Costello College of Business
FUTURE OF FINANCE
LAB
George Mason University

Perpetual Futures Fragility: The Effect of Stablecoin Funding and Reference-Price Construction *

Bera Gumruk Aarush Baradwaj Krishna Malghan
Derek Horstmeyer

Department of Finance, Costello College of Business, George Mason University

Abstract

Perpetual futures are financial contracts that never expire and are rapidly gaining popularity on regulated exchanges worldwide. Perpetuals, in theory, are able to maintain a stable price tied to a real-world, continuously traded "reference price" by having periodic payment between buyers and sellers to maintain the peg. That reference price, however, is only ever as dependable as the market it is drawn from. We test what happens when that underlying market is thin and barely traded. Through the worst crypto crash in our sample's history (October 10, 2025), the contract with the thinnest reference market saw its reference price swing 64.5% in an hour, six times more than a comparable crypto contract, with trading costs spiking to nearly 100%, even though it used exactly the same standard pricing method as the others. A second contract, one whose price the exchange sets directly but which tracks a deep and liquid market, stayed calm throughout. We then compare 20 pairs of contracts that track the exact same asset on the same exchange in the same hour. The exchange-deployed version is consistently further from fair value, in 19 of the 20 pairs and by about ten times at the median. Using the underlying price feeds directly, we find this is almost entirely because those contracts barely trade, not because the exchange-set

*We appreciate the support of the Advisory Board of the Future of Finance Lab at George Mason University. This is a rough draft. Please do not cite without permission. Corresponding author: Costello College of Business, George Mason University, 4400 University Drive, Fairfax, VA 22030; dhorstme@gmu.edu

price is worse: the two price feeds agree to within a twenty-thousandth of a percent for nearly every pair. What the feeds do reveal is variation. Several underlyings are listed by multiple independent operators at once, and comparing those feeds against each other shows agreement differing by more than a factor of twenty on Bitcoin alone, and update frequency differing roughly twofold to threefold among operators tracking the identical index. One S&P 500 contract carrying several million dollars a day runs on a price that does not move 55% of the time, while a competing contract on the same index updates more than twice as often. Nothing visible to a trader distinguishes them. The exposure is not the pricing method, it is who operates it. Stress in the stablecoin (USDC) used to settle these contracts also spills over into the contracts themselves during extreme stress, though not during everyday fluctuations. These results show that venue-deployed perpetuals are persistently more dislocated than validator-priced contracts on the identical underlying ($p = 2.0 \times 10^{-5}$ across 20 matched pairs), but that oracle-level decomposition attributes this to sparse trading rather than to reference-price construction, which contributes only 3–10% of the gap and does not deteriorate under stress. Comparing independent deployers listing the identical underlying instead reveals wide dispersion in reference-price infrastructure: return agreement against the validator median differs by a factor of twenty on Bitcoin, and oracle update frequency varies roughly twofold to threefold across the four verified S&P 500 and gold listings, which face identical trading calendars. That dispersion, undisclosed at the interface level, is the finding most relevant as the CFTC moves toward approving listed perpetual futures in the U.S.

1 Introduction

A perpetual futures contract, or perpetual, is a derivative with no expiration date. Unlike a traditional futures contract, which forces its price to converge to the underlying asset's spot price by a fixed delivery date, a perpetual has no delivery date to force that convergence. Instead, it relies on a periodic funding payment, exchanged directly between traders holding long positions and traders holding short positions, whichever side is pushing the contract's traded price away from a reference price pays the other side. If the funding mechanism works as intended, this payment continuously pulls the traded price back toward the reference, doing the job that physical delivery does for a traditional futures contract. The reference price itself, commonly called the oracle price, is therefore not a minor implementation detail but the anchor the entire mechanism depends on. Shiller (1993) first proposed this construction, cash-settled contracts referenced to an index rather than a deliverable asset, as a way to create markets for goods that cannot be physically delivered. Crypto exchanges have since built the idea into one of the largest derivatives markets in the world.

The traditional futures market offers a useful point of contrast. CME's West Texas Intermediate crude oil contract enforces mandatory physical delivery at Cushing, Oklahoma, backed by roughly two dozen pipelines and fifteen storage terminals with a combined capacity near 90 million barrels; in 2024 alone, more than 317,000 contracts, over 317 million barrels, were actually physically delivered. Because delivery is real and enforced, the futures price cannot permanently diverge from the physical market: convergence at expiration is not optional. A perpetual on a continuously traded, deeply liquid asset like Bitcoin can replicate something close to this discipline, since the oracle price itself is built from live trading on major exchanges. But a growing share of the perpetuals market now references assets that are not continuously, publicly traded at all: pre-IPO equities, commodities without a matching perpetual-friendly spot venue, and other thinly traded reference instruments. On these contracts, known as HIP-3 (builder-deployed) perpetuals on the Hyperliquid exchange studied in this paper, the oracle price is not computed from a multi-exchange median the way native contracts' prices are; it is set directly by the deployer that built the market. There is no equivalent of Cushing, Oklahoma, no physical constraint forcing the reference price to reflect reality. This is not merely a theoretical construction: onchain perpetuals on U.S. equities and commodities already carry real price information during the hours cash markets are closed, single-name equity perps trade within 50 basis points of the next cash open six hours ahead of it, and within 25 basis points one hour out, while the same signal at the weekend close is a full 113 basis points wide, evidence that this asset class serves a genuine price-discovery function precisely when the underlying market cannot itself be

observed (Shehdula, 2026b). Whether this construction meaningfully weakens the funding mechanism's ability to anchor price, particularly during periods of market stress, is the empirical question this paper answers.

History offers a caution about what happens when a large derivatives market rests on a reference price that turns out not to be as solid as assumed. The London Interbank Offered Rate (LIBOR) served as the reference rate for hundreds of trillions of dollars in financial contracts for decades before regulators discovered that submitting banks had been manipulating it, motivating a wave of research on how reference rates should be constructed to resist exactly this kind of fragility (Duffie and Stein, 2015; Duffie and Dworczak, 2021). One proposed remedy, moving LIBOR from a survey of banks' self-reported estimates to a fixing computed directly from actual transactions, ran into the same underlying problem this paper studies: at many currency-maturity pairs, transaction volume was simply too thin to support a reliable daily fixing, forcing a tradeoff between a multi-day sampling window wide enough to reduce noise and a window narrow enough to stay timely (Duffie et al., 2013). A reference price's vulnerability to thin, sparse trading is not unique to crypto; the same tension between construction and observability that a deployer-set HIP-3 oracle faces on a much shorter timescale already shaped how regulators tried to fix LIBOR itself. Foreign exchange benchmarks saw a parallel episode: the WM/Reuters 4pm London fix, the reference price used to value trillions of dollars in FX exposure, was found to have been manipulated by traders at several major banks, and the CFTC fined five of them more than $1.4 billion in a single 2014 enforcement action. Subsequent academic work has continued to examine whether the post-scandal reforms to that benchmark's methodology actually made it more robust (Benenchia et al., 2024). Both episodes share a structural feature relevant here: a reference price that looked authoritative because it was widely used, not because its construction was actually resistant to manipulation or stress. Perpetual futures on deployer-priced assets raise the same question in a new setting, one that, unlike LIBOR and the FX fix, has not yet been examined empirically.

The question also arrives at a moment of real regulatory consequence. The Commodity Futures Trading Commission has moved in 2026 toward formally approving listed perpetual futures for U.S. markets, a shift its own chair has publicly discussed and defended against common misconceptions about the product. That regulatory opening has already produced friction: CME Group filed suit against the CFTC on June 18, 2026, after the agency cleared Kalshi to list its own perpetual-style contracts (WSJ, 2026a). The dispute is fundamentally about classification: CME argues that Kalshi's product is properly a swap, which would place it under a heavier regulatory rulebook effectively out of reach for retail traders, while the CFTC treated it as a futures contract, subject to lighter oversight and tradeable on a

regulated exchange by ordinary investors; the CFTC has called the suit frivolous. This is not a technicality. Which side of that classification line a perpetual contract falls on determines who is legally permitted to trade it and how much investor protection surrounds it, exactly the kind of design question that evidence on reference-price fragility should inform.

The stakes of getting that design question wrong extend beyond market structure to individual investors. As perpetuals expand from crypto-native assets into U.S. single-stock and commodity references, distinct from CME's own newly launched single-stock futures, which retain a fixed expiration, financial commentary has raised the concern that the product's core features, no expiration date, leverage reported as high as 100x, and reliance on potentially thinner counterparties than a traditional futures market provides, could let forced liquidations during a selloff cascade and spill losses onto ordinary investors well beyond the traders directly holding the position (Jakab, 2026). This paper's own evidence, that deployer-priced reference construction concentrates exactly this kind of fragility during acute stress, speaks directly to that concern: a perpetual whose oracle is not built from a continuously observed, multi-exchange market is not just theoretically weaker, it is empirically more prone to the reference-price and liquidity breakdowns that turn an ordinary selloff into a cascading one.

As U.S. exchanges and regulators work out what a well-constructed, exchange-listed perpetual should look like, evidence on which reference-price designs hold up under stress, and which do not, is directly useful to that process, not merely of academic interest.

This paper asks two related questions about reference-price construction and perpetual futures market quality, both using real, minute-level data on six candidate contracts drawn from Hyperliquid via Allium: five native, market-priced contracts (ETH, SOL, XRP, HYPE, and SPX) and one deployer-priced HIP-3 contract (xyz:GOLD). Critically, these two attributes vary independently across the sample: SPX carries the standard validator-median oracle but references a thin, sparsely traded token, while xyz:GOLD carries a deployer-set oracle but references gold, among the deepest and most continuously observed markets in the world. That independent variation is what allows the two channels to be separated. The first question, RQ1, asks whether reference-asset volatility weakens the funding mechanism's anchoring power, and whether this effect is driven by the thinness of the underlying reference market or by the oracle's construction, using the October 10, 2025 crypto liquidation cascade, the single largest deleveraging event in the sample's history, as the paper's central stress test. The second question, RQ2, asks whether volatility in a perpetual's settlement stablecoin, principally USDC, spills over into the quality of the perpetual market itself, even when the underlying reference asset is entirely unaffected, examined through both the full history of USDC's minor day-to-day price fluctuations and, separately, the March 2023 USDC/Silicon Valley Bank depeg as a genuine natural experiment.

Section 2 reviews the existing literature on perpetual futures pricing and reference-benchmark fragility, and situates this paper's contribution within it. Section 3 presents the empirical analysis: data sources, the divergence measure used throughout, and the results for both research questions, including the October 10, 2025 event study, full-history structural checks, a series of panel regressions, and the March 2023 cross-exchange comparison. Section 4 concludes.

2 Literature Review

Perpetual futures are derivative contracts with no expiration date that use periodic funding payments to keep their traded price anchored to a reference price (Shiller, 1993). Unlike traditional futures, which force price convergence at expiration, perpetuals rely entirely on the funding mechanism to self-correct. Foundational theoretical work establishes that the strength of this anchoring depends on the funding formula's design, assuming the reference price is continuously observable and the settlement currency is stable (Angeris et al., 2023; Ackerer et al., 2024). The industry and critics remain divided on whether this mechanism is truly self-correcting or not viable, and this question is largely untested empirically, particularly as perp markets expand beyond Bitcoin into thinly traded tokens and pre-IPO assets where neither assumption holds cleanly.

