

The Scale–Shape Divide in Bitcoin and Ether Perpetuals: Roughness States, Quarticity, and Bipower Residuals

Fuli Yang

GARCH Institute, Harvard Square, Cambridge, MA 02138, USA

3 September 2026

Abstract

This paper asks whether rolling roughness estimates that label volatility states also rank the shape of near-term tail variation. The design freezes the January 2020–July 2026 Binance BTCUSDT and ETHUSDT tape, 9,560 six-hour origins per contract, and the three functionals used in companion work. Forward 8-hour and 24-hour outcomes are reconstructed from one-minute sufficient statistics. At eight hours, a one-standard-deviation increase in $\hat{H}^{IV}$ is associated with 0.186 log points more realized quarticity in Bitcoin and 0.166 in Ether; the corresponding $\hat{H}^{PA}$ effects are 0.177 and 0.157. Origin-level 14-day block-cluster inference preserves all four effects, whereas $\hat{H}^{PV}$ remains null. The accounting identity $\log RQ = 2 \log RV + \log NRQ$ locates the association: it loads through realized-variance scale. Log normalized quarticity, the raw bipower-residual share, and a conservative RQ-scaled significance-filtered residual all satisfy prespecified equivalence tests. Recursive raw-quarticity forecasts improve on both historical and state benchmarks in Bitcoin, but Ether does not reproduce the short-horizon gain; normalized-tail and residual-event results are still less stable. Roughness therefore reaches short-horizon fourth-moment scale in these contracts without furnishing a common tail-shape coordinate.

Keywords: rough volatility; realized quarticity; normalized quarticity; bipower residual; perpetual futures; tail forecasting

JEL codes: C22; C58; G13; G17

Plain Language Summary

Large future quarticity can mean that the whole volatility level rose, or that unusually large minute returns became more concentrated relative to that level. The two interpretations are not interchangeable. Using Binance Bitcoin and Ether perpetuals from January 2020 through July 2026, I separate them. Two rolling roughness estimates accompany higher quarticity over the next eight hours, and this result survives block-cluster inference. The same coefficients shrink to economically small values after quarticity is divided by squared realized variance. A second check based on the part of realized variance left after bipower variation leads to the same conclusion. Forecasting is more uneven: Bitcoin raw-quarticity losses fall, but Ether does not deliver the same short-horizon improvement. For these contracts, roughness is informative about the scale of near-term variation. It is not a stable ranking of tail shape, discontinuous variation, or trading opportunity.

1. Introduction

Rough-volatility research made the short-scale behavior of volatility an empirical object (Gatheral et al., 2018; Bennedsen et al., 2022). It also left an awkward empirical question. Increment slopes, GMM estimators, and normalized variation can react differently to noise and aggregation (Bolko et al., 2023; Cont and Das, 2024). Even when a rolling exponent tracks a persistent market state, that fact alone says little about whether future variation is merely larger or differently shaped.

Three companion studies narrowed the state interpretation before the present exercise. Absolute funding covaries persistently with an increment-variance estimate, but adds little to transition or HAR forecasts (Yang, 2026a). Pre-averaging lowers the measured level without materially changing that funding slope; normalized p-variation breaks the association (Yang, 2026b). Splitting funding by sign then places the earlier result on positive-funding observations (Yang, 2026c). Those findings concern the state label itself. They leave open whether the label reaches future fourth moments.

I therefore change the dependent variable, not the estimator menu. The original USD-margined BTCUSDT and ETHUSDT tape is held fixed from January 2020 through July 2026. Apparent roughness is updated every six hours from the preceding 336 hours; forward outcomes use complete one-minute windows of 8 and 24 hours. Raw realized

quarticity is reported first. It is then divided by squared realized variance, separating fourth-moment scale from relative concentration. Bipower variation supplies a distinct residual measure.

The realized-measure literature gives the decomposition economic content. Realized variance and quarticity summarize integrated variation and its fourth-moment scale (Barndorff-Nielsen and Shephard, 2002; Andersen et al., 2003). Bipower variation is designed to isolate a continuous component under semimartingale conditions (Barndorff-Nielsen and Shephard, 2006). At one-minute frequency, however, noise and market discreteness remain material (Hansen and Lunde, 2006; Barndorff-Nielsen et al., 2008). I call RV minus BV a bipower residual unless it passes an explicit significance screen; it is not treated as an exchange-labelled jump.

The numbers draw a clear line. At eight hours, the IV effects on log RQ are 0.186 in Bitcoin and 0.166 in Ether; PA gives 0.177 and 0.157. PV is near zero. The positive IV/PA coefficients survive 14-day block-cluster inference, although the more conservative 112-lag Newey–West family leaves the Bitcoin adjusted p-values just above 0.05. With common samples and controls, nearly all of the RQ loading is accounted for by twice the log-RV loading. Log NRQ coefficients are economically small under prespecified equivalence bounds. Raw and significance-filtered bipower-residual shares give the same verdict.

