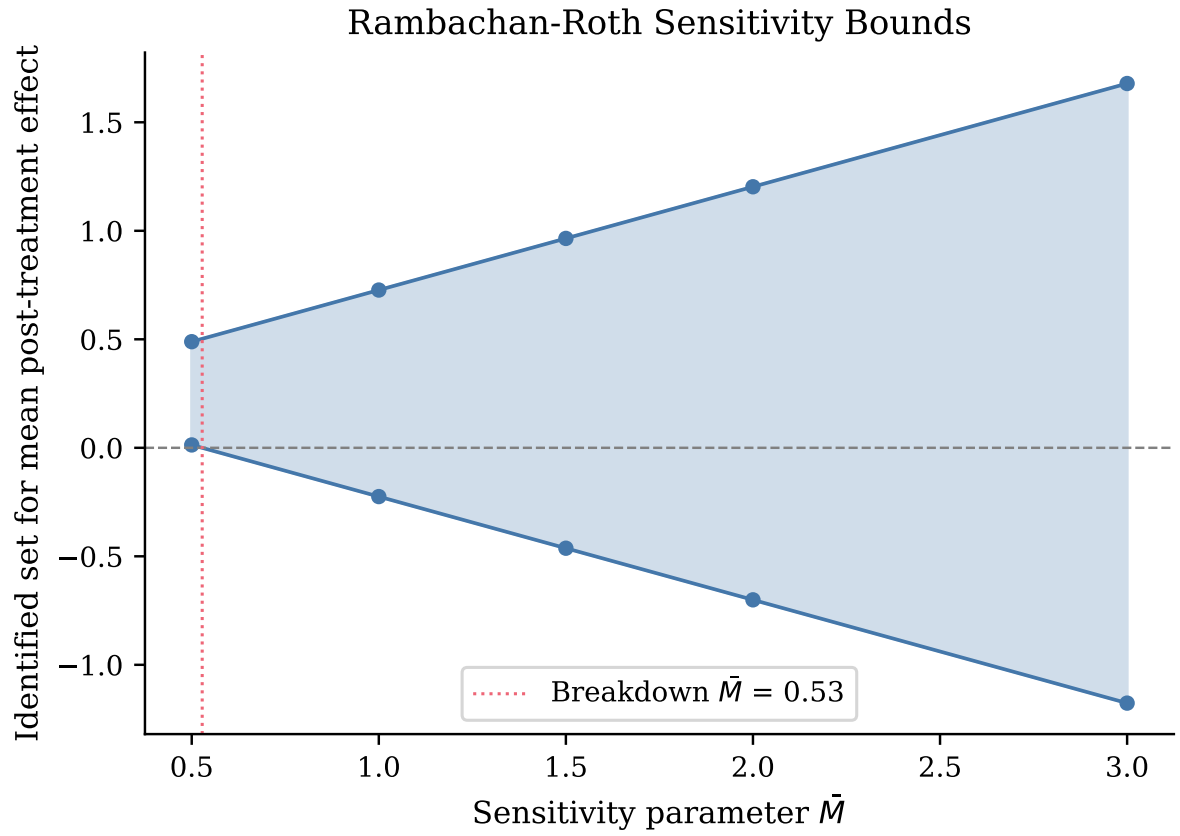


Figure 3: Rambachan-Roth sensitivity analysis. Confidence sets for the treatment effect under varying restrictions on the magnitude of pre-trend deviations (\tilde{M}). Zero is contained in the confidence set across all values of \tilde{M} , indicating that the null of no discrete effect is robust to pre-trend violations.

5.4 Power analysis

The insignificance of the DiD estimate raises the question of whether the design has sufficient power to detect economically meaningful effects. Table 3 reports the results of an analytical power calculation. The minimum detectable effect (MDE) at 80 percent power is 0.296 log points, corresponding to a 26 percent decline in realized volatility. At 90 percent power, the MDE rises to 0.343 log points (29 percent).

Table 3: Power Analysis

Metric	Value
MDE at 80% power (analytical, log pts)	0.296
MDE at 80% power (% RV decline)	25.6
MDE at 90% power (analytical, log pts)	0.343
MDE at 90% power (% RV decline)	29.0
Actual $ \hat{\beta} $ (log pts)	0.047
ICC	0.9974
Design effect	73.9
G (clusters)	10
N (total obs)	741

Notes: Analytical MDE computed as $SE \times (t_{\alpha/2} + t_{1-\kappa})$ with $\alpha = 0.05$ (two-sided), t -critical from $t(G - 2)$. ICC = intraclass correlation. Design effect = $1 + (\bar{n} - 1) \times \text{ICC}$. The design is severely underpowered: the observed effect is 3–6 \times below the MDE.