A closely related empirical literature studies the funding mechanism indirectly, through the returns it generates. Christin et al. (2023) document that a strategy short the perpetual and long the spot market, the crypto analogue of a traditional carry trade, earns an unusually high Sharpe ratio (8.76 annually for a Tether-margined Bitcoin contract in their sample), evidence that cryptocurrency exchanges were effectively supplying cheap, ready leverage to long-side traders willing to pay a persistent premium for it. Schmeling et al. (2025) extend this line of work with a broader panel and find that the futures-spot basis, what they term crypto carry, can exceed 40% annualized and cannot be explained by interest-rate differentials or storage costs alone. The authors attribute the gap instead to a negative convenience yield driven by demand from smaller, trend-following investors and by real limits to arbitrage, regulatory and margin frictions that stop sophisticated traders from closing the basis even when apparent arbitrage opportunities are large. Both papers study funding-rate and carry dynamics on native, continuously spot-traded contracts, the same class this paper calls market-priced; neither examines what happens to that same mechanism once the reference price itself is no longer a continuously observable spot market, the gap this paper's first research question (RQ1) addresses directly.

Prior empirical research has examined how perpetual markets function under normal

conditions. Studies find that crypto derivatives contribute meaningfully to price discovery, often leading spot markets in incorporating new information (Alexander et al., 2020; He et al., 2022). However, the 8-hour funding settlement structure creates predictable windows of informed trading that cause market makers to widen bid-ask spreads defensively, meaning perpetuals can simultaneously increase volume while worsening spread quality (Ruan and Streltsov, 2022). Research also shows that reported open interest is an unreliable liquidity signal, particularly during stress (Giagkiozis and Said, 2024), and that the fixed interest parameter embedded in most exchanges' funding formulas is mathematically unnecessary for anchoring to work (Angeris et al., 2023). Every existing empirical study, however, examines contracts with actively traded spot markets as their reference, leaving deployer-estimated oracle pricing entirely unexplored.

The perpetual contract's own design choices, documented directly by the exchange that popularized the modern product, help explain why reference-price construction matters so much once markets are stressed. BitMEX's original XBTUSD contract was built specifically to consolidate liquidity that would otherwise fragment across multiple expiration dates onto a single, continuously traded instrument, on the premise that combining long-term and short-term traders in one product deepens the market relative to a traditional futures curve (Hayes, 2016). That same design later extended to assets the exchange could not safely custody: BitMEX's ETHUSD contract uses a quanto structure, paying out in Bitcoin rather than Ether, precisely because holding Ether as collateral carried custody risk the exchange was unwilling to accept (Hayes, 2018). The quanto structure lets a speculator's Bitcoin-denominated return move linearly with the ETHUSD price, but it does so by inserting an additional layer, cross-asset margining, between the contract and the reference price it tracks, introducing basis risk and non-linear hedging dynamics absent from a standard linear contract. Read together with the deployer-set oracle mechanism this paper studies, both are examples of the same underlying pattern: perpetual futures on assets the exchange cannot cleanly, continuously observe or custody require some substitute construction, whether a synthetic settlement currency or a deployer-estimated price, and that substitution is exactly where this paper's evidence shows fragility concentrates under stress.

A separate body of work on stablecoins reveals a second untested channel. Stablecoins settle trillions of dollars in derivatives and serve as the unit of account for nearly every perpetual futures contract. Existing research uses perpetual funding rates as a tool to measure stablecoin devaluation risk during stress events, showing the causation from stablecoin instability to funding signals (Eichengreen et al., 2025). Separately, event-study evidence on Tether (USDT) itself finds that its role as a flight-to-safety asset during market stress is chain-specific rather than uniform, a primary liquidity lifeline for Ethereum holders but a

more muted, stabilizing presence for Bitcoin holders, and stronger even in Wrapped Bitcoin than in native Bitcoin (Chernoff et al., 2026), underscoring that a single stablecoin's stress behavior can propagate very differently depending on the venue and asset it touches. No study has reversed this direction to ask whether stablecoin stress damages the perpetual market itself: widening spreads, increasing price-reference deviations, and disrupting the funding mechanism, even when the underlying asset is unaffected, are not looked at. This gap motivates our second research question, which treats the March 2023 USDC depeg as a natural experiment where the settlement currency broke while Bitcoin remained stable, isolating the event as a distinct and previously unmeasured source of tracking breakdown.

3 Empirical Analysis

3.1 Data Sources

Minute-level Hyperliquid data are obtained from Allium and queried directly using SQL. Replication queries are provided in the supplementary materials (`Allium Queries.sql`).

The core table, `hyperliquid.raw.perpetual_market_asset_contexts`, provides real, minute-level historical snapshots for six candidate perpetual contracts: ETH, SOL, XRP, HYPE, and SPX from May 9, 2025 (∼586,000+ rows each), and xyz:GOLD from November 20, 2025 (∼311,000 rows). Fields used include timestamp, coin, oracle price, mark price, funding rate, open interest, premium, impact (bid/ask) prices, and daily volume. One of the six contracts, xyz:GOLD, is a HIP-3 (builder-deployed) perpetual whose oracle price is set by a deployer-designated address rather than the standard eight-exchange validator median used by native contracts. The other five, including SPX, are native contracts carrying the standard validator-median oracle. `SPX` is unrelated to the S&P 500 index. This is established here from return behavior rather than from the ticker: over roughly 4,050 hourly observations, `SPX` returns correlate 0.647 with BTC, 0.649 with SOL and 0.646 with ETH, and carry a beta of 1.88 to Bitcoin, against a correlation of only 0.321 with `xyz:SP500`, Hyperliquid's actual S&P 500 perpetual. For scale, the venue's genuine S&P 500 listings correlate with one another at 0.94–0.98 (Section 3.4.10). `SPX` therefore trades as a high-beta crypto asset, co-moving with memecoins (0.589 with FARTCOIN, 0.492 with PUMP) rather than with equities. The Hyperliquid Policy Center identifies the underlying as the SPX6900 token; the return evidence above independently establishes only that the contract is not equity-index-linked. Hyperliquid does list genuine S&P 500 perpetuals, but they are HIP-3 markets carrying a deployer namespace (`xyz:SP500`, `km:US500`, and others examined in Section 3.4.10). Three properties distinguish native from HIP-3 listings throughout this paper

and are used consistently: HIP-3 markets are absent from Hyperliquid's main perpetuals universe and addressable only through a deployer-prefixed namespace; no HIP-3 market carries data prior to December 2025, which places the entire category after the October 10, 2025 event studied in Section 3.4.1; and HIP-3 oracle updates are written by a deployer-controlled address rather than the validator set.

For the stablecoin-spillover analysis, USDC's own historical price is drawn from `crosschain.dex.token_prices_hourly`, filtered by canonical contract address (the `symbol` field alone returns unrelated or spurious tokens sharing the same ticker). Hyperliquid's own internal USDC pricing was checked and ruled out for this purpose: it is a flat $1.00 with zero trading volume at every hour tested, since Hyperliquid treats USDC as a fixed unit of account rather than a traded, price-discovered pair.

Because Hyperliquid's own data does not reach back to the March 11, 2023 USDC/Silicon Valley Bank depeg, that event is studied separately through five external, publicly accessible sources: Binance (BTC/ETH spot, hourly), Deribit (BTC-PERPETUAL vs. BTC_USDC-PERPETUAL, and BTC/ETH-PERPETUAL vs. Binance spot), Bybit (BTCPERP vs. BTCUSDT), BitMEX (XBTUSD, cross-venue timing check), and OKX (BTC-USDT-SWAP). A full cross-exchange search of ten venues confirmed Deribit as the only same-exchange, publicly accessible, March-2023-reaching USDC-margined-vs.-non-USDC-margined comparison available.

3.2 Divergence Measure and Units

The core dependent variable throughout is percentage divergence,

$$\text{divergence\_pct} = \frac{\text{mark price} - \text{oracle price}}{\text{oracle price}} \times 100,$$

computed per asset per period. Percentage divergence is used throughout, rather than raw dollar divergence, to ensure comparability across contracts with substantially different reference price levels (SPX ∼$1.20 vs. xyz:GOLD ∼$4,125); a dollar-denominated measure would let the highest-priced asset dominate a pooled regression. All regression results reported below use percentage divergence.

3.3 Summary Statistics

Before turning to the regression results, Table 1 reports full-history descriptive statistics on the paper's core variables for all six perpetual contracts, computed over each asset's entire available trading history on Hyperliquid (roughly 408 days for ETH, SOL, XRP, HYPE,

and SPX; roughly 217 days for xyz:GOLD, which listed later). Panel A covers percentage divergence, the paper's central dependent variable. Panel B covers bid-ask spread, parsed from impact prices. Panel C covers the funding rate; its mean and standard deviation round to approximately zero for every asset at six decimal places, so only the min/max range is reported, the range itself is the meaningful full-history fact given how tightly funding sits near zero.

Table 1: Summary Statistics. All figures are percentages, computed across each asset's full available history on `hyperliquid.raw.perpetual_market_asset_contexts`. Divergence and spread statistics are absolute values; funding reports the signed min/max range.

Panel A: Percentage Divergence (mark price vs. oracle price)				
Coin	N (minutes)	Mean |Divergence|	Std. Dev.	Max
SPX	587,796	0.07	0.10	15.26
HYPE	587,805	0.05	0.07	7.41
xyz:GOLD	312,587	0.04	0.07	4.90
SOL	587,983	0.05	0.05	10.54
XRP	587,910	0.05	0.05	3.30
ETH	587,994	0.04	0.04	3.00
Panel B: Bid-Ask Spread				
Coin		Mean	Std. Dev.	Max
SPX		0.12	0.31	99.69
xyz:GOLD		0.01	0.04	9.47
XRP		0.01	0.03	22.09
HYPE		0.02	0.02	7.23
ETH		0.01	0.01	4.17
SOL		0.01	0.01	3.77
Panel C: Funding Rate (min/max, mean and std. dev. ≈ 0 for all six assets)				
Coin			Min	Max
SPX			-0.47	0.09
SOL			-0.43	0.02
HYPE			-0.28	0.06
XRP			-0.13	0.03
ETH			-0.10	0.02
xyz:GOLD			-0.10	0.09

Table 1 previews the paper's central empirical pattern before any regression is run: SPX

ranks worst on every full-history metric across all three panels, divergence, spread, and funding range, evidence that its reference-price fragility is a persistent structural feature rather than a single-event artifact. This is notable precisely because SPX is priced by the same eight-exchange validator median as the four majors it is being compared against; what separates it from them is the thinness of its underlying reference market, not the mechanism computing its oracle. The sample's one deployer-priced contract, xyz:GOLD, shows the mirror image: despite its deployer-set oracle construction, it sits close to the calm native cluster throughout, particularly on spread (Panel B), because the gold market underneath it is deep and continuously observed.

Table 1 (continued): Oracle Price, Open Interest, and Daily Volume. N is the total row count per asset (not the non-null count behind each individual average; open interest and daily volume are NULL for a real fraction of the panel, consistent with the sample-size gap already noted in Section 3.4.3's controls-adjusted regression). Open interest and volume are reported in the raw units returned by `hyperliquid.raw.perpetual_market_asset_contexts`.

