

The Crypto Fear Channel: Asymmetric and Regime-Dependent Volatility Spillover between Bitcoin and Equity Markets

Yi-Hao Lai*

September 26, 2026

Abstract

Using daily returns on SPY, BTC-USD, and VIX from 2015-02 to 2026-04 ($N = 2,812$), we document two stylized facts about the crypto-equity volatility spillover and one negative result. The channel is bidirectional: symmetric Granger tests reject non-causality in the $VIX \rightarrow BTC$ RV direction at every lag from 1 to 10, and at lag 1 that reverse leg is the stronger of the two ($F = 7.77$, $p = 0.00535$, against a forward $F = 0.59$, $p = 0.444$); the $BTC \rightarrow VIX$ leg dominates from lag 2 onward. Our findings characterize that dominant leg and do not license a one-way reading. First, the spillover is *asymmetric*: Bitcoin downside volatility Granger-causes VIX at lags 1–5, whereas Bitcoin upside volatility does not (symmetric Granger tests covering lags 1–10 corroborate the forward leg at lags 2–10). Second, the spillover is *regime-dependent*: in a five-subperiod breakdown, only 2020 survives Bonferroni correction ($F = 12.31$, $p < 10^{-6}$); 2015–2017 is marginally significant at the uncorrected 5% level ($p = 0.014$) and the remaining three subperiods are clearly non-significant. The negative result concerns tail concentration. A quantile regression of VIX on lagged BTC realized variance alone shows a slope that is negative in the lower quantiles and large and positive in the upper tail, but that pattern does not survive once lagged VIX is controlled: the upper-tail slope falls to $+0.42$ at $\tau = 0.95$ with a 95% moving-block bootstrap interval of $[-0.61, +1.14]$, and what remains is a small negative partial slope at $\tau \leq 0.5$. The unconditional pattern is therefore absorbed by the persistence of VIX itself and carries no evidence of tail amplification. A generalized (order-invariant) connectedness decomposition places Bitcoin only marginally on the receiving side (net -0.95 percentage points, negative in 72.5% of rolling windows), too weakly to carry a directional claim on its own; the “fear amplifier conditional on pre-existing equity stress” reading therefore rests on the Granger evidence alone. Out-of-sample forecasts show that adding BTC realized volatility does not improve an AR(3) model of VIX (Diebold-Mariano $t = -0.99$, $p = 0.32$; Clark-West $t = -0.12$ for the nested comparison), highlighting that Granger causality is a necessary but not sufficient condition for forecastability. Taken together, our findings describe the crypto-equity spillover as a *bidirectional, crisis-concentrated* channel whose dominant leg runs from Bitcoin downside volatility to equity fear: informative for ex-post risk attribution, but inadequate as an ex-ante predictor of equity fear.

Keywords: volatility spillover, Bitcoin, VIX, asymmetric causality, quantile regression, Diebold-Yilmaz, crypto-equity linkage

JEL: G11, G15, G17, C22

1 Introduction

Since the institutionalization of Bitcoin trading and the 2024 approval of U.S. spot BTC ETFs, the question of whether cryptocurrency markets transmit fear to traditional equity markets

*Department of Finance, Da-Yeh University, Taiwan. Email: yihao.lai@gmail.com.

has acquired first-order practical importance for portfolio risk management, prudential supervision, and margin design. Yet the empirical literature has yielded an uneasy consensus. On one hand, correlation-based studies report rising BTC-equity co-movement in stress periods, suggesting that cryptocurrency is no longer a standalone asset class.¹ On the other hand, return-predictability tests often fail to find that Bitcoin-based signals improve forecasts of equity volatility at practical horizons. This tension motivates a central question: is the crypto-equity fear channel *real but nonlinear*, *real but regime-dependent*, or an artifact of tail-period correlation without predictive content?

We resolve this tension by decomposing the spillover along three dimensions that the prior literature has typically examined in isolation: directional asymmetry, tail dependence, and regime stability. Asymmetry and regime stability each yield a stylized fact that materially changes how one should interpret correlation-based evidence; the tail-dependence test yields a negative result that is instructive in its own right.

Asymmetry. Using the asymmetric Granger-causality framework of Hatemi-J (2012), we separate Bitcoin positive and negative realized volatility and test whether each Granger-causes VIX at lags 1 through 5. We find that the negative branch carries the entire spillover: Bitcoin downside volatility Granger-causes VIX at every lag tested (1–5), while the upside branch is uniformly non-significant across the same lag window. This is not merely a statistical artifact of a leverage effect — the asymmetry is corroborated by standard symmetric Granger tests over the broader 1 through 10 lag range, where BTC-to-VIX causality is significant at lags 2–10. Our interpretation is behavioral: retail-heavy crypto markets transmit *fear*, not *euphoria*, into the institutional equity complex.

No tail amplification once VIX persistence is controlled. A quantile regression of VIX on lagged BTC realized variance alone produces a striking picture: a slope that is negative in the lower quantiles and rises steeply into the upper tail (+16.29 at $\tau = 0.95$). That picture is an omitted-variable artifact. VIX is highly persistent, and lagged BTC realized variance partly proxies for lagged VIX; once VIX_{t-1} enters the quantile regression, the upper-tail slope falls to +0.42 with a 95% moving-block bootstrap interval of $[-0.61, +1.14]$, and the slope at $\tau = 0.75$ is -0.02 , also insignificant. What survives is a small, negative partial slope at the lower and middle quantiles ($-0.33, -0.20, -0.18$ at $\tau = 0.05, 0.25, 0.50$), which runs *against* a tail-amplification reading. Earlier drafts of this paper reported the unconditional pattern as a headline “sign reversal with $7.0\times$ upper-tail amplification”; that claim is withdrawn (§5.2).

Regime dependence. A five-subperiod breakdown reveals that Granger significance is concentrated in 2020 after multiple-testing correction ($F = 12.31, p < 10^{-6}$), with 2015–2017 only marginally significant at the uncorrected 5% level ($p = 0.014$). This suggests that the spillover is not a permanent feature of cross-market architecture but a *conditional amplification mechanism* that activates mainly during pre-existing equity stress. (Earlier drafts of this paper also invoked a Diebold and Yilmaz (2012) spillover-index result showing Bitcoin to be a strong net receiver. That result was an artifact and has been withdrawn; see §6.2.)

The forecastability gap. A natural next question is whether the in-sample Granger structure translates into out-of-sample predictive power. It does not. Comparing an AR(3) specification of VIX against one augmented with lagged BTC realized volatility over an OOS window of 1,826 days (2019–2026), the Diebold-Mariano test returns $t = -0.99, p = 0.32$, and because the augmented model nests the baseline we also report the Clark and West (2007) statistic,

¹See, e.g., Bouri et al. (2020), Corbet et al. (2018), and Matkovskyy and Jalan (2019) for the early post-2017 literature.

$t = -0.12$, far from its one-sided 5% critical value. This result echoes a broader methodological point: *Granger causality is a necessary but not sufficient condition for forecastability*. An informative signal that is confined to downside innovations and to one crisis regime may be too sparse to outperform a well-specified autoregressive benchmark on average over a long OOS sample.

We make three contributions to the literature. First, within what we show to be a bidirectional channel (§5.1, Table 3), the BTC-to-VIX leg is carried entirely by Bitcoin *downside* volatility, and the COVID-2020 subperiod is the only window that survives Bonferroni multiple-testing correction across five non-overlapping subperiods, so the channel is asymmetric and regime-conditional rather than a permanent feature of cross-market architecture. An order-invariant connectedness decomposition places Bitcoin only marginally on the receiving side (net -0.95 percentage points), too weakly to carry a directional claim on its own; the reading of Bitcoin as a fear *amplifier* (recipient) rather than originator (sender) therefore rests on the Granger evidence alone. Second, we show that an apparently strong tail pattern — a quantile slope that switches sign and rises steeply into the upper tail of VIX — disappears once the persistence of VIX is controlled (§5.2); unconditional quantile regressions of a persistent fear index on a cross-market predictor can manufacture tail heterogeneity that a quantile-Granger specification removes. That Granger evidence is itself two-sided — the equity-to-Bitcoin leg rejects at every lag tested — so what we defend is *predominance* in the BTC-to-VIX direction, not exclusivity.² Third, we pair the in-sample findings with an out-of-sample evaluation under Diebold-Mariano and, because the models are nested, Clark-West comparisons, and the resulting OOS null (DM $t = -0.99$, $p = 0.32$; CW $t = -0.12$) supplies a useful counterweight against promoting in-sample structure into unsupported forecasting claims (Conrad and Kleen, 2020). Each contribution rests on the four methodological building blocks — asymmetric Granger causality, quantile regression for tail dependence, rolling VAR spillover decomposition, and Diebold-Mariano forecast comparison — developed in §4; the novelty lies in their joint application to the BTC-equity fear channel — treated throughout as bidirectional with a dominant BTC-to-VIX leg — and in the empirical patterns they collectively reveal.