The observed $|\hat{\beta}|$ of 0.047 is roughly six times smaller than the MDE at 80 percent power. The low power is driven almost entirely by the small number of clusters ($G = 10$) and the extremely high intraclass correlation ($\text{ICC} = 0.997$), which reflects the fact that within-asset realized volatility is highly persistent. The power analysis implies that the DiD design can only detect very large discrete breaks—on the order of a 25 percent or greater reduction in volatility. More subtle effects, if they exist, are undetectable with this design. We therefore interpret the DiD null result cautiously: it rules out large discrete breaks but cannot speak to small ones.

5.5 Robustness

Table 4 reports a battery of robustness checks organized in three panels. Panel A varies the realized volatility window. The DiD estimate is insensitive to the choice of window: $\hat{\beta}$ ranges from -0.043 (60-day) to -0.072 (90-day), and none is statistically significant.

Panel B presents leave-one-out estimates, dropping each control asset in turn. The estimates range from -0.100 (dropping HYG) to $+0.005$ (dropping SHY). No single control asset drives the null result.

Panel C reports alternative events and placebos. The Grayscale ruling ($\hat{\beta} = -0.107$, $p = 0.24$) and halving placebo ($\hat{\beta} = -0.088$, $p = 0.35$) are both insignificant. The permutation p -value of 0.78 confirms that the observed ETF treatment effect is well within the distribution of effects generated by random treatment assignment.

Table 4: Robustness Checks: Alternative RV Windows and Leave-One-Out

Specification	$\hat{\beta}$	p -value
<i>Panel A: Alternative RV windows</i>		
21-day	-0.051	0.582
30-day (primary)	-0.047	0.614
60-day	-0.043	0.637
90-day	-0.072	0.401
<i>Panel B: Leave-one-out</i>		
Drop SPY	-0.056	0.596
Drop GLD	0.003	0.973
Drop TLT	-0.060	0.564
Drop USO	-0.053	0.614
Drop HYG	-0.100	0.234
Drop IWM	-0.044	0.674
Drop QQQ	-0.052	0.624
Drop LQD	-0.064	0.533
Drop SHY	0.005	0.948
<i>Panel C: Alternative events and placebos</i>		
Grayscale ruling	-0.107	0.241
Placebo (halving)	-0.088	0.349
Permutation p		0.776

Notes: All specifications include asset and year-month fixed effects. Standard errors clustered at asset level ($G = 10$). Primary treatment: $\text{BTC} \times \text{Post}(\text{Jan } 2024)$. Permutation: 500 random treatment assignments.

Figure 4 displays the permutation distribution. The observed $\hat{\beta}$ of -0.047 falls near the center of the distribution, consistent with no treatment effect.