Forecasts qualify the conclusion rather than overturning it. Every Bitcoin State+H specification lowers raw-RQ loss against both simple benchmarks; Ether worsens at eight hours and improves only in isolated 24-hour cells. The paper's contribution is thus a scale–shape boundary. A state variable may help locate the level of near-term fourth-moment variation while carrying little information about concentration after scale is removed.

2. Measurement boundary and hypotheses

For an h -hour forward window containing n_h valid one-minute returns, realized variance, realized quarticity, bipower variation, normalized quarticity, and the bipower residual share are

$$RV_{t,h} = \sum_i r_i^2, \quad RQ_{t,h} = (n_h/3) \sum_i r_i^4, \quad BV_{t,h} = (\pi/2) \sum_i |r_i| |r_{i-1}|, \\ NRQ_{t,h} = RQ_{t,h} / RV_{t,h}^2, \quad JS_{t,h} = \max(RV_{t,h} - BV_{t,h}, 0) / RV_{t,h}.$$

The normalization is an accounting decomposition. Up to numerical precision, log RQ equals twice log RV plus log NRQ in every retained window. Under independent Gaussian minute returns, expected NRQ is $n_h/(n_h+2)$, close to one at both horizons; with generic minute-return kurtosis κ it is approximately $\kappa/3$ for large n_h , absent dependence and intrawindow scale variation. The observed statistic should therefore be read as relative fourth-moment concentration, not as a structural kurtosis parameter. The non-negative RV-minus-BV share is called a bipower residual below.

H1 (short-horizon scale). At eight hours, $\hat{H}^{IV}$ and $\hat{H}^{PA}$ have positive log-RQ coefficients in both contracts. Confirmation requires Holm-adjusted $p < 0.05$ in the 12-cell RQ family; $\hat{H}^{PV}$ is retained as a functional-boundary test.

H2 (shape equivalence). For log NRQ, each standardized-H coefficient lies inside ± 0.20 log points under two one-sided tests, with Holm adjustment across the 12 asset-horizon-estimator cells. The bound marks the smallest association treated here as economically material.

H3 (bipower-residual equivalence). For both the raw and significance-filtered bipower-residual shares, each standardized-H coefficient lies inside ± 0.01 share units under the same familywise equivalence rule. Event regressions must also produce a corrected result to count as contrary evidence.

H4 (recursive value). State+H must lower loss relative to Historical and State on identical forecast dates, pass Holm correction within each outcome-benchmark family, and recur across both contracts and horizons. A gain against only one benchmark is insufficient.

3. Data and frozen design

The source is Binance Vision's public USD-margined perpetual-futures archive (Binance, 2026). BTCUSDT and ETHUSDT one-minute klines, mark prices, index prices, and official funding settlements cover 1 January 2020 through 31 July 2026. Consecutive one-minute bars should be 60 seconds apart; the 50–90 second rule is a timestamp-validation tolerance, not an irregular-sampling choice. A return is rejected when that validation fails, and

missing intervals are never bridged. The tape contains 57,696 hourly rows per asset. Twelve minutes are absent in each of 2020 and 2021, all at month starts; none are absent from 2022 onward.

The roughness panel is inherited without refitting choices. $\hat{H}^{IV}$ is the slope of log increment variance across the frozen mesh set. $\hat{H}^{PA}$ applies the same slope after Jacod et al. (2009) pre-averaging and noise correction. $\hat{H}^{PV}$ is the admissible normalized p-variation root following the implementation documented in Yang (2026b). Each estimate uses the preceding 336 hours and updates on an exact six-hour grid. Negative IV and PA draws remain uncensored.

There are 9,560 forecast origins per asset. Table 1 reconciles every loss. At eight hours, one final-window truncation and 23 month-start gaps leave 9,536 complete windows. At 24 hours, four final truncations and 92 gap-contaminated windows leave 9,464. Lagged controls reduce the IV/PA regression samples to 9,513 and 9,372; PV samples are smaller because an inadmissible root is left missing rather than imputed. The minute-based RV differs from the companion hourly panel in 55 month-boundary hours, where the present concatenated construction retains valid adjacent returns. Aggregate RV changes by about 0.002%, and cumulative returns agree to machine precision.

Table 1. Sample reconciliation

Asset	h	Origins	End	Month gaps	Complete	IV/PA N	PV N
BTC	8	9,560	1	23	9,536	9,513	7,969
BTC	24	9,560	4	92	9,464	9,372	7,862
ETH	8	9,560	1	23	9,536	9,513	7,720
ETH	24	9,560	4	92	9,464	9,372	7,607

Month-gap windows arise from the 2020-02 through 2021-12 month-start archive gaps. End denotes right-edge truncation after 31 July 2026.