The remainder of the paper is organized as follows. Section 2 surveys the crypto-equity spillover literature and positions our contribution. Section 3 describes the data and preliminaries. Section 4 details the four core methodological building blocks (asymmetric Granger, quantile regression, Diebold-Yilmaz, Diebold-Mariano), preceded by a symmetric-Granger baseline. Sections 5 and 6 present the main results and robustness checks. Section 7 reports the out-of-sample forecasting evaluation. Section 8 discusses mechanisms and policy implications. Section 9 concludes.

2 Literature Review

We organize the related literature along three threads: (i) the empirical literature on crypto-equity volatility spillover and the safe-haven debate, (ii) the methodological literature underlying our four building blocks (asymmetric Granger, quantile regression, Diebold-Yilmaz, Diebold-Mariano), and (iii) the recent calls for honest out-of-sample evaluation in volatility forecasting. Each thread has progressed independently; our contribution is to combine them within a single framework that decomposes the spillover into asymmetry, tail concentration, regime conditionality, and predictive power.

²An earlier draft stated a much stronger directional claim at this point, taken from a spillover-index result that has since been withdrawn as a coding artifact; see §6.2.

2.1 Crypto-equity volatility spillover and the safe-haven debate

The empirical literature on Bitcoin’s role in equity portfolios has evolved through three phases. The first phase, exemplified by Corbet et al. (2018) and Matkovskyy and Jalan (2019), documented that cryptocurrency price movements respond meaningfully to macro-financial news and that contagion-style co-movement with equity markets emerges in stress periods, contradicting the early “digital gold” framing. The second phase, motivated by COVID-19 and reviewed in Bouri et al. (2020), Conlon and McGee (2020), Shahzad et al. (2019), and Klein et al. (2018), sharpened the safe-haven question with mixed verdicts: wavelet- and DCC-based correlation analysis showed that Bitcoin’s risk profile is closer to a high-volatility risky asset than to gold, and that its safe-haven properties are at best regime-dependent. The third phase, post-2020, has documented increasing institutional integration: Yarovaya et al. (2022) report that information transmission between cryptocurrency and traditional asset markets intensified during the pandemic, and Akyildirim et al. (2020) link implied-volatility shocks to cryptocurrency returns.

Three gaps motivate our analysis. First, most prior studies test *symmetric* causality or *symmetric* correlation, conflating the upside and downside volatility channels that we will show carry economically distinct information. Second, the tail dependence of the spillover is typically inferred indirectly from rolling DCC or wavelet coherence, rather than directly estimated through quantile regression. Third, sample windows often end at 2020 or include 2020 as a single observation, leaving the post-COVID regime-stability question unaddressed — a gap that matters precisely because COVID-2020 will turn out to be the structural watershed of the spillover. Our analysis closes these gaps by replacing each symmetric or correlation-based summary with a direct, regime-aware, asymmetric counterpart, while extending the sample through April 2026.

2.2 Methodological building blocks

Our four building blocks each have a settled methodological literature. Asymmetric Granger causality, the technique behind our headline asymmetry result, was formalized by Hatemi-J (2012), who decompose each series into positive and negative cumulative innovations and test their respective Granger predictability. Earlier nonparametric extensions of Granger causality (Diks and Panchenko, 2006) and asymmetric volatility-spillover tests (Hong, 2001) provide complementary nonparametric diagnostic tools for asymmetric-transmission testing. Each of these methods was developed to address the specific concern that linear Granger tests can mask sign-dependent transmission — precisely the asymmetry the present paper documents in the BTC-to-VIX channel.

Quantile regression, our tool for tail dependence, originates with Koenker and Bassett (1978) and has subsequently been extended into systemic-risk applications. Most relevant to our setting is Adrian and Brunnermeier (2016)’s CoVaR framework, in which the conditional value-at-risk of one institution given another’s stress is estimated through quantile regression on returns. We adopt the quantile-regression framework rather than CoVaR *per se* because our object of interest is the conditional response of equity *volatility* (VIX), not the conditional value-at-risk of equity returns; but the underlying statistical engine is the same.

The Diebold-Yilmaz spillover index is now a canonical tool for measuring directional volatility transmission across asset classes. Its development spans Diebold and Yilmaz (2009), Diebold and Yilmaz (2012), and Diebold and Yilmaz (2014), and a recent IMF policy note (Iyer, 2022) applies the framework specifically to cryptocurrency-equity spillovers. In the archived implementation used by this draft, we fit a rolling VAR on the ordered system $\{VIX, RV_{\text{spy}}^{(20)}, RV_{\text{btc}}^{(20)}\}$, select lags by AIC within each 252-day window (capped at 5), and compute a 10-step orthogonalized Cholesky FEVD via `statsmodels`. We therefore interpret the resulting spillover shares as ordering-dependent connectedness diagnostics rather than as fully generalized FEVD estimates.

Finally, our forecasting evaluation rests on the Diebold and Mariano (1995) test of equal predictive accuracy with the small-sample adjustment of Harvey et al. (1997). We report the

conventional two-sided p -value and treat the OOS comparison as a separate object from the in-sample discovery exercise, rather than importing factor-discovery thresholds into forecast-comparison inference.

2.3 Honest out-of-sample evaluation in volatility forecasting

A recurring tension in the volatility-forecasting literature is the gap between in-sample statistical significance and out-of-sample economic value. Conrad and Kleen (2020)’s GARCH-MIDAS application demonstrates that incorporating macro-economic information through long-run components can deliver in-sample fit improvements that fail to translate into forecast accuracy gains. The cumulative evidence reviewed there motivates an evaluation discipline in which a new predictor must clear both (i) in-sample significance and (ii) out-of-sample improvement under Diebold-Mariano before being promoted to a structural claim. The present paper adopts this discipline and reports a transparent out-of-sample null — BTC realized volatility fails to improve AR-VIX forecasts under both Diebold-Mariano and Clark-West — alongside its in-sample asymmetry and regime findings and its negative tail-dependence result.

3 Data and Preliminaries

3.1 Sample and variables

Our sample consists of daily data on three series: the SPDR S&P 500 ETF (SPY), Bitcoin in U.S. dollar terms (BTC-USD), and the CBOE Volatility Index (VIX). The reported estimates come from two archived runs of the same experiment. The Granger, subperiod, and EWMA-correlation results come from `k1025_v2.py`, which originally downloaded all three series from Yahoo Finance via `yfinance` with `auto_adjust=True`; its data layer has since been frozen to a pinned snapshot (`experiments/k1025/data/spy_btc_usd_vix_2015-2026-04-08.csv`), from which it reproduces the archived values within a relative tolerance of 5×10^{-4} . The connectedness, quantile-regression, and out-of-sample results come from `k1025_v3.py`, which reads only the paper-local snapshot (`paper/crypto-fear-channel/data/spy_btc_usd_vix_2015-2026.csv`, `auto_adjust=False`), uses log returns for both SPY and BTC, and does no live fetching. The two runs differ in panel construction (§8.4), and the rebuilt run does not yet have its own reproduction specification; for both reasons this manuscript remains a revision draft rather than a submission-ready replication package. The full sample spans 2015-02-02 through 2026-04-08, yielding $N = 2,812$ trading-day observations after aligning BTC to the SPY/VIX trading calendar (we drop weekend BTC observations because VIX is undefined). The out-of-sample window used in §4.5 begins 2019-01-01 and runs through 2026-04-08, providing 1,826 OOS observations — spanning the COVID-19 shock (2020), the late-cycle bull market (2021), the 2022 bear, and the post-ETF recovery (2023–2026).