Coin	N	Oracle Price	Open Interest		Daily Volume	
		(Mean)	Mean	Std. Dev.	Mean	Std. Dev.
SPX	668,739	0.74	11,340,426.53	5,797,555.85	12,174,814.24	8,861,764.10
HYPE	668,751	42.86	24,505,239.91	4,017,611.93	13,523,717.79	5,747,770.48
xyz:GOLD	358,670	4,544.88	29,485.09	19,151.04	539.95	772.23
SOL	668,995	130.25	4,196,889.42	1,051,662.50	4,745,937.72	2,433,681.98
XRP	668,886	1.96	81,135,991.21	27,752,398.37	72,234,888.67	62,755,724.19
ETH	669,010	2,791.38	673,034.21	198,114.30	805,435.49	438,172.87

This completes the paper's core-variable summary: open interest and daily volume both vary by one to two orders of magnitude across assets (xyz:GOLD's raw open interest and volume are far smaller than the other five, consistent with it being the newest, most thinly traded listing), and every asset's coefficient of variation on both series is large (std. dev. comparable to or exceeding the mean), consistent with the heavy-tailed, event-driven behavior documented throughout Section 3.4.

3.4 RQ1: Reference-Asset Volatility and Funding-Rate Anchoring

3.4.1 Event Identification

A scan of each asset's full history for its largest divergence event shows that all six contracts' largest deviations cluster in the same hour: October 10, 2025, 21:00 UTC. This is the well-documented $19 billion crypto liquidation cascade, the largest single-day deleveraging event

in crypto history, triggered by a surprise 100% tariff announcement on Chinese imports. Peak divergence in that hour was 15.26% for SPX, 10.54% for SOL, 7.41% for HYPE, 3.30% for XRP, and 3.00% for ETH – roughly 66 to 218 times each asset's own full-history average divergence (Table 1, Panel A). Every asset's peak-divergence minute fell in a tight 21:21–21:50 UTC window, and every peak was negative (mark price below oracle), consistent with a single, coherent, market-wide liquidation-cascade signature rather than per-asset noise.

The oracle prices themselves, not just how closely the perpetual tracked them, were also far more volatile for the thinly-referenced contract than for the deeply-referenced ones. Comparing each asset's own oracle price swing within that same hour: SPX's oracle swung −64.5%, versus −44.0% for XRP, −18.8% for SOL, −17.1% for HYPE, and −11.7% for ETH. SPX's oracle was nearly six times as volatile as ETH's, evidence that the fragility documented here originates in the reference price itself, not merely in how well the funding mechanism tracks it. Since SPX and ETH are priced by the identical validator-median mechanism, that gap cannot be attributed to oracle construction; what differs between them is the depth of the market each oracle observes. XRP's −44.0% swing is a genuine outlier in this ranking, since XRP is a deep, heavily traded asset that should sit with the calm cluster on a pure depth story, so the gradient is real but not perfectly clean.

Multi-metric confirmation across the same event: open interest fell 42–72% for every asset (a genuine market-wide deleveraging), while peak bid-ask spread was highly uneven – SPX 99.7%, XRP 22.1%, HYPE 7.2%, SOL 3.8%, ETH 1.35% – tracking each asset's reference-price fragility rather than the deleveraging itself. Funding rate moved measurably more negative across every asset during the event – SOL averaged −0.07% (minimum −0.43%), SPX averaged −0.05% (minimum −0.47%), HYPE averaged −0.03% (minimum −0.28%), with ETH and XRP moving smaller amounts in the same direction – consistent with genuine short-side funding pressure during the cascade, but nowhere near a blowout: on an ordinary trading day, funding sits at essentially 0% for every asset in the sample, both signs (Table 1, Panel C). This rules out a funding-rate blowout as the mechanism. SPX's underlying oracle price itself crashed 61% ($1.17 to $0.46) in seven minutes before whipsawing violently; the reference price collapsing faster than the perpetual could track it, not a breakdown in the funding-rate anchor itself, is the empirical driver of the divergence.

3.4.2 Full-History Structural Check

Beyond the single event, full-history dispersion statistics (standard deviation, mean, and maximum of divergence across each asset's entire trading history) confirm that SPX ranks worst on every metric across roughly 408 days of data, indicating a structural pattern rather than a one-day artifact. The sample's one deployer-priced asset, xyz:GOLD, does not repli-

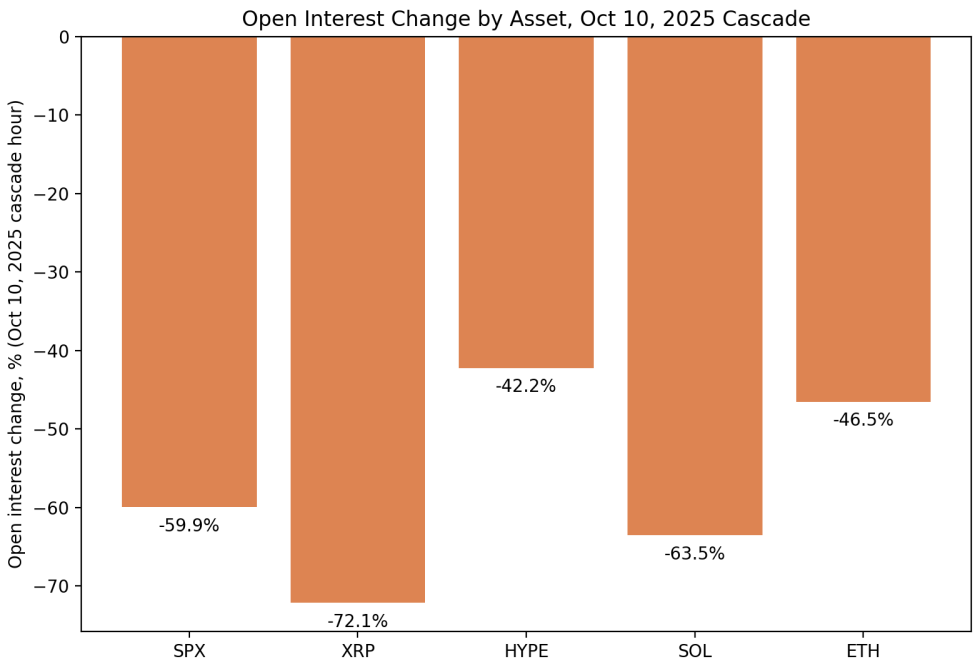

Figure 1: Open interest change by asset during the October 10, 2025 cascade hour. Every asset saw a genuine, market-wide deleveraging (42–72% decline), confirming the event's severity was not confined to the HIP-3 contracts.

cate this pattern on any metric (e.g., its full-history spread volatility is close to the calm native-asset cluster), which is itself informative: deployer-set pricing is evidently not sufficient to produce fragility when the referenced market is deep. What separates SPX from the calm native contracts is therefore not oracle construction, since it shares their mechanism. Sections 3.4.9 and 3.4.10 identify the operative difference directly.

3.4.3 Regression Results

The core specification regresses percentage divergence on oracle volatility (rolling 30-hour standard deviation of oracle returns, lagged one period), estimated server-side in Allium via Snowflake's `REGR_SLOPE`/`REGR_INTERCEPT`/`REGR_R2` functions.

Table 2: Reference-Asset Volatility and Perpetual-Oracle Divergence. This table reports estimates of

$$\text{divergence\_pct}_{i,t} = \alpha + \beta \cdot \text{OracleVol}_{i,t-1} + \gamma_1 \log(\text{Volume}_{i,t}) + \gamma_2 \log(\text{OpenInterest}_{i,t}) + \delta_i + \varepsilon_{i,t},$$

where $\text{divergence\_pct}_{i,t}$ is the percentage gap between contract i's mark price and its oracle price at time t (Section 3.2), $\text{OracleVol}_{i,t-1}$ is the rolling 30-hour standard deviation of

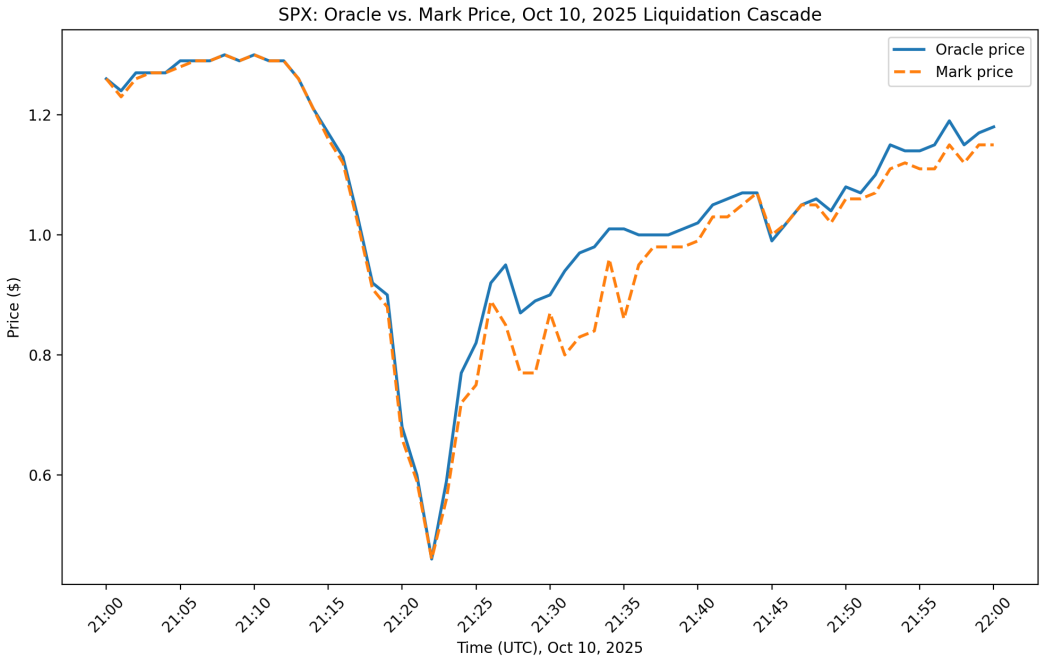

Figure 2: SPX oracle price versus mark price around the October 10, 2025 crash window. The oracle collapses and whipsaws while the perpetual's mark price lags, producing the paper's largest single divergence event.

contract i's oracle returns, lagged one period, and δ_i is a coin (asset) fixed effect, included in the columns so marked. Columns (1)-(2) use the full trading history of all six candidate contracts (ETH, SOL, XRP, HYPE, SPX, xyz:GOLD); columns (3)-(5) restrict to the October 10, 2025 crash window, the single hour in which every contract's largest full-history divergence occurred (Section 3.4.1). Column (5) adds log(Volume) and log(Open Interest) as controls, partialled out via Frisch-Waugh-Lovell. Standard errors, in parentheses below each coefficient, are clustered by coin following Petersen (2009), with the standard small-sample correction $\frac{G}{G-1} \cdot \frac{N-1}{N-K}$; column (5)'s $K = 8$ (intercept, three slopes, and $G-1 = 4$ absorbed asset dummies), reflecting the full model actually estimated rather than the FWL-residualized single-regressor form. Because the data platform does not permit row-level export, each specification's cluster-robust sandwich estimator, $\text{Var}(\hat{\beta}) = (X'X)^{-1}\left[\sum_g X_g' e_g e_g' X_g\right](X'X)^{-1}$, was implemented directly in SQL. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$, using the t-distribution with $G - 1$ degrees of freedom.