Table 2. Minute-based tail-shape outcomes

Asset	h	Measure	N	Median	p95	p99
BTC	8	Normalized RQ	9,536	2.138	11.345	29.695
BTC	8	Bipower residual share	9,536	0.044	0.204	0.321
BTC	24	Normalized RQ	9,464	3.159	23.486	69.451
BTC	24	Bipower residual share	9,464	0.043	0.176	0.279
ETH	8	Normalized RQ	9,536	2.091	10.210	28.476
ETH	8	Bipower residual share	9,536	0.039	0.188	0.298
ETH	24	Normalized RQ	9,464	3.074	21.935	68.769
ETH	24	Bipower residual share	9,464	0.040	0.160	0.271

RQ is reported in the regression tables in logs because its raw distribution is extremely right-skewed. The bipower residual share is $\max(RV - BV, 0)/RV$ and is bounded below at zero.

4. Empirical specifications

The association regressions use one outcome at a time. The roughness estimate and all continuous controls are standardized within the estimation sample. Controls are absolute funding, absolute basis, current log RV, and the matching lagged h-hour outcome; month fixed effects absorb slow shifts. The frozen primary inference is Newey–West (1987) with 112 six-hour lags, a conservative 28-day bandwidth inherited from the companion design. I also report lags 4, 8, and 20, and origin-level 14-day block-cluster standard errors. Averaging all variables within non-overlapping blocks is retained only as a frequency-aggregation diagnostic because it changes the estimand. Binary association regressions are linear probability models, so their coefficients are risk differences.

Tail thresholds are asset-specific and expanding. At origin t , the empirical quantile uses only outcomes whose forward horizon has ended by t . The primary event threshold is 95%; 90% and 99% thresholds are reported as sensitivity checks. Holm (1979) adjustment is applied within each outcome family across the two assets, two horizons, and three estimators.

Recursive forecasts begin in January 2023 after an initial 2020–2022 training period and refit monthly. Historical models use the expanding outcome mean or event rate. State models use funding, basis, current log RV, and the lagged outcome, a deliberately lean alternative to multi-horizon HAR dynamics (Corsi, 2009); State+H adds one roughness estimator. Continuous outcomes use ordinary least squares and squared error. Binary forecasts use logistic regression and Brier loss, so fitted probabilities remain in $[0,1]$. Each candidate is compared separately with Historical and State on identical dates, with HAC Diebold–Mariano tests and Holm adjustment by outcome and benchmark (Diebold and Mariano, 1995).

Use of generative AI tools in this study was limited to language editing, drafting support, and code review. The author independently verified all data construction, code, estimates, citations, and interpretations.

5. Results

5.1 A short-horizon quarticity association

Table 3 begins with the frozen 112-lag specification. IV and PA move together at eight hours: 0.186 and 0.177 log points in Bitcoin, 0.166 and 0.157 in Ether. Ether passes the prespecified Holm rule; Bitcoin does not, with adjusted p-values of 0.092 and 0.085. PV is close to zero. H1 is therefore only partially supported under the primary rule, despite four positive IV/PA estimates.

Table 3. Eight-hour scale, shape, and bipower-residual associations

Asset	H	Log RQ	Log NRQ	Residual share	Holm p (RQ)
BTC	IV	0.186 (2.57)	-0.0023 (-0.13)	-0.0011 (-0.60)	0.092
BTC	PA	0.177 (2.63)	-0.0086 (-0.51)	-0.0018 (-1.08)	0.085
BTC	PV	0.0076 (0.31)	0.0006 (0.07)	-0.0007 (-0.88)	1.000
ETH	IV	0.166 (3.28)	-0.0027 (-0.21)	0.0009 (0.65)	0.013
ETH	PA	0.157 (3.27)	-0.015 (-1.24)	0.0002 (0.15)	0.013
ETH	PV	0.023 (1.03)	0.0030 (0.40)	0.0014 (2.04)	1.000

Effects per one-standard-deviation increase in H; 112-lag HAC t-statistics in parentheses. Controls: absolute funding, absolute basis, current log RV, matching lagged outcome, and month fixed effects.

Figure 1. Continuous-outcome coefficients with 95% Newey–West intervals. The residual-share panels use $\max(RV - BV, 0)/RV$.

5.2 The RQ loading is scale accounting

The exact identity permits a coefficient-level audit. I re-estimate log RQ, log RV, and log NRQ on one common sample with one common design matrix. For every row, β_{RQ} equals $2\beta_{\text{RV}} + \beta_{\text{NRQ}}$ to numerical precision. Table 4 reports the eight-hour coefficients. The IV/PA RQ loading is almost entirely twice the RV loading; NRQ contributes little and sometimes offsets it. This is more informative than noting that normalization lowers significance: it identifies the channel mechanically and empirically.

Table 4. Common-sample accounting decomposition at eight hours

Asset	H	β : log RQ	β : log RV	β : log NRQ	$2\beta_{\text{RV}} + \beta_{\text{NRQ}}$
BTC	IV	0.139 (2.09)	0.067 (2.62)	0.004 (0.20)	0.139
BTC	PA	0.132 (2.11)	0.067 (2.74)	-0.002 (-0.13)	0.132
BTC	PV	0.012 (0.52)	0.007 (0.74)	-0.001 (-0.11)	0.012
ETH	IV	0.122 (2.63)	0.058 (3.15)	0.005 (0.38)	0.122
ETH	PA	0.112 (2.58)	0.059 (3.26)	-0.006 (-0.48)	0.112
ETH	PV	0.020 (0.93)	0.009 (1.04)	0.002 (0.33)	0.020

All three equations use the same observations, standardized regressors, funding, basis, current RV, past RV, past NRQ, and month effects. Across both horizons, the largest absolute coefficient-identity gap is below 3×10^{-14} .