3.2 Constructed variables

We work with three constructed series. In the current archived implementation, SPY returns are computed as simple daily percentage changes, while BTC returns are computed as log returns from adjusted closes: $r_t^{\text{spy}} = P_t^{\text{spy}}/P_{t-1}^{\text{spy}} - 1$ and $r_t^{\text{btc}} = \log(P_t^{\text{btc}}/P_{t-1}^{\text{btc}})$. Realized volatility is then constructed as a 20-day rolling standard deviation of those return series, annualized by $\sqrt{252}$: $\text{RV}_t^{(20)} = \text{std}_{[t-19, t]}(r) \cdot \sqrt{252}$, computed separately for SPY and BTC. The VIX is used in level (it is already an annualized measure, %). For all forward-looking specifications we lag explanatory variables by one day (`signal.shift(1)` convention) to enforce causal predictability and rule out look-ahead.

For the asymmetric Granger tests of §4.2 we follow Hatemi-J (2012)’s sign-decomposition approach with one adaptation: rather than test the cumulative-positive and cumulative-negative

innovations of BTC realized volatility directly, we first decompose BTC *returns* into positive and negative components ($r_t^{\text{btc},+} = \max(r_t^{\text{btc}}, 0)$ and $r_t^{\text{btc},-} = -\min(r_t^{\text{btc}}, 0)$) and then construct directional realized-volatility series from each component using the same 20-day rolling quadratic-mean as in the symmetric specification: $\text{RV}_t^{\text{btc},\pm,(20)} = \sqrt{\frac{1}{20} \sum_{j=t-19}^t (r_j^{\text{btc},\pm})^2} \cdot \sqrt{252}$. We take first differences of these directional RV series to ensure stationarity in the Granger specification (the level series carry near-unit-root persistence in 20-day windows), and use the first-differenced series $\Delta \text{RV}_t^{\text{btc},\pm,(20)}$ as the explanatory variables in the asymmetric Granger tests. This adaptation preserves the key identification of Hatemi-J (2012) — separate sign-conditional Granger tests — while allowing the volatility-channel interpretation that motivates our application.

3.3 Descriptive statistics and stationarity

Table 1 reports descriptive statistics for the four primary series. BTC daily returns exhibit mean 0.158%, daily standard deviation 3.79%, and excess kurtosis 11.83 with pronounced negative skewness (−0.74). SPY daily returns are an order of magnitude less volatile in standard deviation (1.12%) and exhibit a still-higher excess kurtosis (14.15) in this sample window, concentrated in a handful of COVID-2020 and 2022 crisis days; the two series’ tail thickness is thus of comparable order in this window, with SPY’s edge reflecting a small set of extreme equity outliers rather than a reversal of the well-documented unconditional ranking of BTC over equities. The VIX averages 18.4 with maximum 82.7 (March 2020) and minimum 9.1, spanning the calm 2017–2019 vol regime, the COVID shock, the 2022 inflation regime, and the 2024–2026 normalization.

Table 1: Descriptive statistics for daily series, 2015-02-02 to 2026-04-08 ($N = 2,812$).

Series	Mean	Std. dev.	Skewness	Excess kurtosis
r^{btc} (%)	0.158	3.791	−0.739	11.826
r^{spy} (%)	0.056	1.117	−0.307	14.150
VIX	18.382	7.110	—	—
$\text{RV}_{\text{btc}}^{(20)}$	0.543	0.266	—	—

VIX min/max are 9.14 / 82.69. $\text{RV}_{\text{btc}}^{(20)}$ min/max are 0.098 / 1.960 (annualized). Source: `experiments/k1025/k1025_v2_results.json` .descriptive_statistics (audit fix 2026-06-10: values aligned to the cited v2 source, replacing v1 numbers).

Augmented Dickey-Fuller (ADF) stationarity tests reject the unit-root null for all three modeling series at conventional levels: $\text{RV}_{\text{btc}}^{(20)}$ ($t = -5.10$, $p < 0.001$, lags 20), VIX ($t = -5.71$, $p < 10^{-6}$, lags 9), and $\text{RV}_{\text{spy}}^{(20)}$ ($t = -4.46$, $p < 0.001$, lags 20). Stationarity is a precondition for the Granger and Diebold-Yilmaz analyses below, both of which assume covariance stationarity. We do not differentiate further; the analyses use the level series.

4 Methodology

We employ four methodological building blocks. The first three (§4.2–4.4) characterize *in-sample* cross-asset structure along three dimensions: directional asymmetry, tail dependence, and regime stability. The fourth (§4.5) tests *out-of-sample* predictive content.

4.1 Symmetric Granger causality (baseline)

As a baseline we test whether BTC realized volatility Granger-causes VIX in a standard linear specification with lags $\ell = 1, \dots, 10$. For each lag ℓ we estimate the unrestricted regression

$$\text{VIX}_t = \alpha_0 + \sum_{j=1}^{\ell} \alpha_j \text{VIX}_{t-j} + \sum_{j=1}^{\ell} \beta_j \text{RV}_{\text{btc}, t-j}^{(20)} + \varepsilon_t \quad (1)$$

and the restricted regression that drops the β_j terms. The joint null $H_0 : \beta_1 = \dots = \beta_\ell = 0$ is tested with the standard OLS F -test returned by `statsmodels.tsa.stattools.grangercausalitytests`. We therefore interpret these as plain lag-augmentation Granger diagnostics rather than as HAC-adjusted Wald tests. We tabulate F -statistics and p -values across lags 1 through 10 and treat the symmetric Granger result as the linear-causality benchmark against which the asymmetric and tail-conditional specifications below are compared.

4.2 Asymmetric Granger causality (Hatemi-J)

Following the sign-decomposition approach of Hatemi-J (2012), we test whether the directional RV series defined in §3.2 (the first-differenced rolling-RV series constructed separately from positive-only and negative-only BTC daily returns) carry distinct Granger predictive content for VIX. The asymmetric Granger specifications are estimated as two separate regressions, one per branch:

$$\begin{aligned} \text{VIX}_t &= \alpha_0^+ + \sum_{j=1}^{\ell} \alpha_j^+ \text{VIX}_{t-j} + \sum_{j=1}^{\ell} \beta_j^+ \Delta \text{RV}_{t-j}^{\text{btc}, +, (20)} + \varepsilon_t^+, \\ \text{VIX}_t &= \alpha_0^- + \sum_{j=1}^{\ell} \alpha_j^- \text{VIX}_{t-j} + \sum_{j=1}^{\ell} \beta_j^- \Delta \text{RV}_{t-j}^{\text{btc}, -, (20)} + \varepsilon_t^-. \end{aligned} \quad (2)$$

We test the two null hypotheses, $H_0^+ : \beta_1^+ = \dots = \beta_\ell^+ = 0$ (BTC upside RV innovations do not Granger-cause VIX) and $H_0^- : \beta_1^- = \dots = \beta_\ell^- = 0$ (BTC downside RV innovations do not Granger-cause VIX), each with a standard F -test against the corresponding restricted regression. The Hatemi-J framework is appropriate when the suspected transmission channel is sign-dependent — our prior, motivated by retail-driven crypto trading and the asymmetric leverage effect in equity volatility, is precisely such asymmetry. We test lags $\ell \in \{1, 2, 3, 4, 5\}$.³

4.3 Quantile regression for tail dependence

Linear regression measures the conditional mean response of VIX to BTC RV; given the heavy tails of both series, the conditional mean may be uninformative about state-specific transmission, so we test for tail heterogeneity directly. We follow Koenker and Bassett (1978) and estimate quantile regressions of VIX on lagged BTC realized variance at five representative quantiles $\tau \in \{0.05, 0.25, 0.50, 0.75, 0.95\}$, with and without lagged VIX as a control:

$$Q_\tau(\text{VIX}_t \mid \cdot) = \alpha_\tau + \beta_\tau \cdot \text{RV}_{\text{btc}, t-1}^{(20)} \left[+ \delta_\tau \cdot \text{VIX}_{t-1} \right], \quad (3)$$

estimated via the standard Koenker-Bassett check-loss minimization. The controlled version is the quantile-Granger specification: because VIX is highly persistent, a slope on lagged BTC realized variance without VIX_{t-1} can simply pick up VIX's own dynamics, so only β_τ in the controlled specification measures incremental tail information from Bitcoin. Confidence intervals come from a moving-block bootstrap (1,000 replications, block length $\lceil n^{1/3} \rceil = 15$, seed 42),

³The lag window 1–5 is chosen to match the timescale at which retail-investor sentiment plausibly transmits across asset markets; the symmetric Granger test in §4.1 extends to $\ell = 10$ to confirm robustness across longer horizons.

which respects the serial dependence that an i.i.d. pairs bootstrap would ignore; draws in which the solver fails to converge are excluded and reported. The conceptual link to Adrian and Brunnermeier (2016)’s CoVaR framework is direct: both rely on quantile regression to recover state-dependent cross-asset transmission, with the present paper applying the technique to volatility (VIX) rather than tail risk.