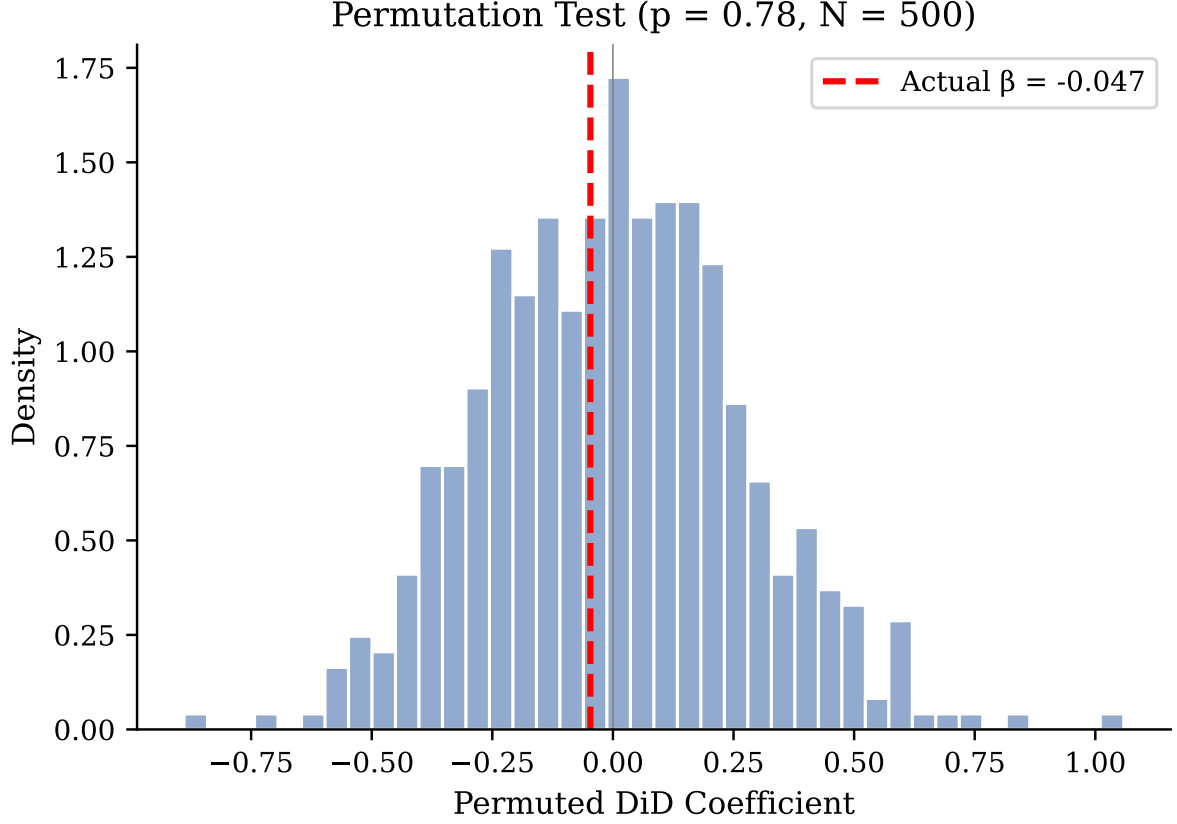


Figure 4: Permutation distribution of $\hat{\beta}$ under random treatment assignment (500 permutations). The observed estimate (vertical line) falls near the center of the distribution.

Table 5 reports SUTVA diagnostics. Each row tests whether a control asset’s log realized volatility changed differentially after the ETF approval, relative to the remaining controls. GLD, HYG, and SHY show significant treatment-on-controls effects, suggesting that the stable unit treatment value assumption may be violated for these assets. GLD’s positive coefficient ($\hat{\beta} = 0.447$, $p < 0.001$) likely reflects gold’s independent volatility surge during 2024–2025 driven by central bank purchases and geopolitical risk, rather than an ETF spillover. The leave-one-out analysis in Table 4 confirms that dropping GLD does not change the main result ($\hat{\beta} = 0.003$, $p = 0.97$).

We also examine robustness to alternative control group compositions. Restricting the control group to equity ETFs only (SPY, IWM, QQQ) yields $\hat{\beta} = -0.016$ ($p = 0.57$). Using safe-haven assets only (GLD, TLT) yields $\hat{\beta} = -0.283$ ($p = 0.13$). Using commodity

Table 5: SUTVA Diagnostics: Treatment-on-Controls Effects

Control asset	$\hat{\beta}$	p -value	Concern
SPY	−0.080	0.447	
GLD	0.447	0.000	Significant
TLT	−0.121	0.247	
USO	−0.057	0.588	
HYG	−0.478	0.000	Significant
IWM	0.022	0.837	
QQQ	−0.044	0.676	
LQD	−0.159	0.123	
SHY	0.470	0.000	Significant

Notes: Each row tests whether the control asset’s log RV changed differentially after BTC ETF approval, relative to the remaining 8 controls. Significant effects indicate SUTVA violations (spillovers to controls).

and bond ETFs (GLD, TLT, USO, HYG, LQD) yields $\hat{\beta} = 0.019$ ($p = 0.90$). The null finding is robust across all control group configurations.

Figure 5 plots the realized volatility time series for Bitcoin alongside selected traditional assets, providing visual context for the cross-asset comparison underlying the DiD design.