5.3 Shape and bipower-residual effects are economically small

Failure to reject zero would be weak evidence for a boundary. H2 and H3 instead specify equivalence regions. With ± 0.20 log points for NRQ and ± 0.01 share units for the raw bipower residual, all 24 asset-horizon-estimator coefficients pass Holm-adjusted two one-sided tests. The 90% confidence intervals sit well inside their respective regions. A conservative significance-filtered residual produces the same result in all 12 cells. H2 and H3 are supported at the stated resolutions.

Table 5. Familywise equivalence tests

Outcome	Bound	Cells	Envelope of 90% CIs	Max Holm TOST p	Pass
Log NRQ	± 0.20	12	[-0.114, 0.027]	<0.001	12
Raw residual share	± 0.01	12	[-0.006, 0.004]	<0.001	12
Filtered residual share	± 0.01	12	[-0.005, 0.001]	<0.001	12

TOST rejects effects outside the prespecified equivalence region. Pass is the number of cells meeting the familywise 5% rule. The filtered residual uses a one-sided 1% RQ-scaled BNS-style screen with the finite-sample BV correction.

5.4 Dependence corrections change precision, not the scale–shape split

The 112-lag bandwidth is deliberately severe. At eight hours, IV and PA remain positive and individually significant at lags 4, 8, 20, and 112. Origin-level 14-day block clustering is also favorable: all four cross-contract IV/PA intervals exclude zero after Holm correction, while PV remains null. Temporal aggregation tells a different story—the coefficient on a 14-day block mean can reverse sign—but that regression asks whether low-frequency block averages co-move. It is not a standard-error calculation for the six-hour estimand.

Table 6. Origin-level 14-day block-cluster inference for log RQ

Asset	h	H	β	95% CI	Holm p	80% MDE
BTC	8	IV	0.186	[0.062, 0.310]	0.030	0.177
BTC	8	PA	0.177	[0.061, 0.294]	0.028	0.167
BTC	8	PV	0.008	[-0.036, 0.051]	1.000	0.063
BTC	24	IV	0.120	[-0.032, 0.272]	0.843	0.217
BTC	24	PA	0.095	[-0.057, 0.248]	1.000	0.218
BTC	24	PV	-0.003	[-0.047, 0.041]	1.000	0.063
ETH	8	IV	0.166	[0.074, 0.258]	0.004	0.132
ETH	8	PA	0.157	[0.071, 0.243]	0.004	0.123
ETH	8	PV	0.023	[-0.022, 0.068]	1.000	0.064
ETH	24	IV	0.103	[-0.014, 0.219]	0.667	0.166
ETH	24	PA	0.081	[-0.033, 0.196]	0.977	0.164
ETH	24	PV	0.015	[-0.028, 0.058]	1.000	0.061

There are 171 calendar blocks. Holm adjustment is across the 12 asset-horizon-estimator log-RQ cells. MDE is the two-sided 5% minimum detectable coefficient at 80% power using the estimated cluster SE.

5.5 Forecast value is real in Bitcoin and nonportable to Ether

Table 7 compares each raw-RQ forecast with both simple benchmarks. All six Bitcoin specifications reduce loss relative to Historical and State; the gains against State range from 0.51% to 1.21% at eight hours and 3.10% to 3.94% at 24 hours, with corrected equal-loss rejections. Ether behaves differently. Every eight-hour addition raises loss relative to State, and at 24 hours only PV yields a corrected gain. H4 fails because portability was part of the rule, not because the Bitcoin gains are numerically negligible.

Table 7. Recursive log-RQ forecast comparisons

Asset	h	H	vs Historical	Holm p	vs State	Holm p
BTC	8	IV	+44.24%	<0.001	+0.51%	0.028
BTC	8	PA	+44.42%	<0.001	+0.83%	0.003
BTC	8	PV	+44.65%	<0.001	+1.21%	0.028
BTC	24	IV	+31.17%	<0.001	+3.10%	<0.001
BTC	24	PA	+31.69%	<0.001	+3.83%	<0.001
BTC	24	PV	+31.69%	<0.001	+3.94%	0.001
ETH	8	IV	+41.20%	<0.001	-0.51%	<0.001
ETH	8	PA	+41.25%	<0.001	-0.42%	<0.001
ETH	8	PV	+41.21%	<0.001	-0.35%	0.081
ETH	24	IV	+31.26%	<0.001	-0.13%	1.000
ETH	24	PA	+31.43%	<0.001	+0.12%	1.000
ETH	24	PV	+32.58%	<0.001	+2.30%	0.028

Positive percentages mean lower squared-error loss. Forecasts start in January 2023, refit monthly, and are evaluated on model-common dates. P-values are HAC Diebold–Mariano tests, adjusted within the log-RQ benchmark family.