A per-asset breakdown during the Oct 10 window (untabulated) shows ETH and SOL with by far the tightest within-asset fit ($R^2 = 0.368$ and 0.402), well above SPX ($R^2 = 0.058$) – the opposite ranking from the full-history divergence-*magnitude* story, in which SPX leads. Fit and magnitude are different questions.

Column (5) flips the Oct 10 coefficient's sign again (from −17.40 to +57.14) and triples

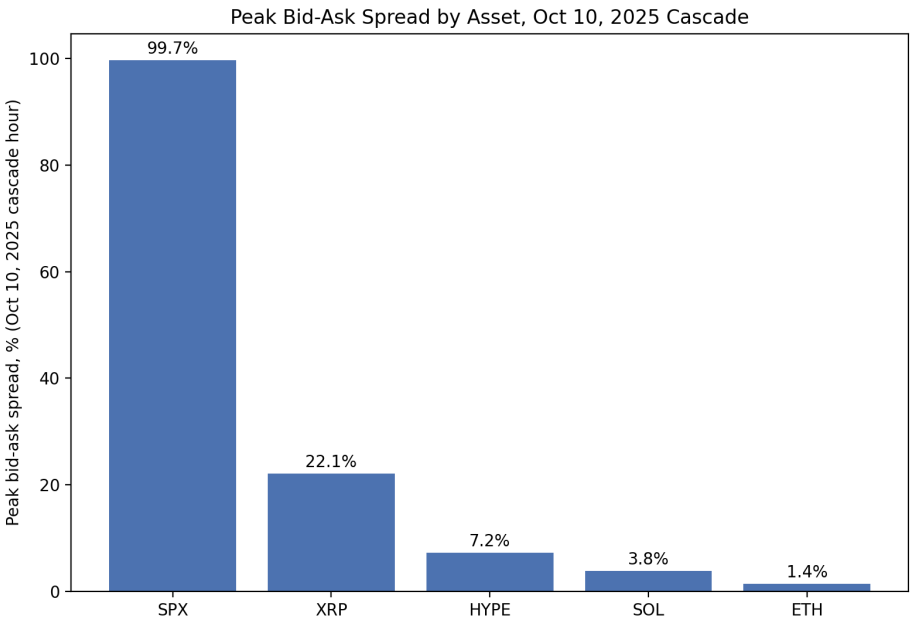

Figure 3: Peak bid-ask spread by asset during the October 10, 2025 window. SPX's spread blowout to 99.7% dwarfs every native, market-priced contract, tracking reference-price fragility rather than the market-wide deleveraging itself.

the fit relative to column (4) (R^2: 0.085 to 0.351). Three robustness checks confirm this is not an artifact: (1) the full-history sample-size drop that would otherwise accompany the controls was traced to a genuine upstream Allium data gap (the `day_base_volume` field is intermittently missing from January 5–31, 2026 across all six assets identically), which does not touch the Oct 10 window; (2) leave-one-asset-out re-estimation returned a positive slope in every one of five folds (range +46.4 to +69.3, $R^2 = 0.31$–0.47), confirming the sign flip is broad-based rather than driven by a single asset; and (3) clustering by coin, extended to this specification directly in SQL (column 5's clustered SE of 18.60, $t = 3.07$, significant at the 5% level with $G - 1 = 4$ degrees of freedom), confirms the sign flip survives the same correction applied to every other specification in this table, not an artifact of understated standard errors. Real multicollinearity between the two controls and the regressor (correlations up to 0.89) means the coefficient remains statistically significant, although its magnitude is imprecisely estimated.

3.4.4 Cluster-Robust Standard Errors

Every coefficient in Table 2 is reported with the clustered standard error described in that table's note, following Petersen (2009): observations within the same contract are not independent across time, so treating them as such understates the true standard error. The

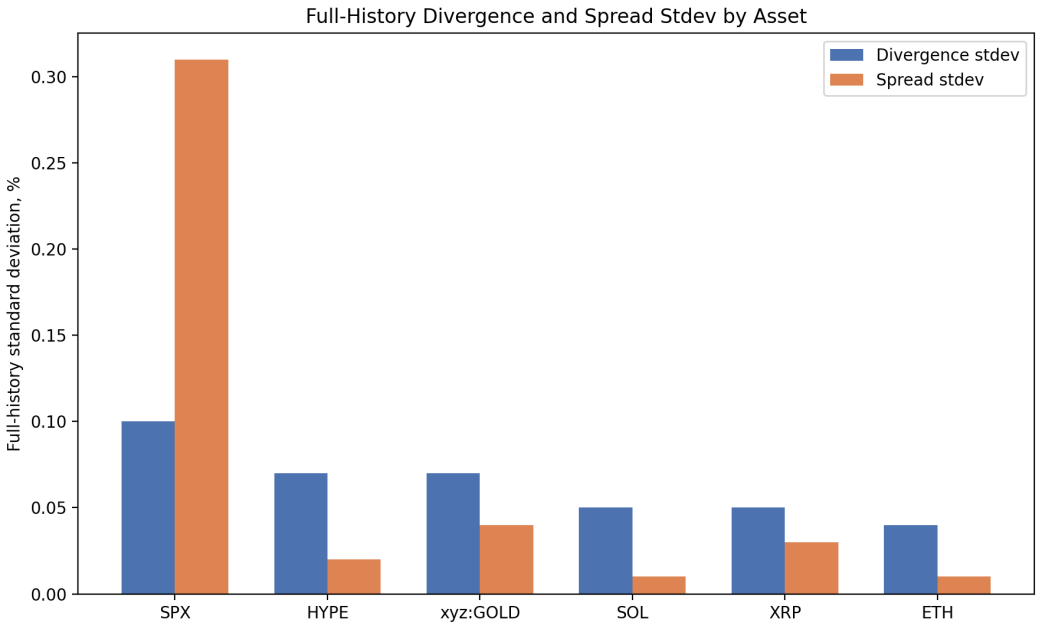

Figure 4: Full-history divergence and bid-ask spread standard deviation by asset. SPX leads on both dimensions by a wide margin despite carrying the same validator-median oracle as the calm native contracts, so its behaviour here cannot be attributed to oracle construction. Sections 3.4.9 and 3.4.10 identify what does drive the difference.

October 10, 2025 asset-fixed-effects slope (column 4) is essentially unchanged from its non-clustered estimate (-17.40 vs. -17.43), but with only five clusters available in that single-day window, small-cluster panels typically widen standard errors considerably; here the clustered t-statistic of -3.04 remains significant at the 5% level, closer to conventional 5% than 1% significance once the small number of clusters is accounted for, but not overturned. The pooled Oct 10 specification (column 3) is stronger still ($t = -3.71$, significant at 5%). The controls-adjusted specification also remains significant after clustering: its sign-flipped coefficient (+57.14) survives at the 5% level ($t = 3.07$, clustered SE 18.60), confirming the earlier leave-one-asset-out check that this result is not an artifact of understated standard errors. The full-history specifications (columns 1-2) and both RQ2 specifications in Table 10 below are also nominally significant after clustering, but their R^2 values are all under 1%, so that significance reflects the very large sample size ($n \approx 61{,}000$) rather than an economically meaningful effect, consistent with the paper's broader finding that the real signal concentrates in acute stress windows rather than full-history averages. The minute-level specifications in Table 4 also use clustered standard errors (columns 1-2; see Section 3.4.6, Additional Robustness Specifications), which sharpen rather than blur the paper's central contrast. It is not applied to the Winsorized per-coin regressions below, since each is esti-

	(1) Full Hist.	(2) Full Hist.	(3) Oct. 10, 2025	(4) Oct. 10, 2025	(5) Oct. 10, 2025
Oracle Volatility$_{t-1}$	0.044	1.721*	−15.177**	−17.398**	+57.14**
	(1.261)	(0.708)	(4.089)	(5.718)	(18.60)
log(Volume)	–	–	–	–	included
log(Open Interest)	–	–	–	–	included
N	61,485	61,485	120	120	120
R^2	0.00001	0.010	0.083	0.085	0.351
Clusters	6	6	5	5	5
Coin FE	No	Yes	No	Yes	Yes
Time FE	No	No	No	No	No
Controls (log Vol., OI)	No	No	No	No	Yes

mated on a single asset and coin-clustering does not apply there, or to the daily specification.

3.4.5 Winsorized Per-Coin Regressions

To test whether extreme outliers were suppressing a real relationship rather than driving a spurious one, each asset's series (percentage divergence and lagged oracle volatility) was Winsorized at the 1st/99th percentile before re-running the core specification separately for each coin, no pooling, no fixed effects, since each column is already a single asset.

Table 3: Winsorized Per-Coin Regressions. This table reports estimates of $\text{divergence\_pct}_{i,t} = \alpha_i + \beta_i \cdot \text{OracleVol}_{i,t-1} + \varepsilon_{i,t}$, estimated separately for each contract i (no pooling), after Winsorizing divergence_pct and OracleVol_{t-1} at their own 1st/99th percentiles. Panel A uses full trading history; Panel B restricts to the October 10, 2025 window. `xyz:GOLD` has no Panel B column because its listing (November 20, 2025) postdates the event.

Panel A: Full History	(1) ETH	(2) HYPE	(3) SOL	(4) SPX	(5) XRP	(6) xyz:GOLD
Oracle Volatility$_{t-1}$	0.644	2.374	−0.638	3.927	1.086	2.468
N	11,085	11,085	11,085	11,085	11,085	5,898
R^2	0.0015	0.0204	0.0015	0.0483	0.0037	0.0104
Panel B: October 10, 2025 Window	(1) ETH	(2) HYPE	(3) SOL	(4) SPX	(5) XRP	(6) xyz:GOLD
Oracle Volatility$_{t-1}$	−11.67	−7.33	−11.20	−20.13	−8.43	n/a
N	24	24	24	24	24	n/a
R^2	0.533	0.171	0.644	0.506	0.657	n/a

Winsorizing dramatically strengthens the Oct 10 fit for every asset relative to the un-Winsorized per-asset regressions reported in Section 3.4.3: SPX rises from $R^2 = 0.058$ to 0.506; XRP from 0.129 to 0.657; SOL from 0.402 to 0.644; ETH from 0.368 to 0.533; HYPE from 0.028 to 0.171. These results indicate that extreme outliers, not a genuine absence of relationship, were suppressing the fit in the un-Winsorized version. Full-history Winsorized fits (Panel A) remain small ($R^2 < 0.05$ throughout), consistent with the paper's broader finding that the real signal concentrates in acute stress windows.