4.4 Diebold-Yilmaz spillover index

To measure directional volatility transmission, we estimate a rolling VAR on the 3-variable system $\{\text{VIX}, \text{RV}_{\text{spy}}^{(20)}, \text{RV}_{\text{btc}}^{(20)}\}$ and compute a 10-step orthogonalized forecast-error variance decomposition (FEVD) within each window. In the archived implementation, the lag order is selected by AIC in each rolling window (capped at 5 for stability), and the FEVD is generated by `statsmodels` using the Cholesky ordering shown above. The directional spillover from variable i to variable j is therefore an ordering-dependent share of variable j ’s forecast-error variance attributable to shocks in variable i . We report the total spillover index, the directional “to” and “from” indices, and BTC’s net spillover (to minus from) across rolling 252-trading-day windows. This should be read as a rolling connectedness diagnostic rather than as a fully generalized FEVD implementation.

Generalized (KPPS) decomposition. The Cholesky shares defined above are not invariant to the ordering of the system, and in our data that is not a hypothetical concern: reordering moves BTC’s net share by 18.66 percentage points and reverses its sign (§5.3). Every connectedness figure reported in this paper is therefore computed from the generalized forecast-error variance decomposition of Koop et al. (1996) and Pesaran and Shin (1998), which does not depend on the ordering. Writing the vector moving-average representation of the estimated VAR as $x_t = \sum_{h=0}^{\infty} A_h \varepsilon_{t-h}$ with $E[\varepsilon_t \varepsilon_t'] = \Sigma$, the H -step generalized share of variable i ’s forecast-error variance attributable to shocks in variable j is

$$\theta_{ij}(H) = \frac{\sigma_{jj}^{-1} \sum_{h=0}^{H-1} (e_i' A_h \Sigma e_j)^2}{\sum_{h=0}^{H-1} e_i' A_h \Sigma A_h' e_i}, \quad (4)$$

where e_i is the i -th selection vector and σ_{jj} the j -th diagonal element of Σ . Generalized shares do not sum to one across j , so we use the row-normalized $\tilde{\theta}_{ij}(H) = \theta_{ij}(H) / \sum_k \theta_{ik}(H)$. The total connectedness index and BTC’s net directional connectedness are

$$C(H) = \frac{100}{N} \sum_{i \neq j} \tilde{\theta}_{ij}(H), \quad C_{\text{btc}}(H) = \sum_{j \neq \text{btc}} \tilde{\theta}_{j,\text{btc}}(H) - \sum_{j \neq \text{btc}} \tilde{\theta}_{\text{btc},j}(H), \quad (5)$$

with $N = 3$ and $H = 10$. The Cholesky implementation described in the preceding paragraph is retained only as the archived ordering-dependent diagnostic against which the ordering sensitivity is measured. The 18.66-point spread quoted above is the single figure taken from it, and it is reported to justify the choice of estimator rather than as a finding about the market; the Cholesky path supplies no connectedness estimate that this paper reports as a result. The generalized quantities were produced by a separate archived run of the same experiment, and the source annotations on the connectedness results in §§5.3 and 6.1 name that run’s results file rather than the one behind the remaining tables.

4.5 Out-of-sample forecasting framework

We test whether the in-sample asymmetry and regime findings translate into out-of-sample predictive power for VIX. The base forecast is an autoregressive $\text{AR}(p)$ model whose lag length is selected by AIC over $p = 1, \dots, 22$ on the in-sample window (2015-02 to 2018-12), with every candidate fitted on a common sample (22 held-back observations, 942 fitted observations) so

that AIC compares like with like; this selects $p = 3$. The augmented forecast adds lagged BTC realized volatility:

$$\text{VIX}_t = \alpha_0 + \sum_{j=1}^p \alpha_j \text{VIX}_{t-j} + \gamma \cdot \text{RV}_{\text{btc}, t-1}^{(20)} + u_t. \quad (6)$$

Both models are re-estimated on a rolling 756-day window, and one-day-ahead point forecasts are generated for each day in the OOS window 2019-01-01 to 2026-04-08 (1,826 observations). Forecast errors are evaluated under squared-error loss. The loss differential is tested with the Diebold and Mariano (1995) statistic using HAC standard errors with bandwidth $\max(h - 1, \lceil h^{1/3} n^{1/3} \rceil)$. Because the augmented model nests the baseline, the Diebold-Mariano statistic is biased against the larger model under the null, so we also report the Clark and West (2007) adjusted statistic, which is the appropriate test for this comparison. Failure to reject the null is interpreted as a transparent signal that, despite in-sample causal structure, BTC realized volatility does not improve VIX point forecasts on average over the long OOS sample.

5 Main Results

This section presents the directional asymmetry of the BTC-to-VIX channel (§5.1), the test for tail concentration in the conditional response, which is negative once VIX persistence is controlled (§5.2), and the regime-dependence of the Granger structure (§5.3).

5.1 Asymmetric Granger causality: only BTC downside transmits

Table 2 reports the central asymmetric-Granger result. The negative branch ($\Delta \text{RV}^{\text{btc}, -}$, the first-differenced downside directional RV series) Granger-causes VIX at every tested lag from 1 through 5, with F -statistics ranging from 8.27 to 21.78 and p -values uniformly below 10^{-5} . The positive branch ($\Delta \text{RV}^{\text{btc}, +}$) is not significant at any lag, with F -statistics between 0.17 and 2.04 and p -values between 0.15 and 0.95. The asymmetry is monolithic: there is no overlap region in which both branches contribute — the channel is fully one-sided.

Table 2: Asymmetric Granger causality: BTC realized-volatility innovations to VIX, lags 1–5.

Lag	BTC ⁺ → VIX		BTC [−] → VIX	
	F	p	F	p
1	2.04	0.153	21.78	3.2×10^{-6}
2	0.21	0.812	19.57	$< 10^{-8}$
3	0.20	0.896	13.11	1.7×10^{-8}
4	0.17	0.951	9.49	1.3×10^{-7}
5	0.29	0.919	8.27	$< 10^{-6}$

F -tests of the joint null that the lagged $\text{BTC}^{\pm}$ coefficients are zero in Eq. (2). The negative-branch test rejects the null at every lag with $p < 10^{-5}$; the positive branch fails to reject at any lag. Source: `experiments/k1025/k1025_v2_results.json .asymmetric_granger`.

The asymmetry is corroborated by symmetric Granger tests over a broader lag window. Pooling the positive and negative innovations into a single specification, the symmetric $\text{BTC RV} \rightarrow \text{VIX}$ test rejects the joint null at lags 2 through 10 with p -values from 4.9×10^{-3} (lag 2) down to below 10^{-6} (lags 3–10); the lag-1 specification is non-significant ($F = 0.59$, $p = 0.44$).

Both directions of the symmetric test. The symmetric specification above reports only the BTC-to-VIX direction. Table 3 prints both directions of the same test over the full lag

1–10 window, estimated on the same sample and the same artifact. The two directions are not symmetric in strength or in lag profile. In the BTC RV $\rightarrow$ VIX direction the F -statistic is non-significant at lag 1 ($F = 0.59$, $p = 0.444$), rises to a peak of $F = 14.05$ at lag 3, and remains significant through lag 10. In the VIX $\rightarrow$ BTC RV direction the null is rejected at *every* lag from 1 to 10, including the lag at which the forward direction fails: the reverse lag-1 test gives $F = 7.77$ ($p = 0.00535$) against the forward $F = 0.59$ ($p = 0.444$). The reverse-direction F -statistics are, however, uniformly smaller than the forward ones from lag 2 onward (3.81–5.13 versus 5.33–14.05), so the evidence is one of a dominant BTC-to-VIX channel operating alongside a persistent, weaker equity-to-BTC feedback — not of a one-way channel.