The time split resolves an apparent tension with the weakening post-2023 full-sample slope. Bitcoin's gains against State increase in 2025–2026: for example, IV rises from 0.25% to 0.93% at eight hours and from 1.35% to 4.91% at 24 hours. A changing recursive interaction can improve prediction without restoring one stable full-sample coefficient. Ether's eight-hour losses persist in both evaluation subperiods. Binary models are monthly-refit logits; realized event rates range from 3.12% to 6.23%, and no fitted probability leaves the unit interval.

6. Robustness and falsification checks

The scale identity is exact to numerical precision. Trimming or winsorizing log RQ at the asset-horizon 0.5% and 99.5% quantiles leaves the eight-hour IV/PA estimates close to baseline. q90, q95, and q99 event definitions yield no

corrected NRQ or residual-share association; the smallest adjusted event p-value is 0.098. Restricting IV and PA to dates on which PV has an admissible root also preserves the scale–shape reading.

The significance screen is intentionally conservative because the minute-sufficient tape does not retain tripower variation. BV is multiplied by $n/(n-1)$, and the positive RV–BV residual must exceed a one-sided 1% statistic scaled by RQ. The screen retains 8.49% and 14.77% of Bitcoin's 8-hour and 24-hour windows, and 6.62% and 13.53% of Ether's. Roughness coefficients remain small; none has raw p below 0.095 and all satisfy the ± 0.01 familywise equivalence bound. I treat this as a robustness screen, not as a canonical BNS jump test.

Subperiod slopes weaken after January 2023, whereas some recursive Bitcoin gains strengthen later. Those statements concern different objects. The former estimates one constant association within each subperiod; the latter asks whether a monthly refitted model lowers realized loss. Finally, 14-day block-average regressions reverse some signs. Their unit is a low-frequency average rather than a six-hour origin, so they are evidence that the association does not aggregate cleanly—not evidence that the origin-level standard errors were miscomputed.

7. Discussion

The eight-hour result is not a generic null. IV and PA both anticipate the scale of fourth-moment variation, and block-cluster inference makes the cross-contract pattern more—not less—visible. The accounting decomposition then shows why the result stops at raw quarticity. Roughness loads on future RV; because RQ contains RV squared, that scale loading passes into RQ almost arithmetically. Little remains in NRQ.

The estimator comparison matters for that judgment. IV and PA again act as one family, whereas PV contributes little to the origin-level scale association. This echoes the functional boundary in Yang (2026b), now on a future outcome rather than contemporaneous funding. The result is not another funding regression. It identifies the part of tail measurement to which that estimator family is connected.

The forecast results are stronger than the original draft allowed for Bitcoin. Each H specification beats both benchmarks for raw RQ, and the 24-hour improvements are not tiny. But Ether refuses the same interpretation at eight hours. This contract split is substantive: a predictor can be useful in one liquid perpetual without defining a portable crypto tail factor. PV's smaller admissible-root sample adds another reason to resist pooling the forecast map.

For empirical risk work, the practical lesson is procedural. Report raw quarticity together with its RV and NRQ components. Call $\max(RV-BV, 0)/RV$ a residual share unless an explicit jump test supports a stronger name. And compare a proposed signal with Historical as well as State on the same dates. Those three steps prevent a useful scale predictor from being promoted into a tail-shape claim.

8. Limitations

The market boundary remains one exchange and two perpetual contracts. That continuity makes the comparison with the companion studies exact, but it limits external validity. Liquidations are omitted because the frozen package has no stable full-window public archive. The bipower residual is a high-frequency decomposition, not a record of exchange-labelled liquidations.

One-minute returns make the construction transparent but do not eliminate noise. RQ is especially sensitive to isolated minute moves. Complete windows, no interpolation, outlier checks, and NRQ address distinct parts of that problem; none turns RQ into an efficient noise-robust quarticity estimator. The RQ-scaled filter is constrained by the sufficient-statistic archive and cannot reproduce a tripower-based canonical BNS statistic.

The event threshold is expanding rather than structural. It adapts to history and prevents look-ahead, but its realized event rate need not equal 5% in a nonstationary market. Forecast evaluation likewise covers only January 2023–July 2026. These choices support a disciplined pseudo-real-time exercise; they do not identify a trading strategy or causal stress channel.

9. Conclusion

On the frozen Bitcoin and Ether perpetual-futures tape, IV and PA roughness estimates anticipate eight-hour realized quarticity. The coefficient identity locates that association in future variance scale, and familywise equivalence tests

rule out effects of the prespecified size on normalized quarticity and bipower-residual shares. Bitcoin also delivers corrected raw-RQ forecast gains against both simple benchmarks. Ether does not reproduce the eight-hour gain.

My reading is correspondingly specific. Apparent roughness can help rank how large near-term fourth-moment variation becomes in a contract. In these data it does not rank tail concentration after scale is removed, and its predictive value does not travel uniformly from Bitcoin to Ether. That is the scale–shape divide.