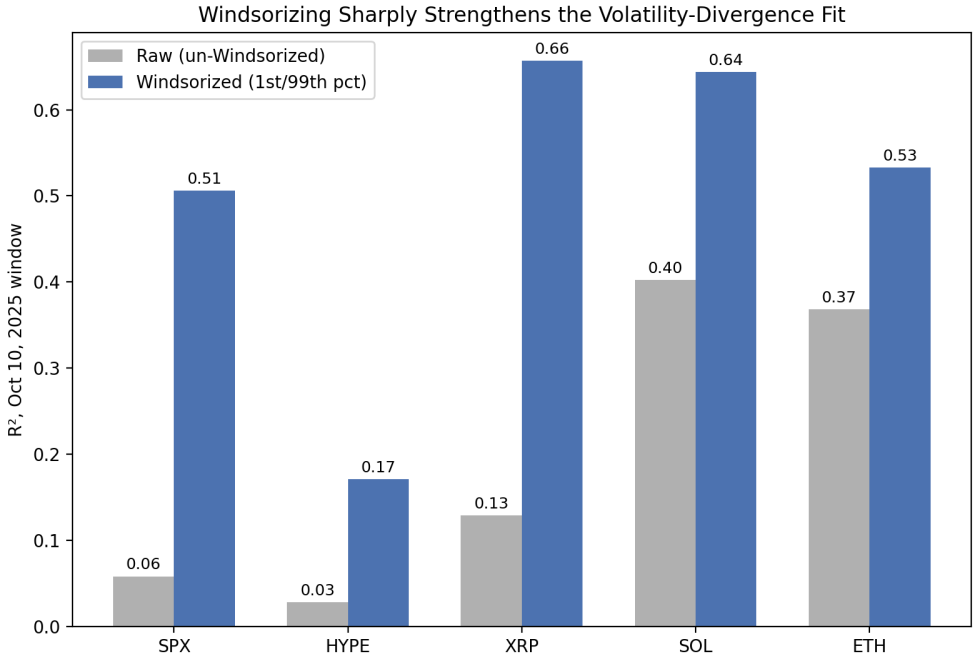

Figure 5: Oct. 10, 2025 fit (R^2) by asset, Winsorized versus raw. Every asset's fit improves sharply once the 1st/99th-percentile outliers are capped.

3.4.6 Additional Robustness Specifications

Three further checks were run: whether the core Oct 10 result survives at minute-level (rather than hourly) granularity; whether a daily-frequency specification with macroeconomic controls changes the full-history null; and whether a Volatility×Open Interest interaction term, hypothesized to carry a negative coefficient, actually does.

Table 4: Additional Robustness Specifications. Columns (1)-(2) re-estimate the core specification at minute-level granularity, no aggregation, pooled across assets, no fixed effects; standard errors for these two columns, in parentheses, are clustered by coin following Petersen (2009), using the same one-way sandwich estimator as Table 2. Column (3) aggregates to

daily frequency and adds VIX and the effective federal funds rate as separate macroeconomic controls (both from FRED), with asset fixed effects, testing whether a daily specification could recover a full-history signal the finer-grained specifications did not. Columns (4)-(5) replace the volatility regressor with $\text{OracleVol}_{i,t-1} \times \log(\text{OpenInterest}_{i,t})$, net of log(Volume) and log(Open Interest) as controls and asset fixed effects, via the same Frisch-Waugh-Lovell approach as Table 2 column (5). Columns (3)-(5) report unclustered standard errors (Section 3.4.4). $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

	(1) Minute Full Hist.	(2) Minute Oct. 10, 2025	(3) Daily Full Hist.	(4) Interaction Full Hist.	(5) Interaction Oct. 10, 2025
Oracle Volatility$_{t-1}$	-2.871	-107.29^{***}	-0.0609	–	–
	(7.27)	(10.11)			
Oracle Vol.$_{t-1} \times \log(\text{OI})$	–	–	–	-0.035	$+1.567$
N	3,648,385	7,170	1,801	27,597	120
R^2	0.00096	0.147	0.00019	0.0006	0.094
Clusters (cols 1-2, clustered)	6	5	–	–	–
Coin FE	No	No	Yes	Yes	Yes
Time FE	No	No	No	No	No
Controls (log Vol., OI)	No	No	No	Yes	Yes
VIX, Fed Funds Rate	No	No	Yes	No	No

The minute-level results confirm the full-history null is not an artifact of hourly aggregation (column 1, $R^2 = 0.00096$), while the Oct 10 window still shows a real, meaningful relationship at the finest available granularity (column 2, $R^2 = 0.147$), not directly comparable to the hourly asset-FE version's $R^2 = 0.085$ since this specification is pooled rather than asset-FE, but confirming the crash-window signal survives regardless. Clustering by coin, extended to both minute-level columns directly in SQL, sharpens this contrast rather than blurring it: the full-history slope is nowhere near significant once clustered ($t = -0.38$, only 6 clusters), consistent with a genuine null, while the Oct 10 window's slope remains overwhelmingly significant despite the small-cluster correction ($t = -10.61$, $p < 0.01$ with only 5 clusters), the strongest clustered result in the paper and further evidence the crash-window signal is not a large-sample or aggregation artifact. The daily specification (column 3) is essentially null once asset fixed effects, VIX, and the federal funds rate are all controlled for; the real signal concentrates in event windows like Oct 10, not full-history or daily averages. The interaction term (columns 4-5) returns the opposite sign from the hypothesized direction in the Oct 10 window (+1.567 rather than negative) and is weak in both windows relative to the plain controls regression ($R^2 = 0.351$, Table 2 column 5) or the Winsorized per-coin regressions above ($R^2 = 0.17$–0.66).

This sign puzzle resolves once two further interaction terms, $\text{OracleVol}_{t-1} \times$ Funding

and log(OI) × Funding, are added to the same Oct 10 specification alongside the original volatility×open-interest term: the original term collapses to essentially nothing once both new terms are included, while the funding-interaction term alone explains the large majority of the Oct 10 divergence (Table 5, Panel A). This is a plausible resolution to the wrong-sign puzzle: the original interaction term was not capturing a genuine volatility/open-interest self-correction effect, it was proxying for a funding-rate relationship the earlier specification had not isolated, and swapping log(Volume) in for log(Open Interest) reproduces essentially the same pattern, confirming the finding is not an artifact of the open-interest choice specifically.

An R^2 above 0.75 on a funding-interaction term is itself informative, however. Funding is the very mechanism a perpetual uses to pull its mark price back toward the oracle, so including it as a regressor risks mechanically inflating fit rather than reflecting an independent relationship. An R^2 that high indicates that omitting funding entirely is the more defensible specification, not merely a hedge against endogeneity. The paper's reported specification therefore drops the funding rate entirely, keeping only $\text{OracleVol}_{t-1} \times \log(\text{Volume})$ (Table 5, Panel B): this recovers a real, moderate fit ($R^2 = 0.0482$, $n = 120$, slope $+0.966$), roughly half of the volatility×open-interest interaction specification's $R^2 = 0.094$ (Table 4, column 5) and well above the near-zero fit the volatility×open-interest term shows once funding's own interaction is present. A direct multicollinearity check among the three-interaction-term specifications, paralleling the Table 2 controls check above, confirms severe collinearity in both variants (pairwise correlations ranging from -0.58 to $+0.88$), further support for reporting the simpler no-funding specification as primary rather than leaning on individually estimated coefficients from a family of specifications that cannot cleanly separate volatility, liquidity, and funding effects.

Table 5: Funding-Rate Interaction Diagnostics and the Reported No-Funding Specification. Panel A reports the three-interaction diagnostic described above, Oct. 10, 2025 window, asset fixed effects, each coefficient net of the other two interaction terms via the same Frisch-Waugh-Lovell approach as Table 2 column 5; columns (1) and (2) differ only in whether open interest or volume is paired with the volatility and funding terms. Panel B reports the paper's primary specification, funding dropped entirely.

3.4.7 A Second HIP-3 Case Study: xyz:SKHX

The paper's core sample contains only one thinly-referenced contract (SPX) and only one deployer-priced contract (xyz:GOLD), a real limitation on the six-asset design, which motivates the matched-pair and oracle-level analyses that follow. This limitation motivates a second independent event that partially addresses the identification problem, holding thin-

Panel A: Three-Interaction Diagnostic (Oct. 10, 2025 window)	(1) OI Variant	(2) Volume Variant
$\text{OracleVol}_{t-1} \times \log(\text{OI or Volume})$	−0.058	−0.053
R^2	0.0016	0.0015
$\text{OracleVol}_{t-1} \times \text{Funding}$	−14,308.9	−16,317.0
R^2	0.1392	0.1509
$\log(\text{OI or Volume}) \times \text{Funding}$	+72.61	+69.89
R^2	0.7794	0.7583
N	120	120

Panel B: Reported Specification, Funding Dropped	(3) $\text{OracleVol}_{t-1} \times \log(\text{Volume})$
Coefficient	+0.966
R^2	0.0482
N	120

ness fixed while flipping the oracle mechanism. On July 27, 2026, around 23:00 UTC, the price feed behind SK Hynix's HIP-3 perpetual, `xyz:SKHX`, not part of this paper's six-asset core sample, imported a real but thin-liquidity trade from Korea's pre-market session and printed a sharp, short-lived collapse, first reported publicly by Allium's own research team (Shehdula, 2026a).

Querying Allium's primary `hyperliquid.raw.perpetual_market_asset_contexts` table directly, the same source used throughout this paper, independently confirms the event rather than relying on the secondary writeup alone. Over the roughly 90-minute window surrounding the event (22:30–00:00 UTC), the oracle price fell to a minimum of $872.25, exactly matching the trough separately reported by Allium's research team, before recovering. Mark-oracle divergence swung from −11.33% to +14.63% within this window, a range comparable in magnitude to SPX's 15.3% peak divergence during the October 10, 2025 cascade, despite involving a different underlying asset entirely, a different calendar month, and the opposite oracle mechanism. `xyz:SKHX` is deployer-priced where SPX is validator-priced. This is reported as a single case study rather than as evidence for a mechanism; Sections 3.4.9 and 3.4.10 test the mechanism directly on the full sample. What the episode does establish is that an imported reference price can dislocate a live, leverage-bearing market severely and within minutes. Open interest fell from a pre-event baseline of approximately 428,068 (averaged over the two hours preceding the window) to a low of 345,166, a roughly 19.4% decline consistent with the mass liquidation the event triggered. Funding stayed within a

narrow band throughout (−0.00076 to +0.00688), the same pattern already documented for the Oct 10 cascade: the failure mode is a reference-price and liquidity breakdown, not a funding-mechanism blowout.

This is a second, independent instance of a HIP-3 deployer-priced contract exhibiting the same structural signature as this paper's central finding: an externally sourced, deployer-controlled reference price that can move sharply on thin, sometimes questionable underlying liquidity, producing outsized divergence and open-interest stress that a market-priced native perpetual's continuously observed, multi-exchange oracle does not experience to the same degree. Because `xyz:SKHX` was not part of the original six-asset sample and this is a single asset on a single day, it is presented as corroborating evidence alongside the paper's formal regression results, not as a third regression specification; a full event-study treatment matching Section 3.4.1's Oct 10 analysis is a natural extension for future work (Section 4).

The parallel is not exact, however, and the difference is itself informative. Unlike SPX's Oct 10 bid-ask spread, which blew out to 91–99%, `xyz:SKHX`'s spread barely widened in absolute terms: it averaged 0.0225% during the event window versus a 0.0181% pre-event baseline, peaking at 0.19%, roughly ten times the baseline in relative terms but still a fraction of SPX's collapse. The reference-price and open-interest fragility this paper documents does not automatically imply a liquidity collapse of the same magnitude on every thinly-referenced asset; `xyz:SKHX`'s order book stayed comparatively orderly even as its oracle and mark price diverged sharply, a reminder that reference-price fragility and order-book fragility are related but separable failure modes.

3.4.8 Matched-Pair Identification: The Same Asset Under Two Oracle Mechanisms

The analysis to this point compares contracts that differ in oracle construction *and* in what they reference: native contracts track major cryptocurrencies, while HIP-3 contracts track equities, commodities, and thinly traded tokens. That confound is unavoidable in the six-asset core sample. This section removes it, and in doing so reaches a more limited conclusion than the earlier framing claimed.