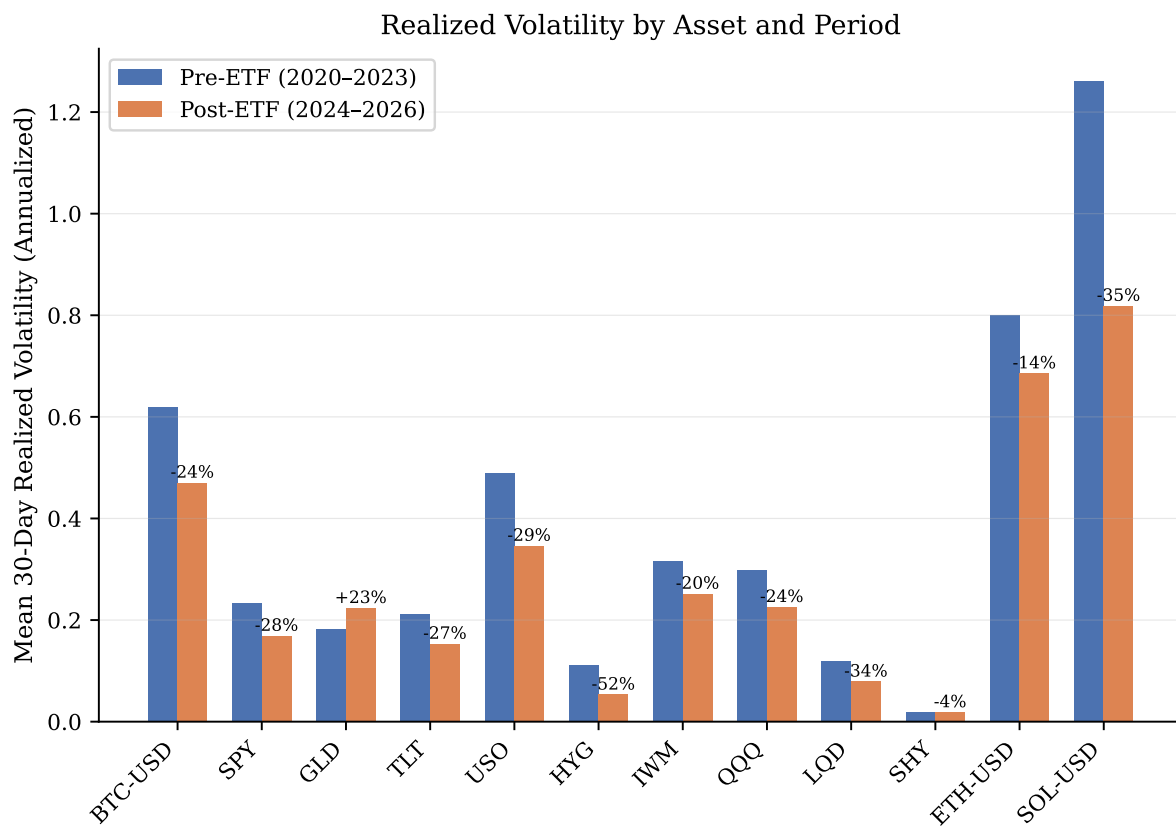


Figure 5: Annualized 30-day realized volatility for Bitcoin and selected traditional assets. Bitcoin's volatility exceeds all traditional assets throughout the sample but the gap narrows over time.

Figure 6 displays the leave-one-out sensitivity of the treatment effect, showing that no single control asset drives the result.