Appendix A. Additional numerical checks

Table A1. Smallest p-value in each event-sensitivity family

Measure	Threshold	Asset	H	Effect (t)	Holm p
Jump Share	q90	BTC	PA	-0.014 (-2.65)	0.098
Jump Share	q95	BTC	PV	-0.0033 (-1.72)	1.000
Jump Share	q99	ETH	PV	0.0026 (2.31)	0.248
Normalized Quarticity	q90	BTC	PA	-0.018 (-2.16)	0.368
Normalized Quarticity	q95	BTC	PA	-0.010 (-1.56)	1.000
Normalized Quarticity	q99	ETH	PV	0.0023 (1.74)	0.985

Each Holm family contains 12 asset-horizon-estimator tests. Reporting the smallest cell makes the multiplicity boundary explicit.

Table A2. Eight-hour raw-quarticity outlier sensitivity

Asset	H	Treatment	Effect (t)	Holm p
BTC	IV	Winsor	0.181 (2.54)	0.099
BTC	IV	Trim	0.175 (2.53)	0.103
BTC	PA	Winsor	0.175 (2.62)	0.088
BTC	PA	Trim	0.180 (2.74)	0.061
BTC	PV	Winsor	0.0074 (0.31)	1.000
BTC	PV	Trim	0.0034 (0.15)	1.000
ETH	IV	Winsor	0.164 (3.24)	0.014
ETH	IV	Trim	0.164 (3.28)	0.011
ETH	PA	Winsor	0.157 (3.24)	0.014
ETH	PA	Trim	0.164 (3.46)	0.006
ETH	PV	Winsor	0.020 (0.93)	1.000
ETH	PV	Trim	0.012 (0.55)	1.000

Outcomes are winsorized or trimmed at the asset-horizon 0.5% and 99.5% quantiles. Controls and 12-test Holm families match the baseline specification.

Table A3. Twenty-four-hour continuous associations

Asset	H	Log RQ	Log NRQ	Residual share
BTC	IV	0.120 (1.38)	-0.044 (-1.59)	-0.0020 (-0.97)
BTC	PA	0.095 (1.09)	-0.067 (-2.30)	-0.0032 (-1.64)
BTC	PV	-0.0032 (-0.13)	-0.0003 (-0.03)	-0.0000 (-0.06)
ETH	IV	0.103 (1.63)	-0.018 (-0.76)	0.0015 (0.94)
ETH	PA	0.081 (1.31)	-0.043 (-1.93)	0.0005 (0.31)
ETH	PV	0.015 (0.65)	0.0049 (0.48)	0.0016 (2.49)

Effects per one-standard-deviation H; 112-lag HAC t-statistics in parentheses. No coefficient in this table survives its 12-cell outcome-family Holm adjustment.

Table A4. Eight-hour log-RQ bandwidth sensitivity

Asset	H	4 lags	8 lags	20 lags	112 lags
BTC	IV	0.186 (4.35)	0.186 (3.97)	0.186 (3.57)	0.186 (2.57)
BTC	PA	0.177 (4.14)	0.177 (3.83)	0.177 (3.54)	0.177 (2.63)
ETH	IV	0.166 (4.82)	0.166 (4.43)	0.166 (4.06)	0.166 (3.28)
ETH	PA	0.157 (4.77)	0.157 (4.43)	0.157 (4.15)	0.157 (3.27)

Coefficients are identical across columns; HAC t-statistics are in parentheses. Four and 112 six-hour lags correspond to one day and 28 days.

Table A5. Full dual-benchmark forecast audit

Asset	h	Outcome	IV: H/S (p)	PA: H/S (p)	PV: H/S (p)
BTC	8	Log RQ	+44.2/+0.5 (0.028)	+44.4/+0.8 (0.003)	+44.7/+1.2 (0.028)
BTC	8	Log NRQ	-1.6/+4.5 (1.000)	-2.2/+3.9 (1.000)	-0.4/+5.4 (1.000)
BTC	8	NRQ q95	-32.3/+3.8 (<0.001)	-30.9/+4.8 (<0.001)	-0.4/+25.9 (1.000)
BTC	8	Residual q95	+12.1/+1.0 (0.281)	+12.3/+1.3 (0.080)	+11.3/+0.3 (0.807)