A subset of Hyperliquid's HIP-3 markets list an underlying that is *also* traded as a native perpetual on the same exchange, permitting a matched pair: two contracts on the identical underlying, same venue, same margin currency, same hour, differing in how the reference price is constructed. Enumerating all HIP-3 sub-dexes through Hyperliquid's public `perpDexs` and `metaAndAssetCtxs` endpoints yields 278 HIP-3 markets across ten deployers, of which 26 share a ticker with a native perpetual. Requiring at least 400 overlapping hourly observations and requiring the two contracts' median prices to agree within 20%

(a validation gate that rejected `para:STX` and `flx:GAS`, whose price levels differ by orders of magnitude and which therefore do not reference the same asset) leaves **20 validated matched pairs**, roughly 95,000 pair-hours between February and September 2026. Because Hyperliquid's API serves live data, all figures below are as of September 2, 2026; re-running `matched_pair_analysis.py` on a later date will shift them slightly as new hours enter each window. For each pair,

$$b_{i,t} \;=\; \frac{P_{i,t}^{\text{HIP-3}} - P_{i,t}^{\text{native}}}{P_{i,t}^{\text{native}}} \times 100,$$

the hourly percentage gap between the two contracts' closing prices, which nets out the underlying's own price path. Stress hours are those in which the native leg's absolute log return exceeds two standard deviations of its own distribution. Liquidity is measured in-sample, as mean hourly notional volume over the pair's own overlapping window, rather than from a point-in-time snapshot. All $|b|$ figures are Winsorized at the 1st and 99th percentiles.

Table 6: Matched-Pair Dislocation, Deployer-Priced versus Validator-Priced Contracts on the Identical Underlying. Ratio is mean $|b|$ in stress hours over mean $|b|$ in calm hours. Volumes are in-sample mean hourly notional, thousands of dollars. The final column reports the share of sample hours in which the HIP-3 leg traded at all.

HIP-3 contract	Underlying	N	Med. $\lvert b\rvert$	Ratio	Native \$k/hr	HIP-3 \$k/hr	Hrs traded
`hyna:XPL`	XPL	4,938	0.6521	1.83	469	3.8	45%
`hyna:ADA`	ADA	4,940	0.5664	1.92	238	1.2	30%
`hyna:BCH`	BCH	3,813	0.5640	3.96	160	1.6	22%
`hyna:IP`	IP	3,836	0.5568	2.02	54	3.3	40%
`hyna:XMR`	XMR	3,821	0.5437	1.75	267	2.4	40%
`hyna:ENA`	ENA	4,988	0.5295	2.17	423	2.6	45%
`flx:XMR`	XMR	3,216	0.5267	1.95	282	1.0	43%
`hyna:LTC`	LTC	4,962	0.4768	1.69	123	1.2	18%
`hyna:SUI`	SUI	4,938	0.3816	1.57	669	3.4	48%
`hyna:FARTCOIN`	FARTCOIN	4,947	0.3200	1.36	706	6.8	69%
`hyna:LINK`	LINK	4,942	0.3064	1.77	266	3.7	48%
`hyna:PUMP`	PUMP	4,986	0.3008	1.83	995	4.1	70%
`hyna:DOGE`	DOGE	4,963	0.2768	1.49	644	2.8	48%
`hyna:ZEC`	ZEC	4,968	0.2465	1.18	5,535	10.0	79%
`hyna:XRP`	XRP	4,966	0.1522	1.65	2,174	6.7	64%
`hyna:BNB`	BNB	4,941	0.1370	1.22	310	6.3	61%
`hyna:SOL`	SOL	4,989	0.0940	0.91	10,509	33.5	88%
`hyna:HYPE`	HYPE	4,988	0.0885	1.05	17,886	66.2	94%
`hyna:ETH`	ETH	4,991	0.0531	0.88	39,524	125.2	95%
`hyna:BTC`	BTC	4,993	0.0422	0.93	99,605	268.8	96%

A persistent level effect. In 19 of the 20 pairs the venue-deployed contract sits further from its native twin than a native contract typically sits from its own oracle (Table 1, Panel A: roughly 0.02–0.05%). The median across pairs is 0.313%, an order of magnitude wider, ranging from 0.042% for BTC to 0.652% for XPL. Under a sign test this is significant at $p = 2.0 \times 10^{-5}$. Whatever its source, the dislocation is real, large, and present on ordinary trading days rather than only under stress.

The magnitude is almost entirely a liquidity phenomenon. Regressing the log median dislocation on log in-sample volume gives $R^2 = 0.799$ using the underlying's native volume ($\beta = -0.853$, SE 0.101, $t = -8.45$) and $R^2 = 0.865$ using the HIP-3 contract's own volume ($\beta = -1.159$, SE 0.108, $t = -10.74$). Liquidity, measured either way, explains roughly four fifths of the cross-sectional variation in how far these contracts sit from fair value.

The two channels cannot be separated in this sample, and this is the section's binding limitation. The two volume measures are correlated at 0.930. Entered together, the HIP-3 contract's own volume retains significance ($t = -3.08$) while the underlying's depth does not ($t = -0.89$); for the stress ratio neither survives ($t = -0.61$ and $t = -0.74$). With $n = 20$ and regressors this collinear, the design cannot establish that deployer-set oracle construction contributes anything beyond the fact that these contracts trade thinly. Table 6's final column makes the concern concrete: the HIP-3 legs trade in as few as 18% of sample hours, and a sparsely traded contract's last price will drift from a continuously traded one's regardless of how either oracle is built. Settling whether the oracle mechanism adds anything on top of that requires comparing each leg's mark price against its *own* oracle rather than against the other leg's price. Section 3.4.9 does exactly that, using oracle-level data, and finds that it does not.

Stress amplification is confined to thin pairs. The five pairs whose underlying trades above \$5M/hour (BTC, ETH, SOL, HYPE, ZEC) show ratios of 0.88–1.18, that is, no amplification whatever, while the fifteen thinner pairs range from 1.22 to 3.96. Bivariately the ratio falls with underlying depth ($\beta = -0.491$, SE 0.136, $t = -3.61$, $R^2 = 0.419$), though as noted the multivariate version cannot attribute this to either channel specifically. This does qualify the paper's central RQ1 claim: fragility does not intensify under stress uniformly, but only where the contract and its reference market are thin.

One robustness check yields a consistent result. Monero is listed by two independent deployers, and the two give nearly identical results (`hyna:XMR` median 0.5437%, ratio 1.75; `flx:XMR` median 0.5267%, ratio 1.95), so the pattern is not an artifact of a single deployer's implementation.

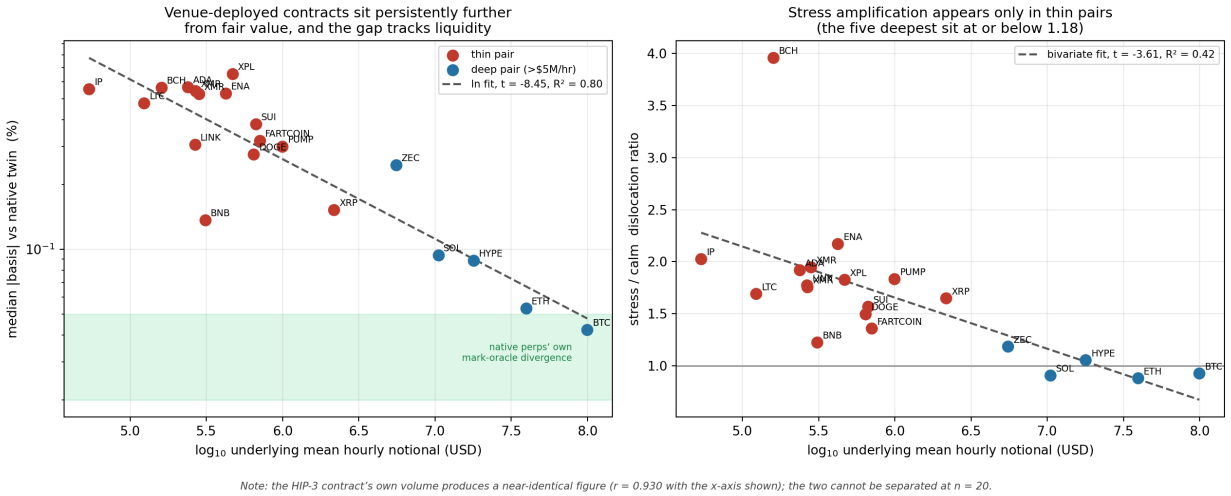

Figure 6: Matched-pair results across 20 validated pairs. Left: median dislocation from the native twin falls sharply with liquidity; the shaded band marks the range within which native contracts track their own oracles. Right: stress amplification is present only for thin pairs, with the deepest sitting at or below 1.0. Both panels are plotted against the underlying's volume; the HIP-3 contract's own volume produces a visually near-identical figure, which is precisely the identification problem discussed in the text.

3.4.9 Decomposing the Gap: Is It the Oracle, or Is It Thin Trading?

Section 3.4.8 established that venue-deployed contracts sit persistently further from fair value than their native twins, and that the magnitude tracks liquidity, but could not attribute that gap to oracle construction rather than to the contracts' own sparse trading. Because Allium's `perpetual_market_asset_contexts` table carries the oracle price alongside the mark price for every contract, that attribution can be settled directly. The contract-to-contract gap decomposes exactly:

$$\underbrace{P^H - P^N}_{\text{observed gap}} = \underbrace{(P^H - O^H)}_{A\text{: HIP-3 tracking error}} + \underbrace{(O^H - O^N)}_{B\text{: \textbf{oracle disagreement}}} + \underbrace{(O^N - P^N)}_{C\text{: native tracking error}},$$

where P denotes mark price and O denotes oracle price. Component B is the quantity of interest. Oracle prices are published continuously whether or not a single contract trades, so B is entirely immune to the sparse-trading confound that limits the previous section. If deployer-set construction is genuinely inferior, B must be large. Table 7 reports all three components at hourly frequency over 5,136 hours per pair, February to September 2026.

Table 7: Decomposition of the Matched-Pair Gap into Oracle and Tracking Components. Medians, in percent. "Share from oracle" is median B over median total gap. The final three columns report the frequency with which oracle disagreement exceeds 0.5% and 2%, and the ratio of mean B in stress hours (native leg's absolute log return above two standard deviations) to mean B in calm hours.