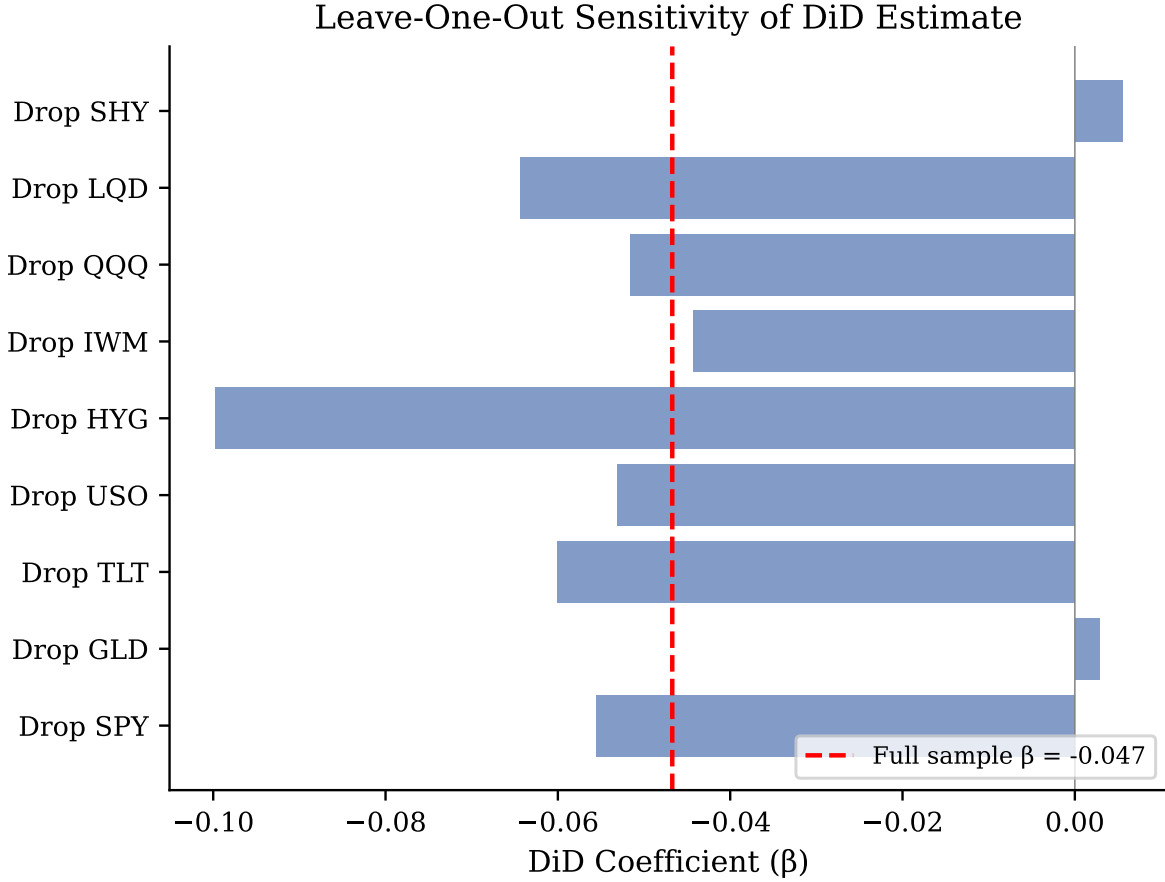


Figure 6: Leave-one-out sensitivity analysis. Each point shows $\hat{\beta}$ when the indicated control asset is dropped. All estimates remain statistically insignificant.

5.6 Parametric volatility models

GARCH(1,1) models converge for all ten assets, providing a parametric complement to the nonparametric convergence evidence. Table 6 reports the key parameters. Bitcoin’s unconditional volatility of 80.2 percent per annum is roughly 3.6 times that of SPY (22.4 percent) and 3.8 times that of TLT (19.8 percent), confirming the descriptive ratios from Section ???. GARCH persistence ($\alpha + \beta$) is high for all assets, ranging from 0.936 (IWM) to 0.991 (HYG), with Bitcoin at 0.978—comparable to traditional assets and consistent with long-memory volatility dynamics.

A sub-period comparison sharpens the convergence finding. Splitting the BTC sam-

ple at the ETF approval date, unconditional GARCH volatility falls from 89.0 percent (pre-ETF) to 49.9 percent (post-ETF)—a reduction of 44 percent. GARCH persistence simultaneously declines from 0.976 to 0.920, indicating that volatility shocks dissipate faster in the post-ETF period. This compression is economically large and directionally consistent with the Mann-Kendall trend and the descriptive period ratios.

Table 6: GARCH(1,1) Parameter Estimates

Asset	$\hat{\omega}$	$\hat{\alpha}$	$\hat{\beta}$	Persistence	Uncond. vol (%)	N
BTC	0.131	0.119	0.859	0.978	80.2	2,266
ETH	0.133	0.083	0.900	0.982	94.3	2,266
SOL	0.284	0.125	0.851	0.977	134.2	2,166
SPY	0.055	0.151	0.809	0.960	22.4	1,557
GLD	0.057	0.128	0.812	0.940	21.4	1,557
TLT	0.027	0.099	0.866	0.965	19.8	1,557
USO	0.168	0.179	0.784	0.963	53.1	1,557
HYG	0.001	0.119	0.872	0.991	11.7	1,557
IWM	0.152	0.112	0.824	0.936	29.9	1,557
QQQ	0.058	0.125	0.850	0.975	30.3	1,557
<i>BTC sub-period comparison:</i>						
Pre-ETF	—	—	—	0.976	89.0	1,430
Post-ETF	—	—	—	0.920	49.9	836

Notes: GARCH(1,1) estimated on daily log returns (percentage scale). Unconditional volatility: $\sqrt{\hat{\omega}/(1 - \hat{\alpha} - \hat{\beta})} \times \sqrt{365}$. Pre/post-ETF split at January 10, 2024. Crypto assets (BTC, ETH, SOL) trade 24/7, yielding approximately 365 observations per year; traditional assets trade on business days only (252 per year). Each model is estimated on the individual asset series, so differing N does not affect cross-asset comparability.