Asset	h	Outcome	IV: H/S (p)	PA: H/S (p)	PV: H/S (p)
BTC	24	Log RQ	+31.2/+3.1 (<0.001)	+31.7/+3.8 (<0.001)	+31.7/+3.9 (0.001)
BTC	24	Log NRQ	-1.3/+5.9 (1.000)	-2.2/+5.1 (1.000)	-2.0/+5.1 (1.000)
BTC	24	NRQ q95	+0.0/-0.3 (1.000)	+0.1/-0.2 (1.000)	-0.5/-0.8 (1.000)
BTC	24	Residual q95	+5.9/+1.8 (0.072)	+5.8/+1.6 (0.043)	+4.8/+0.6 (0.281)
ETH	8	Log RQ	+41.2/-0.5 (<0.001)	+41.3/-0.4 (<0.001)	+41.2/-0.3 (0.081)
ETH	8	Log NRQ	+4.0/+0.4 (0.123)	+3.8/+0.1 (0.527)	+3.7/+0.4 (0.542)
ETH	8	NRQ q95	-11.5/+0.3 (0.163)	-11.3/+0.5 (0.009)	-9.4/+2.4 (0.002)
ETH	8	Residual q95	-1.5/-0.5 (1.000)	-0.9/+0.0 (1.000)	+4.9/+6.2 (0.025)
ETH	24	Log RQ	+31.3/-0.1 (1.000)	+31.4/+0.1 (1.000)	+32.6/+2.3 (0.028)
ETH	24	Log NRQ	+3.6/+0.2 (1.000)	+3.5/+0.1 (1.000)	+3.7/+2.3 (1.000)
ETH	24	NRQ q95	-18.4/+0.7 (0.468)	-18.2/+0.8 (0.123)	-19.2/+1.7 (0.004)
ETH	24	Residual q95	-17.2/+1.2 (0.018)	-17.3/+1.1 (0.018)	-18.8/+1.3 (0.018)

Each cell reports percentage loss improvement relative to Historical/State and, in parentheses, the larger of the two benchmark-specific Holm p -values. A strict H_4 success requires both improvements to be positive and the reported p -value below 0.05. Exact DM statistics and forecast counts are retained in the replication table.

Appendix B. Estimator and outcome algorithms

B.1 Increment-variance functional

Let X_j be log hourly RV in the preceding 336-hour window and $L = \{1, 2, 3, 6, 12\}$ hours. For each ℓ , $V(\ell)$ is the population-normalized sample variance of $X_{j+\ell} - X_j$. $\hat{H}^{IV}$ is one half of the OLS slope in $\log V(\ell) = a + b \log \ell$. A window must be finite and supply at least two positive increment variances.

B.2 Pre-averaged functional

Within each hour, minute returns are pre-averaged with $k = 8$ and $g(u) = u(1-u)$. Writing $\psi_2 = k^{-1}\Sigma g(j/k)^2$ and $\psi_1 = k\Sigma \Delta g(j/k)^2$, the hourly estimator is

$$PAIV = [\psi_2 k]^{-1} \Sigma_i (\Sigma_{j=1}^{k-1} g(j/k) r_{i+j})^2 - [\psi_1 / (2\psi_2 k^2)] \Sigma_i r_i^2.$$

Values are floored at 10^{-18} before logging. $\hat{H}^{PA}$ applies the B.1 slope to the 336 hourly log PAIV values. The construction follows the pre-averaging logic of Jacod et al. (2009); it is used as a noise-robust measurement transform, not as a point estimate of spot regularity.

B.3 Normalized p-variation functional

For the 336-point path $X = \sqrt{RV}$, $L = 335$ fine increments are partitioned into $K = 18$ near-equal coarse intervals. With endpoints b_i and interval weights $(b_{i+1} - b_i)/L$,

$$W(p) = \Sigma_i [(b_{i+1} - b_i)/L] |X(b_{i+1}) - X(b_i)|^p / \Sigma_{j=b_i}^{b_{i+1}-1} |X(j+1) - X(j)|^p.$$

A fixed 77-point grid over $p \in [1.01, 20]$ locates crossings of $\log W(p) = 0$; bisection refines a unique crossing and $\hat{H}^{PV} = 1/p$. Windows with no root, multiple roots, nonfinite inputs, or degenerate increments remain missing. This implements the normalized p -variation diagnostic of Cont and Das (2024).

B.4 Residual significance screen

For n minute returns, BV is multiplied by $n/(n-1)$. The positive residual RV–BV enters only when

$$(RV - [n/(n-1)]BV) / \sqrt{\{[(\pi/2)^2 + \pi - 5]RQ\}/n} > \Phi^{-1}(0.99).$$

Otherwise the filtered share is zero. Because the archive lacks tripower variation, this is reported as an RQ-scaled BNS-style screen rather than the canonical feasible BNS jump statistic.

Declarations

Funding. This research received no external funding.

Declaration of interest. The author declares no competing financial or personal interests.

Ethics statement. This study uses public market data and involves no human participants, animals, personal data, or interventions. Ethics approval and informed consent were not required.

Data and code availability. Source data are publicly available from Binance Vision. Hourly minute-sufficient panels, coverage audits, analysis code, regression outputs, and computational documentation accompany the

working-paper replication package. A versioned public repository and persistent data DOI will be supplied before journal publication; no DOI is claimed in the present draft.

Relation to previous work. This paper reuses the public tape and frozen measurement design of Yang (2026a), SSRN 7378018, Yang (2026b), SSRN 7393578 (<https://doi.org/10.2139/ssrn.7393578>), and Yang (2026c), SSRN 7397158 (<https://doi.org/10.2139/ssrn.7397158>). It estimates new minute-based quarticity, bipower-residual, event, equivalence, block-cluster, and recursive-loss objects. No companion-paper coefficient is presented as a new estimate.