HIP-3 contract	Underlying	Med. B (oracle)	Med. A (HIP-3)	Med. C (native)	Share from oracle	%hrs $B > 0.5$	%hrs $B > 2$	Stress ratio
`hyna:BTC`	BTC	0.00095	0.02864	0.04651	4.4%	0.00%	0.00%	2.40
`hyna:BNB`	BNB	0.00096	0.02367	0.04659	3.5%	0.43%	0.00%	0.65
`hyna:LTC`	LTC	0.00136	0.03450	0.05468	4.8%	0.49%	0.00%	0.89
`hyna:ETH`	ETH	0.00145	0.02681	0.04648	6.3%	0.06%	0.00%	2.29
`hyna:SOL`	SOL	0.00160	0.02345	0.05173	4.5%	0.08%	0.00%	2.16
`hyna:DOGE`	DOGE	0.00162	0.02925	0.04998	5.6%	0.43%	0.12%	0.70
`hyna:LINK`	LINK	0.00171	0.02765	0.04389	6.1%	0.66%	0.16%	0.47
`hyna:XRP`	XRP	0.00179	0.02373	0.05525	5.4%	0.55%	0.12%	0.60
`hyna:SUI`	SUI	0.00198	0.02802	0.05495	6.0%	0.66%	0.14%	1.72
`hyna:ADA`	ADA	0.00202	0.04028	0.06615	5.2%	0.76%	0.12%	0.87
`hyna:BCH`	BCH	0.00213	0.06757	0.06492	4.2%	23.06%	23.02%	0.88
`hyna:HYPE`	HYPE	0.00237	0.03245	0.03456	7.4%	0.08%	0.00%	1.78
`hyna:ENA`	ENA	0.00295	0.03464	0.05765	6.5%	0.16%	0.02%	1.83
`hyna:ZEC`	ZEC	0.00303	0.02588	0.04065	9.1%	0.45%	0.27%	0.89
`hyna:FARTCOIN`	FARTCOIN	0.00303	0.03732	0.04474	9.9%	0.95%	0.58%	0.57
`hyna:XPL`	XPL	0.00323	0.05189	0.05527	6.5%	1.15%	0.92%	0.67
`hyna:XMR`	XMR	0.00339	0.08315	0.06662	3.8%	22.28%	20.57%	1.32
`hyna:PUMP`	PUMP	0.00351	0.03369	0.06192	7.4%	0.20%	0.04%	1.59
`hyna:IP`	IP	0.00440	0.10997	0.13049	6.4%	29.87%	29.21%	0.01
`flx:XMR`	XMR	0.08517	0.07164	0.06662	61.1%	34.59%	32.70%	1.01

The gap is not an oracle effect. For 19 of the 20 pairs the two oracles agree to within 0.005% at the median, and oracle disagreement accounts for only 3.4% to 9.9% of the observed contract-to-contract gap. For BTC the median oracle disagreement is 0.00095%, roughly one part in a hundred thousand. Component A, the HIP-3 contract's failure to track its *own* oracle, is ten to thirty times larger and carries the overwhelming majority of the gap. Read against Section 3.4.8's finding that the gap scales with liquidity, the dislocation documented there is best explained as a stale-price phenomenon on sparsely traded contracts, not evidence that deployer-set reference construction is inferior once the confound is removed.

Nor does oracle disagreement widen under stress. The stress ratios in Table 7's final column cluster at or below unity, with a median of 0.89. The three largest ratios

(BTC 2.40, ETH 2.29, SOL 2.16) sit on pairs whose absolute oracle gaps are approximately 0.004%, economically meaningless in either regime. This provides little support for stress-driven oracle deterioration as a mechanism.

What the oracle data does show is a sharp dispersion in deployer quality. Deployer oracles do not vary along a spectrum; they fall into two clearly separated groups. Sixteen of the twenty mirror the validator median almost exactly, exceeding 0.5% disagreement in under 1.2% of hours and never once exceeding 10%; `hyna:BTC` exceeds 0.5% in zero of 5,135 hours. The remaining four (`flx:XMR`, `hyna:XMR`, `hyna:BCH`, `hyna:IP`) disagree by more than 2% in 20% to 33% of *all* hours, and three of them exceed 10% disagreement in 11% to 20% of hours. These are not transient dislocations during stress events but persistent, structural mispricings on live, leverage-bearing markets.

The failures are asset-driven, not deployer-driven. Monero is listed by two unrelated deployers, and both fail on it: `flx:XMR` disagrees by over 2% in 32.7% of hours and `hyna:XMR` in 20.6%, while the same `hyna` deployer tracks BTC to within 0.001% and never once breaches 0.5%. The same deployer that tracks Bitcoin almost exactly is far less accurate on Monero. This points to the observability of the underlying spot market rather than deployer competence as the binding constraint, though this paper does not test that mechanism directly and the inference should be read as a hypothesis for future work rather than an established result.

Implication. The regulatory concern survives, but relocated. The risk in a venue-deployed perpetual is not that its reference price is systematically worse. It is that reference-price quality varies enormously and invisibly across contracts on the same venue, with no interface-level signal distinguishing them. A trader holding `hyna:BTC` and one holding `hyna:XMR` face products whose reference prices differ in reliability by more than an order of magnitude, presented identically. For a regulator weighing approval of listed perpetual futures, the question suggested by this evidence is not who computes the reference price, but whether per-contract reference-price quality is disclosed, monitored, and bounded.

3.4.10 Oracle Infrastructure Quality Across Deployers

The decomposition in Section 3.4.9 shows that deployer-set oracles are, at the median, nearly indistinguishable from the validator median. That result is about central tendency, and it invites a different question: are all deployer oracles alike, or does the category conceal wide variation? The HIP-3 universe permits a direct test, because several underlyings are listed by multiple independent deployers simultaneously. Gold carries six listings and the S&P 500 five, of which four in each family produce enough price movement to be verified as tracking the stated underlying (Section 3.4.11). Comparing those feeds against *each other* isolates

infrastructure quality with no native benchmark required and no trading-activity confound, since oracle prices publish irrespective of whether the contract trades.

Two measures are used. The first is agreement, computed as the root-mean-square difference between two feeds' hourly log returns, which is scale-free and therefore valid across deployers using different contract notionals. The second is liveness, the share of consecutive minute observations in which a feed's oracle price does not change at all, reported alongside the market's traded volume and open interest so that dormant listings can be separated from live ones.

Agreement varies by more than an order of magnitude, including on Bitcoin. Against the validator median, `hyna:BTC` returns an RMS return difference of 0.01% and a return correlation of 0.9996. On the same asset over the same window, `flx:BTC` returns 0.20% with a correlation of 0.823 and a maximum hourly divergence of 3.22%, and `cash:BTC` returns 0.33%. Bitcoin is the most continuously observed asset in the sample, so no thinness of the underlying can account for a twenty- to thirty-fold spread in feed accuracy. Among the six gold deployers, pairwise RMS differences range from exactly 0.00%, indicating deployers sharing a single upstream feed, to 0.24% with maxima near 2.9%.

Liveness separates the deployers more sharply still, and crypto provides a clean control. Because crypto underlyings trade continuously, a stale crypto feed cannot be attributed to a closed market. Native BTC, ETH and SOL oracles are unchanged in 4.8%, 5.3% and 6.3% of consecutive observations respectively, with longest unchanged runs of nine to seventeen observations. The best HIP-3 deployer matches this closely: `hyna:BTC` 5.7%, `hyna:ETH` 6.9%, `hyna:SOL` 6.9%. Deployer-set construction is therefore fully capable of validator-median liveness, which is worth stating as directly as the failures.

The substantive result is dispersion among deployers of the same asset. Table 8 reports the S&P 500 and gold listings. Frozen shares across the four verified S&P 500 listings run 21.1%, 31.5%, 36.9% and 55.6%, and across the four verified gold listings 25.9%, 38.7%, 50.3% and 55.2%. Every deployer of a given asset faces the identical trading calendar, so the calendar cannot explain a spread of that size. The economically relevant case is `km:US500`, which carries roughly $5.7M in daily notional volume and up to 20,031 in open interest on an oracle frozen 55.6% of the time, while `xyz:SP500` tracks the same index at 21.1%. A trader choosing between two S&P 500 perpetuals on one venue faces a materially different reference-price update regime with nothing at the interface to indicate it.

Table 8: Oracle Liveness by Deployer, Same Underlying. Frozen share is the percentage of consecutive minute observations in which the oracle price is unchanged. Volume is average daily notional; open interest is the sample maximum. Rows marked (unver.) have

oracles that never update and therefore generate no return series, so their stated underlying rests on the ticker alone and cannot be confirmed (Section 3.4.11); they are excluded from every verified count in the text. Dashes denote listings whose observation window differs materially from the others, reported in the text rather than compared directly here.

Underlying	Contract	Operator	Frozen	Avg daily notional	Max OI
S&P 500	`xyz:SP500`	xyz	21.06%	—	—
S&P 500	`cash:USA500`	cash	31.46%	23,314,999	7,854
S&P 500	`mkts:US500`	mkts	36.90%	—	—
S&P 500	`km:US500`	km	55.55%	5,715,142	20,031
S&P 500 (unver.)	`abcd:USA500`	abcd	100.00%	0	0
Gold	`xyz:GOLD`	xyz	25.92%	61,099,565	89,339
Gold	`cash:GOLD`	cash	38.66%	1,217,722	1,602
Gold	`flx:GOLD`	flx	50.32%	342,388	692
Gold	`km:GOLD`	km	55.21%	229,242	740
Gold (unver.)	`hyna:GOLD`	hyna	100.00%	0	0
Gold (unver.)	`mkts:GOLD`	mkts	100.00%	0	0
Bitcoin	BTC (native)	validator	4.82%	2,376,262,430	45,372
Bitcoin	`hyna:BTC`	hyna	5.68%	5,381,275	677
Bitcoin	`flx:BTC`	flx	41.16%	0	0
Bitcoin (unver.)	`cash:BTC`	cash	100.00%	0	0
Ether	ETH (native)	validator	5.26%	932,858,938	1,088,624
Ether	`hyna:ETH`	hyna	6.93%	2,544,316	3,462
Ether (unver.)	`cash:ETH`	cash	100.00%	0	0
Solana	SOL (native)	validator	6.32%	247,645,498	7,181,031
Solana	`hyna:SOL`	hyna	6.94%	755,645	40,102

Fully frozen oracles exist, are confined to dormant markets, and cannot be verified at all. Five feeds in the sample never update across the observation window: `cash:BTC` across 250,849 consecutive observations, `cash:ETH` across 232,112, `abcd:USA500` across 267,362, and `hyna:GOLD` and `mkts:GOLD`. In every case both traded volume and open interest are exactly zero, so these are listed-but-never-launched markets and no trader is exposed to them. They carry a second, subtler property worth stating: because their prices never move, they generate no return series, and so *their identity cannot be confirmed from price behavior at all*. That `abcd:USA500` tracks the S&P 500 or that `hyna:GOLD` tracks gold is an inference from the ticker string and nothing else. This paper therefore excludes them from every verified count and reports them only as dormant listings. The finding is that permissionless listing accumulates abandoned and unverifiable infrastructure, not that live markets run on constant prices.

Two limits bound these results. First, equity and commodity underlyings do not

trade continuously, so a materially non-zero frozen share is expected for those feeds and an absolute liveness figure is not by itself evidence of neglect; only the cross-deployer spread on a single asset is interpretable, which is why the comparison is constructed that way. Second, listing windows differ slightly across deployers (`km:US500` ends in mid-June, `xyz:SP500` begins in mid-March), so level comparisons should be read as indicative rather than exact.

Taken with Section 3.4.9, the picture is coherent. Deployer-set reference pricing is not inherently inferior, and at its best is indistinguishable from the protocol's own validator median. What it introduces is variance: across deployers listing the same asset on the same venue, reference-price agreement differs by more than an order of magnitude and update frequency by a factor of roughly two and a half, with no disclosure distinguishing them. The exposure a trader takes on is not the pricing mechanism but the identity of whoever operates it, which is precisely the attribute the product does not surface.