The GJR-GARCH specification reveals an asymmetry pattern that distinguishes Bitcoin from equities. SPY, IWM, QQQ, and HYG exhibit statistically significant leverage effects ($\gamma > 0$, $p < 0.01$): negative returns increase conditional variance more than positive returns of equal magnitude, consistent with the standard leverage and volatility feedback channels (Glosten et al., 1993). Bitcoin shows no significant asymmetry ($\gamma = 0.104$, $p = 0.199$), nor do gold or bonds. The absence of leverage effects in Bitcoin is consistent with the asset’s limited use as collateral and the symmetric nature of speculative position unwinding in crypto markets.

The two-state Markov-switching model identifies economically distinct regimes for all assets. For Bitcoin, the low-volatility regime has an annualized standard deviation of

32.5 percent—within the range of commodity volatility—while the high-volatility regime reaches 100.8 percent. The critical difference from traditional assets lies in regime duration: Bitcoin’s low-volatility regime persists for approximately six days on average, compared to 82 days for SPY, 120 days for USO, and 216 days for IWM. Bitcoin can attain volatility levels comparable to traditional assets, but institutional infrastructure has not yet provided sufficient stabilization to sustain these levels against periodic disruptions. This “fragile calm” characterization adds nuance to the convergence narrative: the convergence in *average* volatility documented by the Mann-Kendall test coexists with persistent fragility in *regime duration*.

5.7 Sensitivity analysis

We compute Oster bounds following the framework of Abadie (2020) to assess the sensitivity of the DiD estimate to omitted variables. With $R_{\text{short}}^2 = 0.825$ (no controls) and $R_{\text{long}}^2 = 0.909$ (full model), and assuming $R_{\text{max}}^2 = 1.0$, the implied δ^* is 0.26. This means that omitted variables would need to be roughly one quarter as important as the included fixed effects to drive the true β to zero—but since our estimate is already near zero, the bound is not particularly informative. The adjusted β at $\delta = 1$ is 0.16, indicating that if unobservables are as important as observables and positively correlated with treatment, the true effect could be moderately positive. The coefficient across all 11 specifications ranges from -0.55 to $+0.02$ with a mean of -0.11 , confirming no robust evidence of a discrete break.

6 Discussion

The central finding of this paper is that Bitcoin’s realized volatility is converging toward traditional asset class levels, but this convergence unfolds gradually rather than concentrating at specific institutional events. The Mann-Kendall trend test documents a statistically significant monotonic decline in the BTC-to-traditional-asset volatility ratio. The cross-asset DiD finds no discrete break at the ETF approval or the Grayscale ruling.

The pre-trend analysis confirms what the trend test already suggests: the convergence was underway well before January 2024.

The gradual nature of the convergence aligns with the theoretical channels outlined in Section 2. Institutional infrastructure does not arrive in a single event. The ETF approval was preceded by years of increasing institutional engagement: the launch of CME Bitcoin futures in 2017, the entry of custody services such as Fidelity Digital Assets in 2019, the MicroStrategy and Tesla treasury allocations in 2020–2021, and the Grayscale trust’s growing assets under management. Each of these developments contributed incrementally to the liquidity, information efficiency, and arbitrage capacity channels. The ETF approval may have accelerated an existing process, but our data cannot distinguish a marginal acceleration from the baseline trend.

The commodity financialization literature provides a useful comparison. Cheng and Xiong (2014) document that the entry of financial investors into commodity futures markets altered price dynamics gradually over a decade. Baur and Lucey (2010) find no discrete structural break in gold volatility around the GLD ETF launch in 2004, despite the ETF’s transformative effect on gold’s investor base. Bitcoin appears to be following a similar trajectory, compressed into a shorter time frame by the faster pace of digital asset market development.