Use of generative AI. Generative AI tools were used for language editing, drafting assistance, and code review. The author checked and remains responsible for the data construction, empirical design, estimates, citations, interpretation, and final text.

References

- Ackerer, D., Hugonnier, J., and Jermann, U. J. (2026). Perpetual futures pricing. *Mathematical Finance*, 36(3), 481-499. <https://doi.org/10.1111/mafi.70018>.
- Andersen, T. G., Bollerslev, T., Diebold, F. X., and Labys, P. (2003). Modeling and forecasting realized volatility. *Econometrica*, 71(2), 579-625. <https://doi.org/10.1111/1468-0262.00418>.
- Angeris, G., Chitra, T., Evans, A., and Lorig, M. (2023). A primer on perpetuals. *SIAM Journal on Financial Mathematics*, 14(3), SC17-SC30. <https://doi.org/10.1137/22M1520931>.
- Barndorff-Nielsen, O. E., and Shephard, N. (2002). Econometric analysis of realized volatility and its use in estimating stochastic volatility models. *Journal of the Royal Statistical Society: Series B*, 64(2), 253-280. <https://doi.org/10.1111/1467-9868.00336>.
- Barndorff-Nielsen, O. E., and Shephard, N. (2006). Econometrics of testing for jumps in financial economics using bipower variation. *Journal of Financial Econometrics*, 4(1), 1-30. <https://doi.org/10.1093/jfinec/nbi022>.
- Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A., and Shephard, N. (2008). Designing realized kernels to measure the ex post variation of equity prices in the presence of noise. *Econometrica*, 76(6), 1481-1536. <https://doi.org/10.3982/ECTA6495>.
- Binance. (2026). Binance public data. <https://data.binance.vision/>.
- Bennedsen, M., Lunde, A., and Pakkanen, M. S. (2022). Decoupling the short- and long-term behavior of stochastic volatility. *Journal of Financial Econometrics*, 20(5), 961-1006. <https://doi.org/10.1093/jfinec/nbaa049>.
- Bolko, A. E., Christensen, K., Pakkanen, M. S., and Veliyev, B. (2023). A GMM approach to estimate the roughness of stochastic volatility. *Journal of Econometrics*, 235(2), 745-778. <https://doi.org/10.1016/j.jeconom.2022.06.009>.
- Cont, R., and Das, P. (2024). Rough volatility: Fact or artefact? *Sankhya B*, 86(1), 191-223. <https://doi.org/10.1007/s13571-024-00322-2>.
- Corsi, F. (2009). A simple approximate long-memory model of realized volatility. *Journal of Financial Econometrics*, 7(2), 174-196. <https://doi.org/10.1093/jfinec/nbp001>.
- 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>.
- Gatheral, J., Jaisson, T., and Rosenbaum, M. (2018). Volatility is rough. *Quantitative Finance*, 18(6), 933-949. <https://doi.org/10.1080/14697688.2017.1393551>.
- Hansen, P. R., and Lunde, A. (2006). Realized variance and market microstructure noise. *Journal of Business & Economic Statistics*, 24(2), 127-161. <https://doi.org/10.1198/073500106000000071>.
- Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, 6(2), 65-70.
- Jacod, J., Li, Y., Mykland, P. A., Podolskij, M., and Vetter, M. (2009). Microstructure noise in the continuous case: The pre-averaging approach. *Stochastic Processes and their Applications*, 119(7), 2249-2276. <https://doi.org/10.1016/j.spa.2008.11.004>.
- Newey, W. K., and West, K. D. (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, 55(3), 703-708. <https://doi.org/10.2307/1913610>.
- Patton, A. J. (2011). Volatility forecast comparison using imperfect volatility proxies. *Journal of Econometrics*, 160(1), 246-256. <https://doi.org/10.1016/j.jeconom.2010.03.034>.
- Schirmann, D. J. (1987). A comparison of the two one-sided tests procedure and the power approach for assessing the equivalence of average bioavailability. *Journal of Pharmacokinetics and Biopharmaceutics*, 15(6), 657-680. <https://doi.org/10.1007/BF01068419>.
- Schmeling, M., Schrimpf, A., and Todorov, K. (2026). Crypto carry. *Management Science, Articles in Advance*. <https://doi.org/10.1287/mnsc.2024.05069>.
- Yang, F. (2026a). Apparent volatility roughness and perpetual-futures funding: Persistent association, weak transition information. SSRN 7378018. <https://doi.org/10.2139/ssrn.7378018>.
- Yang, F. (2026b). Apparent roughness across estimators: Does the funding association survive noise-robust measurement? SSRN 7393578. <https://doi.org/10.2139/ssrn.7393578>.
- Yang, F. (2026c). Apparent roughness and funding-rate asymmetry: Evidence from Bitcoin and Ether perpetuals. SSRN 7397158 (<https://doi.org/10.2139/ssrn.7397158>).