Table 3: Bidirectional symmetric Granger causality between BTC realized volatility and VIX, lags 1–10.

Lag	BTC RV $\rightarrow$ VIX		VIX $\rightarrow$ BTC RV	
	F	p	F	p
1	0.59	0.444	7.77	0.00535
2	5.33	0.0049	4.05	0.01745
3	14.05	4.3×10^{-9}	4.18	0.0058
4	10.26	3.0×10^{-8}	5.13	0.00040
5	8.02	1.6×10^{-7}	4.74	2.6×10^{-4}
6	7.28	9.9×10^{-8}	4.87	5.9×10^{-5}
7	7.61	4.0×10^{-9}	4.61	3.9×10^{-5}
8	7.02	3.2×10^{-9}	4.35	3.1×10^{-5}
9	6.10	1.6×10^{-8}	4.18	2.2×10^{-5}
10	5.56	3.0×10^{-8}	3.81	3.9×10^{-5}

F -tests of the joint null that the lagged coefficients of the conditioning series are zero, estimated separately in each direction on the same sample ($N = 2,812$). The forward direction (BTC RV $\rightarrow$ VIX) fails to reject at lag 1 and rejects at lags 2–10; the reverse direction (VIX $\rightarrow$ BTC RV) rejects at every lag 1–10 but with uniformly smaller F -statistics from lag 2 onward. Source: `experiments/k1025/k1025_v2_results.json`
`.granger_causality.btc_rv_to_vix` and `.granger_causality.vix_to_btc_rv`.

Asymmetric vs. symmetric Granger: reading the lag-1 result. The lag-1 cell is the one place where all three specifications disagree, and it has to be read with all three in view. The *forward* symmetric specification at lag 1 averages over the (small, possibly zero) upside coefficient and the (large, positive) downside coefficient, producing the pooled $F = 0.59$ ($p = 0.444$) non-rejection. Once the two branches are estimated separately as in Eq. (2), the downside-branch test recovers $F = 21.78$ ($p < 10^{-5}$): the lag-1 BTC-to-VIX channel is real but mechanically invisible to a sign-pooled symmetric test. Sign-pooling is therefore *a* reason the forward lag-1 test fails to reject, but it is not the only thing happening at that lag. The *reverse* symmetric test at the same lag, on the same sample and the same artifact, does reject, and by a wide margin: VIX $\rightarrow$ BTC RV gives $F = 7.77$ ($p = 0.00535$), an order of magnitude larger in F than the forward $F = 0.59$. At a one-day horizon, then, the dominant measurable linear relation between the two series runs from equity fear into Bitcoin volatility, not the other way, and the forward-direction signal at that horizon exists only inside the downside branch. We flag this explicitly because the reverse direction is not a nuisance result we can set aside: it rejects at every lag from 1 to 10 (Table 3). What the lag structure supports is a *predominant* BTC-to-VIX channel — forward F -statistics exceed reverse ones from lag 2 onward, peaking at $F = 14.05$ against a reverse maximum of $F = 5.13$ over the same lags — operating alongside a weaker but persistent feedback from equity fear to Bitcoin volatility, with the feedback leg the stronger of the two at lag 1. Two caveats bound how far this can be pushed. First, the asymmetric decom-

position has been estimated in the forward direction only, so we cannot say whether the reverse leg is itself sign-asymmetric; that is left to future work. Second, the bidirectional pattern is what a Granger test on two contemporaneously correlated volatility series will show whenever both carry common shocks, so it identifies the *direction of incremental linear predictability*, not an economic mechanism. The practical implication for the rest of the paper is narrower than a one-way reading would license: asymmetric and symmetric Granger are not interchangeable diagnostics, and neither licenses describing this channel as unidirectional.

5.2 Tail dependence: no amplification once VIX persistence is controlled

Table 4 reports the quantile-regression slope of VIX on lagged BTC realized variance at five quantiles, first without and then with lagged VIX as a control. Without the control (Panel A) the slope is significantly negative at $\tau = 0.05$ (-2.81) and significantly positive at $\tau = 0.75$ and 0.95 ($+8.23$ and $+16.29$), which read in isolation looks like a sign reversal with strong upper-tail amplification. Adding VIX_{t-1} (Panel B) removes that pattern. The upper-tail slope falls to $+0.42$ with an interval of $[-0.61, +1.14]$, the $\tau = 0.75$ slope is -0.02 with an interval of $[-0.21, +0.23]$, and the three lower and middle quantiles carry small negative slopes (-0.33 , -0.20 , -0.18) whose intervals exclude zero.

Table 4: Quantile regression of VIX on lagged BTC realized variance, without and with lagged VIX, $N = 2,811$.

τ	0.05	0.25	0.50	0.75	0.95
<i>Panel A: no control</i>					
β_τ	-2.81	-1.51	$+2.59$	$+8.23$	$+16.29$
95% CI	$[-3.85, -1.93]$	$[-3.75, +2.08]$	$[-0.80, +7.43]$	$[+1.76, +14.89]$	$[+0.76, +24.69]$
<i>Panel B: with VIX_{t-1} (quantile-Granger)</i>					
β_τ	-0.33	-0.20	-0.18	-0.02	$+0.42$
SE (block boot)	0.162	0.076	0.059	0.121	0.496
t -stat	-2.04	-2.61	-3.00	-0.16	0.84
95% CI	$[-0.63, -0.02]$	$[-0.33, -0.03]$	$[-0.31, -0.07]$	$[-0.21, +0.23]$	$[-0.61, +1.14]$

Slope β_τ on $RV_{\text{btc}, t-1}^{(20)}$ in Eq. (3). Intervals are percentile intervals from a moving-block bootstrap (1,000 replications, block length 15, seed 42); non-converged draws are excluded, leaving at least 997 successful draws in every cell. Panel B is the specification that measures incremental tail information from Bitcoin. Source: `experiments/k1025/k1025_v3_results.json .quantile_regression`.

The comparison between the panels is the finding. The coefficients in Panel A are large because lagged BTC realized variance is correlated with lagged VIX, and VIX is persistent; once VIX’s own lag is in the regression, Bitcoin’s slope shrinks by a factor of roughly 8 at $\tau = 0.05$ and by far more in the upper quantiles. Nothing in Panel B supports tail amplification: the only quantiles with intervals excluding zero are $\tau \leq 0.5$, where the slope is small and negative, and the upper-tail slope is statistically indistinguishable from zero. The “sign switch” between Panels A and B at $\tau \geq 0.5$ is therefore an omitted-variable effect, not a state-dependent transmission mechanism.

Withdrawn: the earlier tail-amplification claim. Earlier drafts of this paper made this quantile pattern a headline finding, reporting the unconditional slope as -2.78 at $\tau = 0.05$ rising to $+19.85$ at $\tau = 0.95$ ($t = -13.84$ and 8.17), a “ $7.0\times$ upper-tail amplification” ($\beta_{0.95}/\beta_{0.5} \approx 7.04$) and a sign reversal “statistically significant in both halves”. Those figures are withdrawn. They came from the specification without lagged VIX, estimated with an i.i.d. pairs bootstrap

that ignores the strong serial dependence of VIX and therefore understated the uncertainty; they also came from an earlier panel construction, which this version replaces with the pinned snapshot, log returns throughout, and Bitcoin aligned to equity trading days; we do not attribute the gap between the earlier unconditional estimates and Panel A to any single one of these changes. The corrected unconditional estimates are those in Panel A, and the claim of tail amplification does not survive Panel B.

5.3 Regime dependence: COVID-2020 as structural watershed

We split the 2015–2026 sample into five non-overlapping subperiods — pre-mania (2015–2017), crypto-winter (2018–2019), COVID (2020), bull-bear (2021–2022), and recovery-plus-ETF (2023–2026) — and re-estimate the BTC-to-VIX Granger test within each subperiod with AIC-selected lag length. Table 5 reports the result. Granger causality survives the Bonferroni multiple-testing correction ($p < 0.01/5$) only during 2020 ($F = 12.31$, $p = 1.6 \times 10^{-7}$, lag 3); among the other four subperiods, 2015–2017 is marginally significant at the uncorrected 5% level ($F = 2.88$, $p = 0.014$, Table 5) but does not pass the corrected threshold, and the remaining three fail to reject at any conventional level.

Table 5: Subperiod Granger tests of BTC realized volatility on VIX, AIC-best lag.