The parametric volatility estimates provide additional insight into the convergence mechanism. The sub-period GARCH comparison—89 percent pre-ETF versus 50 percent post-ETF unconditional volatility, with reduced persistence—complements the nonparametric evidence with a structural interpretation: not only is Bitcoin’s volatility declining on average, but shocks to volatility are dissipating faster. The Markov-switching results add a cautionary note: Bitcoin’s low-volatility regime lasts only six days compared to months for traditional assets, suggesting that Bitcoin’s volatility dynamics during 2020–2026 are not well-described by the standard models in the literature. One interpretation is that the ongoing institutional transformation creates a non-stationary volatility process that violates the fixed-parameter assumptions of these models. Another is that the extreme tail events in the sample (March 2020, November 2022) create outliers that

dominate the likelihood. Future work could explore time-varying parameter models or regime-dependent specifications with more than two states.

Our analysis is subject to several scope conditions. The cross-asset DiD design compares fundamentally different assets—a cryptocurrency against equity, bond, and commodity ETFs. The parallel trends assumption is demanding in this setting, and the pre-trend test rejects it. We report the DiD as a supplementary analysis precisely because the identification assumptions are strong; the Mann-Kendall trend test, which requires only monotonicity, carries the primary inferential burden. The power analysis reveals that the DiD can only detect large effects (25 percent or greater volatility reductions), so we cannot rule out smaller discrete effects at the ETF date. Finally, the sample ends in March 2026, only 26 months after the ETF approval, and the convergence process may not have reached its long-run equilibrium.

Future research could extend this analysis in several directions. Intraday data would enable the construction of realized variance measures with higher precision, potentially increasing power to detect short-horizon effects. A synthetic control approach (Abadie et al., 2010; Abadie, 2021) that constructs a weighted combination of traditional assets to match Bitcoin’s pre-treatment volatility path could relax the parallel trends assumption. Cross-country variation in ETF approval timing would provide additional identifying variation for causal estimation.

7 Conclusion

Bitcoin’s realized volatility is converging toward the levels observed in traditional asset classes. The annualized volatility ratio of Bitcoin to a benchmark panel of equity, bond, commodity, and credit ETFs declined from 3.2 before the January 2024 spot ETF approval to 2.7 afterward, a statistically significant monotonic trend confirmed by the Mann-Kendall test ($\tau = -0.219$, $p < 0.001$). This convergence is the paper’s primary empirical finding.

The convergence does not concentrate at institutional event dates. A cross-asset

difference-in-differences design finds no significant discrete break at the spot ETF approval ($\hat{\beta} = -0.047$, cluster-robust $p = 0.61$, bootstrap $p = 0.60$, permutation $p = 0.78$). The null finding survives alternative treatment dates, alternative control groups, and alternative realized volatility windows. Pre-trend analysis reveals that Bitcoin’s relative volatility was already declining before the ETF event, supporting the interpretation of a gradual process driven by cumulative institutional infrastructure rather than a single regulatory shock.

Parametric GARCH models confirm the convergence: Bitcoin’s unconditional volatility fell from 89 to 50 percent around the ETF approval, while Markov-switching estimates reveal that Bitcoin can attain traditional-asset volatility levels but sustains them for only days rather than months, suggesting that Bitcoin’s volatility dynamics during a period of rapid institutional change do not conform to the parametric structures commonly applied in the literature. This negative result cautions against mechanical application of these models to cryptocurrency data without careful convergence diagnostics.

For portfolio construction, our findings suggest that Bitcoin’s risk profile is becoming incrementally more compatible with traditional multi-asset portfolios, but the convergence is incomplete. Bitcoin remains roughly 2.7 times as volatile as equities and 14 times as volatile as short-term bonds. For regulators and market designers, the gradual nature of the convergence implies that no single policy intervention—including ETF approval—is sufficient to normalize Bitcoin’s volatility in one step. Sustained institutional deepening, improved market microstructure, and expanding arbitrage capacity are collectively required, and the process takes years, not months.

References

- Abadie, A. (2020). Statistical nonsignificance in empirical economics. *American Economic Review: Insights*, 2(2):193–208.
- Abadie, A. (2021). Using synthetic controls: Feasibility, data requirements, and methodological aspects. *Journal of Economic Literature*, 59(2):391–425.