Subperiod	N	Best lag	F	p
2015–2017 (Pre-mania)	735	5	2.88	0.014
2018–2019 (Crypto winter)	503	1	0.37	0.544
2020 (COVID)	253	3	12.31	1.6×10^{-7}
2021–2022 (Bull-Bear)	503	1	2.03	0.155
2023–2026 (Recovery+ETF)	818	2	0.16	0.850

Granger F -tests within each subperiod with AIC-selected lag length. Only 2020 rejects the non-causality null (Bonferroni threshold also cleared). Source: `experiments/k1025/k1025_v2_results.json`
`.subperiod_granger`.

The EWMA dynamic correlation (RiskMetrics, $\lambda = 0.94$) by VIX regime corroborates the regime-dependence finding from a different angle. Sorting trading days into four VIX regimes (Low: $VIX \leq 15$; Normal: $15 < VIX \leq 25$; High: $25 < VIX \leq 35$; Crisis: $VIX > 35$), the mean BTC-SPY EWMA dynamic correlation (RiskMetrics, $\lambda = 0.94$) rises near-monotonically from 0.07 in the Low regime to 0.27 in Normal, peaking at 0.45 in High, and remaining elevated at 0.41 in Crisis (Table 6; the slight Crisis dip from 0.45 to 0.41 reflects the small Crisis-regime sample of $n = 63$). The percentage of days with positive correlation rises from 62% in Low to 96% in High and 98% in Crisis. Bitcoin’s purported diversification value relative to equities is therefore essentially absent precisely when investors most need it.

Table 6: BTC-SPY EWMA dynamic correlation (RiskMetrics, $\lambda = 0.94$) by VIX regime, $N = 2,812$ days total.

VIX regime	N	Mean	Median	Std. dev.	% positive
Low ($VIX \leq 15$)	1,037	0.067	0.053	0.190	62.3
Normal (15–25)	1,386	0.266	0.302	0.258	82.8
High (25–35)	326	0.454	0.484	0.204	95.7
Crisis ($VIX > 35$)	63	0.407	0.484	0.156	98.4

EWMA (RiskMetrics, $\lambda = 0.94$) BTC-SPY dynamic correlation aggregated within each VIX regime (audit fix 2026-06-10: regime cutpoints and correlation-estimator label aligned to k1025_v2.py implementation; an earlier draft labelled this DCC and used 22/30 cutpoints). The High and Crisis regime correlations contradict the early “digital gold” thesis and are consistent with the crisis concentration of the Granger evidence in Tables 2 and 5. Source: `experiments/k1025/k1025_v2_results.json .dcc_correlation_by_regime`.

Finally, the connectedness analysis quantifies directionality. We use the generalized (KPPS) forecast-error variance decomposition of Koop et al. (1996) and Pesaran and Shin (1998) rather than a Cholesky decomposition, because the Cholesky net shares are not invariant to the ordering of the variables: across the six orderings of $\{VIX, RV_{\text{spy}}^{(20)}, RV_{\text{btc}}^{(20)}\}$, BTC’s Cholesky net share spans 18.66 percentage points and changes sign (from +10.55 to -8.11), so no directional claim can be carried by it. Under the order-invariant generalized decomposition, the full-sample total connectedness index is 19.52% and BTC’s net connectedness is -0.95 percentage points. BTC is therefore, at most, a marginal net receiver: its net position is within one percentage point of zero, an order of magnitude smaller than the equity-fear channel itself (VIX net +25.9 pp, SPY net -24.9 pp). The directional evidence is too weak to carry the “fear originator vs. fear amplifier” reframing on its own; that reframing rests on the Granger results alone, and the connectedness decomposition is reported here as a descriptive complement rather than as a supporting pillar.

6 Robustness

This section reports the one robustness exercise that is currently archived in a traceable way for the active manuscript: the time-series stability of the rolling spillover index. Other robustness ideas considered in earlier drafts remain prospective until they are written back into the experiment outputs.

6.1 Spillover-index stability across rolling windows

The generalized total connectedness index is computed on each of 512 rolling 252-trading-day windows. The mean is 22.75% with a standard deviation of 8.79 percentage points, ranging from 6.29% to 47.02%. The index is therefore *not* stable: it moves by a factor of seven across the sample, rising sharply in the COVID-2020 window and subsiding in the calm 2017–2019 period. This is the opposite of what earlier drafts of this paper reported, and the reversal is instructive. Those drafts described the now-withdrawn index as “remarkably stable” at $90.11\% \pm 0.22$ pp — a near-constant, near-total connectedness that would have been surprising for any three-asset system. That constancy was an artifact of a mis-sliced variance-decomposition array, which drove the index mechanically toward 90% on any input, including pure noise (§6.2). The corrected index behaves as connectedness measures normally do: it is regime-sensitive, and it co-moves with the Granger evidence in §5.3 rather than standing apart from it.

The directional decomposition is likewise not stable. Across the 512 rolling windows BTC’s generalized net connectedness is negative in 72.5% of windows — so BTC is a net receiver

more often than not, but in more than one window in four it is a net transmitter. The claim in earlier drafts that “BTC remains a net receiver in every window” does not survive the corrected decomposition and is withdrawn. What survives is weaker and more honest: BTC’s net position leans negative but is neither uniform across windows nor economically large in the full sample (−0.95 pp).

6.2 Erratum: the withdrawn spillover numbers

Earlier drafts of this paper reported a now-withdrawn total spillover index of 90.11% and a withdrawn BTC net spillover of −74.4 percentage points, and read the latter as strong evidence that Bitcoin is a net receiver of volatility. Both numbers are withdrawn. They were produced by a defect in the variance-decomposition code, which indexed the `statsmodels` FEVD array as `(horizon, n, n)` when its true layout is `(n, horizon, n)`. The resulting slice returned the decomposition table of the *last variable* rather than the *last horizon*, so the number of variables was read as the forecast horizon (10) instead of 3. Two consequences follow. The total index is driven mechanically toward 90% for any input — we verified that it does so on simulated white noise, where the true connectedness is zero — which explains the implausible stability the earlier drafts celebrated. And the net share becomes a subtraction across non-commensurable dimensions, which is what manufactured the −74.4 pp figure.

The defect was reproduced bit-for-bit on the same data before it was corrected, so the before/after comparison in this section is measured rather than asserted: the buggy path returns −76.6 pp and 90.09% on the pinned snapshot, matching the published values. The corrected, order-invariant estimates are those reported in §6.1. The direction of the error matters for how the reader should weigh the remaining evidence: the connectedness pillar of this paper was not merely imprecise, it was reporting an artifact. Together with the withdrawn tail-amplification claim (§5.2), this leaves the paper’s positive contribution resting on the Granger evidence alone.

Deferred robustness work. Two additional exercises considered in earlier drafts are not carried as active evidence here: a uniform-lag re-estimation of the subperiod Granger tests and a dedicated pre-/post-ETF split. Neither is currently written back into the experiment JSON that underlies this manuscript. A third exercise, the cross-asset VXN/QQQ replication, has now been re-estimated under the order-invariant decomposition, and it does not replicate: on the QQQ/BTC/VXN panel BTC’s net connectedness is +2.70 pp, which does not even agree in *sign* with the −0.95 pp obtained on the SPY/VIX panel. The agreement that earlier drafts reported between the two panels (−76.64 pp vs. −76.89 pp, described there as “near identical”) was agreement between two scripts that shared the same mis-slicing defect, not corroboration of an effect. That cross-panel replication claim is withdrawn.

7 Out-of-Sample Forecasting Evaluation

The in-sample results in §5 establish that the BTC-to-VIX Granger channel is asymmetric and concentrated in the COVID-2020 subperiod, and that it carries no tail amplification once VIX persistence is controlled. The natural next question is whether the Granger structure translates into out-of-sample forecasting power. We test the augmented specification of Eq. (6) against the AR(3) baseline over the 1,826-day OOS window 2019-01-01 to 2026-04-08.

Over the full OOS window the augmented model’s MSE is 4.527 against 4.512 for the baseline, a 0.32% *deterioration*. The Diebold-Mariano statistic is $t = -0.99$ ($p = 0.32$, HAC bandwidth 13), and the Clark-West statistic for the nested comparison is $t = -0.12$, far below its one-sided 5% critical value of 1.645. The loss differential has negative first-order autocorrelation (−0.135), which shrinks the HAC standard error as the bandwidth grows, so $|t|$ rises with the bandwidth rather than falling: 0.72 at bandwidth 0, 0.99 at 13, and 1.33 at 30. The conclusion

does not depend on that choice. We return to the reconciliation of this OOS null with the in-sample Granger evidence in §8.2.

Table 7: Out-of-sample forecast evaluation: AR(3) baseline vs. AR(3)+BTC RV, by subperiod of the OOS window 2019-01-01 to 2026-04-08.

Period	N	ΔMSE (%)	DM t	DM p	CW t
Full OOS	1,826	−0.32	−0.99	0.32	−0.12
2019	252	−0.02	−0.20	0.84	−0.11
2020 (COVID)	253	−0.07	−0.09	0.92	0.28
2021–2022	503	−0.49	−0.82	0.41	0.21
2023–2026	818	−0.53	−2.75	0.006	−2.55

ΔMSE is the percentage improvement of the augmented model relative to the baseline; negative values indicate deterioration. DM is the Diebold and Mariano (1995) statistic with HAC bandwidth $\max(h - 1, \lceil h^{1/3} n^{1/3} \rceil)$ (13, 7, 7, 8, 10 by row). CW is the Clark and West (2007) statistic for nested models; a value above 1.645 would indicate that the augmented model adds predictive content at the one-sided 5% level. Source: `experiments/k1025/k1025_v3_results.json .oos_forecast`.

Subperiods. A natural concern is that the pooled test obscures predictability inside the crisis window that the Granger results highlight. Table 7 splits the OOS window into four subperiods. The augmented model does not improve on the baseline in any of them: its MSE is higher in every subperiod, and no Clark-West statistic comes close to 1.645, including the 2020 COVID window ($t = 0.28$). The only subperiod with a large statistic is 2023–2026, and there the sign runs the other way: adding Bitcoin realized volatility makes the forecasts significantly *worse* (DM $t = -2.75$, $p = 0.006$). An earlier version of this section reported a single pooled DM statistic of $t = -1.14$ from an AR(10) model whose lag length had been selected by AIC with each candidate fitted on a different sample, which mechanically favoured the longest lag in the grid (10 was its upper bound), together with a VIX-level split into crisis and normal days; those figures are withdrawn and replaced by Table 7. The qualitative conclusion, no improvement, is unchanged.

Interpretation. The forecasting null holds for the pooled window and for every subperiod. We interpret the result as substantive rather than as a power deficiency. Conrad and Kleen (2020)’s methodological argument is directly relevant: an in-sample signal that lives in downside innovations and only during 1 of 5 subperiods is exactly the kind of sparse signal that fails to outperform a well-specified autoregression on average over a long OOS sample. Reporting (i) significant asymmetric Granger with F -statistics up to 21.78, (ii) regime concentration in 2020, (iii) the disappearance of the apparent tail pattern under a control for VIX persistence, and (iv) a failed OOS comparison under both Diebold-Mariano and Clark-West together gives a cleaner template for empirical claims about cross-asset predictors than the “in-sample-only” or “OOS-only” modes that dominate the existing crypto-equity spillover literature.

8 Discussion

This section interprets the empirical findings, addresses the Granger-versus-forecastability tension, and draws policy implications.

8.1 Asymmetry mechanism: retail-driven downside transmission

The asymmetric Granger result (Table 2) shows that BTC downside volatility transmits to VIX while BTC upside volatility does not. This pattern is unlikely to reflect a symmetric

leverage effect within Bitcoin’s own price process, because the leverage effect would generate the asymmetry within the BTC-VIX joint distribution but would not, on its own, prevent a positive BTC volatility shock from spilling over to expected equity fear. One candidate interpretation is behavioral and market-structure based: Bitcoin trading is materially more retail-driven than U.S. large-cap equity trading, and downside moves may carry a more salient risk signal than upside moves. Under that interpretation, negative BTC shocks could co-move with broader risk-off positioning more strongly than positive shocks do. We do not structurally identify this mechanism here, so the mechanism language should be read as suggestive rather than as established evidence of margin-cascade contagion.

8.2 Granger causality $\neq$ forecastability: a methodological lesson

The juxtaposition of the highly significant in-sample asymmetric Granger result (F up to 21.78, $p < 10^{-5}$) with the failed out-of-sample comparison (DM $t = -0.99$, Clark-West $t = -0.12$) is, on its surface, paradoxical. Two reconciliations are consistent with our evidence. First, the in-sample signal is concentrated in one thin slice of the joint distribution: the negative branch of BTC innovations, with no contribution from the positive branch (Table 2). A signal carried by that slice can produce a strong F -statistic over the sample without improving point forecasts of VIX on average over the OOS window, because most BTC innovations are not negative. The quantile evidence does not add a second slice: once VIX persistence is controlled there is no upper-tail concentration for the forecast to exploit (Table 4). Second, the regime-conditional finding (§5.3) localizes the signal to one of five subperiods (2020); even if the predictive content within 2020 is real, it is too brief to dominate the point-forecast loss averaged across 1,826 OOS days. The methodological lesson is that the Granger test and the DM test are answering different questions: “does X predict Y at *any* horizon, conditional on full information about the joint distribution?” versus “does augmenting an autoregressive forecast of Y with X reduce expected loss *on average* across an OOS sample?” Our evidence supports the former but not the latter.

8.3 Implications for prudential policy and margin design

The crisis-conditional spillover may have policy relevance, but the present evidence supports only a narrower statement than some earlier drafts suggested. Where systemic-risk frameworks for positive-feedback equity strategies focus on *intra-equity* concentration of correlated trading flow during stress, the present paper is consistent with an *external-market amplification* channel in which cryptocurrency volatility co-moves more strongly with equity fear during crisis than during calm periods. That is enough to suggest that supervisors and risk managers should not assume complete decoupling between crypto and equity stress states. It is *not* enough, by itself, to justify a numeric stress-test threshold, to infer causal margin-cascade mechanics, or to claim measurable retail-protection externalities. We therefore read the policy implication as conditional and qualitative: crisis-state crypto volatility belongs on the monitoring dashboard, while stronger structural claims require richer identification.

8.4 Limitations and future work

Three limitations qualify our findings. First, the COVID-2020 subperiod is the only window in which Granger causality clears conventional thresholds, and that subperiod includes both the March-2020 panic and the subsequent monetary-policy regime shift; disentangling pure-fear contagion from policy-induced volatility transmission is beyond our reduced-form framework and would require structural identification through high-frequency event studies. Second, our evidence is at the daily frequency; intraday transmission — where liquidation cascades might plausibly originate — is invisible to the daily Granger and quantile tests we deploy. Third,

the evidence comes from two archived runs of the same experiment built on different panel constructions. The Granger, subperiod, and EWMA-correlation results come from the earlier run; the connectedness, quantile-regression, and out-of-sample results come from the rebuilt run on the pinned snapshot, whose panel uses log returns throughout and aligns Bitcoin to equity trading days. The Granger tests have not yet been re-estimated on the rebuilt panel. Because the unconditional quantile slopes moved materially between the two constructions, that re-estimation is the main prerequisite before the Granger magnitudes are cited as final.

9 Conclusion

Using daily data on SPY, BTC-USD, and VIX from February 2015 through April 2026 ($N = 2,812$), we document two stylized facts about the crypto-equity volatility spillover and one negative result. The channel itself is bidirectional — symmetric Granger tests reject non-causality from VIX to BTC realized volatility at every lag from 1 to 10, and at lag 1 that reverse leg ($F = 7.77$) is the stronger of the two ($F = 0.59$ forward) — and the findings below describe the BTC-to-VIX leg, which dominates from lag 2 onward. First, the spillover is *asymmetric*: Bitcoin downside volatility Granger-causes VIX at every tested lag from 1 through 5 (F between 8.27 and 21.78, p uniformly below 10^{-5}), while the upside branch is non-significant at all lags. Second, the spillover is *regime-dependent*: of five non-overlapping subperiods spanning 2015 through 2026, only the 2020 COVID window rejects non-causality at conventional levels ($F = 12.31$, $p < 10^{-6}$). The negative result is that the spillover is *not* tail-concentrated once the persistence of VIX is controlled: the upper-tail quantile slope of VIX on lagged BTC realized variance is +0.42 with a 95% block-bootstrap interval of $[-0.61, +1.14]$, and the steep unconditional tail pattern is an omitted-variable effect. An order-invariant connectedness decomposition places Bitcoin only marginally on the receiving side (net -0.95 percentage points), which is too weak to sustain a directional claim by itself; the reframing from “crypto as fear originator” to “crypto as fear amplifier conditional on pre-existing equity stress” therefore rests on the Granger evidence alone.