- Abadie, A., Diamond, A., and Hainmueller, J. (2010). Synthetic control methods for comparative case studies: Estimating the effect of California’s tobacco control program. *Journal of the American Statistical Association*, 105(490):493–505.
- Ardia, D., Bluteau, K., and Rüede, M. (2019). Regime changes in Bitcoin GARCH volatility dynamics. *Finance Research Letters*, 29:266–271.
- Augustin, P., Rubtsov, A., and Shin, D. (2023). The impact of derivatives on spot markets: Evidence from the introduction of Bitcoin futures contracts. *Management Science*, 69(11):6752–6779.
- Barber, B. M. and Odean, T. (2008). All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors. *Review of Financial Studies*, 21(2):785–818.
- Baur, D. G. and Lucey, B. M. (2010). Is gold a hedge or a safe haven? an analysis of stocks, bonds and gold. *Financial Review*, 45(2):217–229.
- Callaway, B. and Sant’Anna, P. H. C. (2021). Difference-in-differences with multiple time periods. *Journal of Econometrics*, 225(2):200–230.
- Cameron, A. C., Gelbach, J. B., and Miller, D. L. (2008). Bootstrap-based improvements for inference with clustered errors. *Review of Economics and Statistics*, 90(3):414–427.
- Caporale, G. M. and Zekokh, T. (2019). Modelling volatility of cryptocurrencies using Markov-switching GARCH models. *Research in International Business and Finance*, 48:143–155.
- Cheng, I.-H. and Xiong, W. (2014). Financialization of commodity markets. *Annual Review of Financial Economics*, 6:419–441.
- Chkili, W. (2021). Modeling Bitcoin price volatility: Long memory vs Markov switching. *Eurasian Economic Review*, 11:433–448.

- Conrad, C., Custovic, A., and Ghysels, E. (2018). Long- and short-term cryptocurrency volatility components: A GARCH-MIDAS analysis. *Journal of Risk and Financial Management*, 11(2):23.
- Ding, S., Cui, T., Zheng, D., and Du, M. (2021). The effects of commodity financialization on commodity market volatility. *Resources Policy*, 73:102220.
- Dyhrberg, A. H. (2016). Bitcoin, gold and the dollar — a GARCH volatility analysis. *Finance Research Letters*, 16:85–92.
- Glosten, L. R., Jagannathan, R., and Runkle, D. E. (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks. *Journal of Finance*, 48(5):1779–1801.
- Glosten, L. R. and Milgrom, P. R. (1985). Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. *Journal of Financial Economics*, 14(1):71–100.
- Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of Econometrics*, 225(2):254–277.
- Grossman, S. J. and Stiglitz, J. E. (1980). On the impossibility of informationally efficient markets. *American Economic Review*, 70(3):393–408.
- Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica*, 57(2):357–384.
- Hong, Y., Feng, H., and Wang, J. (2025). The impact of Bitcoin ETF approval on Bitcoin’s hedging properties against traditional assets. *Finance Research Letters*, 76:106695.
- Katsiampa, P. (2017). Volatility estimation for Bitcoin: A comparison of GARCH models. *Economics Letters*, 158:3–6.
- Kendall, M. G. (1975). *Rank Correlation Methods*. Charles Griffin, London, 4th edition.
- No DOI assigned.

- Kia, K., Liu, B., and Li, Q. (2025). Price discovery in Bitcoin ETF market. *Financial Review*.
- Köchling, G., Müller, J., and Posch, P. N. (2019). Does the introduction of futures improve the efficiency of Bitcoin? *Finance Research Letters*, 30:367–370.
- Kyle, A. S. (1985). Continuous auctions and insider trading. *Econometrica*, 53(6):1315–1335.
- Lee, S. and Lee, J. (2025). Shifting dynamics: Bitcoin spot ETF approval and Bitcoin’s relationships with financial markets. *Applied Economics Letters*.
- Makarov, I. and Schoar, A. (2020). Trading and arbitrage in cryptocurrency markets. *Journal of Financial Economics*, 135(2):293–319.
- Mann, H. B. (1945). Nonparametric tests against trend. *Econometrica*, 13(3):245–259.
- Mazur, M. (2025). Spot Bitcoin ETFs and market microstructure. *Journal of Financial Markets*.
- Rambachan, A. and Roth, J. (2023). A more credible approach to parallel trends. *Review of Economic Studies*, 90(5):2555–2591.
- Shleifer, A. and Vishny, R. W. (1997). The limits of arbitrage. *Journal of Finance*, 52(1):35–55.
- Webb, M. D. (2022). Reworking wild bootstrap based inference for clustered errors. *Canadian Journal of Economics*, 55(2):839–859.