The corresponding out-of-sample test, however, returns a transparent null: augmenting an AR baseline of VIX with lagged BTC realized volatility does not reduce expected forecast loss over the 1,826-day OOS window 2019–2026 (DM $t = -0.99$, $p = 0.32$; Clark-West $t = -0.12$), nor in any of four OOS subperiods. We interpret this null as evidence that the in-sample signal is too sparse — localized in the BTC negative branch and the COVID subperiod — to dominate point-forecast loss averaged over the long OOS sample, and we treat the joint reporting of in-sample structure with out-of-sample discipline as a contribution in itself, illustrating that Granger causality and forecastability answer related but distinct questions.

The findings reframe the crypto-equity volatility link in three ways. Methodologically, asymmetric estimators recover structure that pooled symmetric tools obscure, while unconditional quantile regressions of a persistent fear index can manufacture tail heterogeneity that a control for the index’s own lag removes. Empirically, the COVID-2020 watershed implies that crypto-equity transmission is a crisis-time phenomenon rather than a steady-state channel, with dormant periods during which the two markets behave as approximately independent. Substantively, the paper is strongest as a reduced-form characterization of when the BTC–VIX link appears and when it vanishes; stronger claims about structural mechanism, cross-asset generality, or operational policy calibration remain future work. We see three natural extensions: high-frequency replication to identify the within-day transmission mechanism, a re-estimation of the Granger tests on the rebuilt panel, and a cross-asset rebuild under a methodologically symmetric specification. The asymmetric, regime-dependent crypto fear channel is, on the evidence, real — but real in a sense that requires conditioning on crisis and on direction to be visible at all, and it carries no tail amplification once the persistence of equity fear is accounted for.

References

- Adrian, T. and Brunnermeier, M.K. (2016). CoVaR. *American Economic Review*, 106(7):1705–1741. <https://doi.org/10.1257/aer.20120555>
- Akyildirim, E., Corbet, S., Lucey, B., Sensoy, A., and Yarovaya, L. (2020). The relationship between implied volatility and cryptocurrency returns. *Finance Research Letters*, 33:101212. <https://doi.org/10.1016/j.frl.2019.06.010>
- Andrews, D.W.K. (1991). Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica*, 59(3):817–858. <https://doi.org/10.2307/2938229>
- Bouri, E., Shahzad, S.J.H., Roubaud, D., Kristoufek, L., and Lucey, B. (2020). Bitcoin, gold, and commodities as safe havens for stocks: New insight through wavelet analysis. *The Quarterly Review of Economics and Finance*, 77:156–164. <https://doi.org/10.1016/j.qref.2020.03.004>
- Clark, T.E. and West, K.D. (2007). Approximately normal tests for equal predictive accuracy in nested models. *Journal of Econometrics*, 138(1):291–311. <https://doi.org/10.1016/j.jeconom.2006.05.023>
- Conlon, T. and McGee, R. (2020). Safe haven or risky hazard? Bitcoin during the Covid-19 bear market. *Finance Research Letters*, 35:101607. <https://doi.org/10.1016/j.frl.2020.101607>
- Conrad, C. and Kleen, O. (2020). Two are better than one: Volatility forecasting using multiplicative component GARCH-MIDAS models. *Journal of Applied Econometrics*, 35(1):19–45. <https://doi.org/10.1002/jae.2742>
- Corbet, S., Meegan, A., Larkin, C., Lucey, B., and Yarovaya, L. (2018). Exploring the dynamic relationships between cryptocurrencies and other financial assets. *Economics Letters*, 165:28–34. <https://doi.org/10.1016/j.econlet.2018.01.004>
- Diebold, F.X. and Mariano, R.S. (1995). Comparing predictive accuracy. *Journal of Business & Economic Statistics*, 13(3):253–263. <https://doi.org/10.1080/07350015.1995.10524599>
- Diebold, F.X. and Yilmaz, K. (2009). Measuring financial asset return and volatility spillovers, with application to global equity markets. *The Economic Journal*, 119(534):158–171. <https://doi.org/10.1111/j.1468-0297.2008.02208.x>
- Diebold, F.X. and Yilmaz, K. (2012). Better to give than to receive: Predictive directional measurement of volatility spillovers. *International Journal of Forecasting*, 28(1):57–66. <https://doi.org/10.1016/j.ijforecast.2011.02.006>
- Diebold, F.X. and Yilmaz, K. (2014). On the network topology of variance decompositions: Measuring the connectedness of financial firms. *Journal of Econometrics*, 182(1):119–134. <https://doi.org/10.1016/j.jeconom.2014.04.012>
- Diks, C. and Panchenko, V. (2006). A new statistic and practical guidelines for nonparametric Granger causality testing. *Journal of Economic Dynamics and Control*, 30(9–10):1647–1669. <https://doi.org/10.1016/j.jedc.2005.08.008>
- Harvey, D., Leybourne, S., and Newbold, P. (1997). Testing the equality of prediction mean squared errors. *International Journal of Forecasting*, 13(2):281–291. [https://doi.org/10.1016/S0169-2070\(96\)00719-4](https://doi.org/10.1016/S0169-2070(96)00719-4)

- Harvey, C.R., Liu, Y., and Zhu, H. (2016). ...and the cross-section of expected returns. *Review of Financial Studies*, 29(1):5–68. <https://doi.org/10.1093/rfs/hhv059>
- Hatemi-J, A. (2012). Asymmetric causality tests with an application. *Empirical Economics*, 43(1):447–456. <https://doi.org/10.1007/s00181-011-0484-x>
- Hong, Y. (2001). A test for volatility spillover with application to exchange rates. *Journal of Econometrics*, 103(1–2):183–224. [https://doi.org/10.1016/S0304-4076\(01\)00043-4](https://doi.org/10.1016/S0304-4076(01)00043-4)
- Iyer, T. (2022). Cryptic connections: Spillovers between crypto and equity markets. *IMF Global Financial Stability Notes*, 2022/01. <https://doi.org/10.5089/9781616358068.065>
- Klein, T., Pham Thu, H., and Walther, T. (2018). Bitcoin is not the new gold — A comparison of volatility, correlation, and portfolio performance. *International Review of Financial Analysis*, 59:105–116. <https://doi.org/10.1016/j.irfa.2018.07.010>
- Koenker, R. and Bassett, G. (1978). Regression quantiles. *Econometrica*, 46(1):33–50. <https://doi.org/10.2307/1913643>
- Koop, G., Pesaran, M.H., and Potter, S.M. (1996). Impulse response analysis in nonlinear multivariate models. *Journal of Econometrics*, 74(1):119–147. [https://doi.org/10.1016/0304-4076\(95\)01753-4](https://doi.org/10.1016/0304-4076(95)01753-4)
- Matkovskyy, R. and Jalan, A. (2019). From financial markets to Bitcoin markets: A fresh look at the contagion effect. *Finance Research Letters*, 31:93–97. <https://doi.org/10.1016/j.frl.2019.04.007>
- Pesaran, M.H. and Shin, Y. (1998). Generalized impulse response analysis in linear multivariate models. *Economics Letters*, 58(1):17–29. [https://doi.org/10.1016/S0165-1765\(97\)00214-0](https://doi.org/10.1016/S0165-1765(97)00214-0)
- Shahzad, S.J.H., Bouri, E., Roubaud, D., Kristoufek, L., and Lucey, B. (2019). Is Bitcoin a better safe-haven investment than gold and commodities? *International Review of Financial Analysis*, 63:322–330. <https://doi.org/10.1016/j.irfa.2019.01.002>
- Yarovaya, L., Brzeszczyński, J., Goodell, J.W., Lucey, B., and Lau, C.K.M. (2022). Rethinking financial contagion: Information transmission mechanism during the COVID-19 pandemic. *Journal of International Financial Markets, Institutions and Money*, 79:101589. <https://doi.org/10.1016/j.intfin.2022.101589>