3.4.11 Asset-Identity Verification

Every contract examined in this paper is identified by a ticker string chosen by whoever deployed it. Ticker strings are not evidence. A contract named `SPX` need not track the S&P 500, a contract named `km:US500` need not track the same index as one named `xyz:SP500`, and two contracts named `GOLD` on different sub-dexes need not track the same metal. Because deployer-chosen naming is unverified by the protocol, and because permissionless listing means no party audits it, this section states the identity evidence for every asset claim made above rather than resting on names.

The test is behavioral. Contracts tracking the same underlying must display near-unit correlation and near-unit regression slope in their hourly oracle returns, computed only over hours in which both feeds actually moved so that stale observations cannot manufacture agreement. Table 9 reports the results.

Table 9: Return-Based Identity Verification. Pairwise hourly oracle-return correlation and slope for contracts claimed to track a common underlying, restricted to hours in which both feeds moved.

Underlying	Pair	N	Correlation	Slope
Bitcoin	BTC vs `hyna:BTC`	4,456	0.9999	0.9999
Bitcoin	BTC vs `flx:BTC`	2,731	0.9998	0.9994
Bitcoin	`flx:BTC` vs `hyna:BTC`	2,731	0.9998	1.0001
Monero	XMR vs `hyna:XMR`	3,280	0.9999	1.0003
Monero	XMR vs `flx:XMR`	2,660	0.9991	0.9897
Monero	`flx:XMR` vs `hyna:XMR`	2,660	0.9991	1.0086
Gold	`cash:GOLD` vs `xyz:GOLD`	3,088	0.9918	0.9994
Gold	`cash:GOLD` vs `km:GOLD`	2,514	0.9814	0.9823
Gold	`km:GOLD` vs `xyz:GOLD`	2,512	0.9781	0.9842
Gold	`cash:GOLD` vs `flx:GOLD`	2,656	0.9299	1.0004
Gold	`flx:GOLD` vs `xyz:GOLD`	2,654	0.9274	0.8685
Gold	`flx:GOLD` vs `km:GOLD`	2,513	0.9210	0.8563
S&P 500	`km:US500` vs `xyz:SP500`	2,183	0.9757	0.9907
S&P 500	`cash:USA500` vs `xyz:SP500`	2,673	0.9677	0.9989
S&P 500	`cash:USA500` vs `km:US500`	2,602	0.9593	0.9749
S&P 500	`mkts:US500` vs `xyz:SP500`	1,516	0.9447	0.9550
S&P 500	`cash:USA500` vs `mkts:US500`	148	0.8122	0.8282

Three points follow.

The scale difference across S&P 500 deployers is a contract-size convention, not a different index. `km:US500` quotes near 750 while `xyz:SP500` quotes near 7,667, a tenfold gap that could plausibly indicate an entirely different underlying. The correlation of 0.9757 at a slope of 0.9907 confirms both contracts track the same index, differing only in contract-size convention.

The matched pairs of Section 3.4.8 carry an independent identity gate. Each pair was required to show median prices agreeing within 20%, a screen that rejected `para:STX` (a price gap of roughly 610,000%) and `flx:GAS` (86%), both of which share a ticker with a native perpetual while referencing something else entirely. Ticker collisions are therefore not hypothetical on this venue; they are present and were removed by test rather than by inspection.

Five listings cannot be verified by any price-based method, and are excluded from every count. `cash:BTC`, `cash:ETH`, `abcd:USA500`, `hyna:GOLD` and `mkts:GOLD` have oracles that never update, so they generate no return series against which identity could be checked. Their stated underlyings rest on the ticker string alone. This paper counts four verified gold listings and four verified S&P 500 listings rather than six and five, and describes the remainder as dormant listings of unconfirmed reference. The observation generalizes: on a permissionless venue, a contract whose reference price never moves cannot be shown to track what it claims to track, and the absence of movement is precisely what removes the

means of checking.

3.5 RQ2: Settlement-Stablecoin Volatility Spillover

3.5.1 Full-History and Extreme-Event Results

Using the same percentage-divergence measure, with USDC's deviation from \$1 as the predictor, a full-history correlation is essentially zero everywhere (≤ 0.04), and the corresponding regression is a clean, doubly null result: R^2 near zero in both pooled and asset-fixed-effects specifications, which agree closely with each other. Everyday USDC fluctuations do not measurably move Hyperliquid perpetual markets. The full-history result is null and remains so under cluster-robust standard errors.

Table 10: Settlement-Stablecoin Deviation and Perpetual-Oracle Divergence. This table reports estimates of

$$\text{divergence\_pct}_{i,t} = \alpha + \beta \cdot \text{USDCDeviation}_t + \delta_i + \varepsilon_{i,t},$$

where USDCDeviation_t is USDC's price deviation from its \$1 peg at time t (Section 3.1), common across all six contracts, and δ_i is a coin fixed effect, included in column (2). Both columns use the full trading history. Standard errors, in parentheses, are clustered by coin following Petersen (2009). $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$, using the t-distribution with $G - 1 = 5$ degrees of freedom.

	(1) Full History, Pooled	(2) Full History, Coin FE
USDC Deviation	-0.279^{*}	-0.290^{**}
	(0.101)	(0.102)
N	61,491	61,491
R^2	0.0004	0.0005
Clusters	6	6
Coin FE	No	Yes
Time FE	No	No

Both coefficients are statistically significant only because of the very large sample size ($n \approx 61{,}000$); the R^2 values are effectively zero, so this significance is not economically meaningful. RQ2's real evidence rests on the event-driven results below (the 2-SD extreme-USDC-hours finding and the March 11, 2023 Deribit event study), not this full-history regression.

Restricting to hours where USDC's deviation from peg exceeds two standard deviations of its own distribution surfaces a more interesting pattern: mean divergence barely moves, but divergence *volatility* rises 1.75 to 3.8 times for five of the six assets (SPX most strongly, at 3.8x), with xyz:GOLD the one clean exception, essentially unaffected. This is the strongest RQ2 evidence within the Hyperliquid sample, and its correlation-scale summary statistics largely confirm the earlier finding, again with SPX as the interesting exception on the mean-divergence dimension specifically – its average divergence during 2-SD hours is actually *lower* than during normal hours (0.060 vs. 0.072), despite higher dispersion.

3.5.2 March 11, 2023 Event Study

Because Hyperliquid's history does not reach March 2023, the USDC/SVB depeg is studied through the external cross-exchange sources described above. The clearest single result comes from Deribit, where the USDC-margined `BTC_USDC-PERPETUAL` is compared directly against the BTC-margined `BTC-PERPETUAL` on the same exchange: divergence between the two peaks at 15.56% at 06:00 UTC and tracks USDC's own real-time depeg almost exactly (correlation 0.775 against $1/\text{usdc\_price} - 1$; mean absolute residual 1.81 percentage points).

Liquidity on the USDC-margined leg collapsed during the worst hours (from a ∼180 BTC/hour calm baseline to 6–16 BTC/hour). Because this comparison holds the exchange and the underlying asset fixed and varies only the settlement currency, it is close to a mechanical demonstration that quoting an asset in a depegged currency inflates its quoted price by roughly the depeg's own size, and is treated as the paper's central same-exchange RQ2 evidence. A companion Bybit comparison (USDT-margined BTCUSDT vs. USDC-margined BTCPERP) during the October 10, 2025 cascade shows the same directional effect at much smaller magnitude (∼0.05% baseline gap widening to 0.9–1.0% under stress), consistent with BTC's far greater liquidity depth relative to the thinner reference assets that anchor the paper's main results; this gap was confirmed to revert to baseline within a normal trading day once the cascade passed.

3.6 Robustness and Data Limitations

Several data-quality checks support the results above. The `day_base_volume` field's January 2026 gap (Section 3.4.3) is a confirmed upstream Allium pipeline issue affecting all six assets identically and does not touch the paper's headline October 2025 results. The full-controls regression's sign flip was independently confirmed via leave-one-asset-out re-estimation. The March 2023 event study's Deribit finding was independently cross-checked against an index-methodology concern (Deribit's BTC-USDC index switched from a market-price-derived

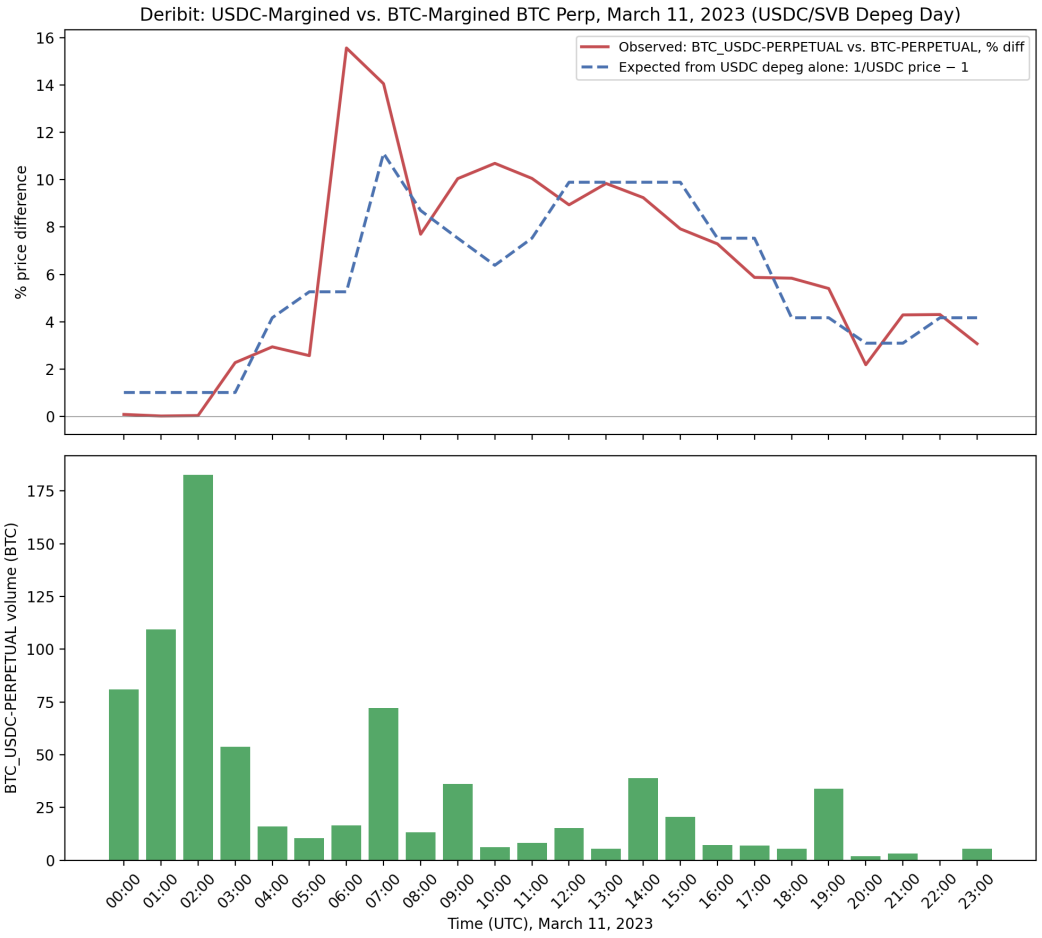

Figure 7: Deribit's USDC-margined vs. BTC-margined BTC perpetual around the March 11, 2023 USDC/SVB depeg. The two contracts hold the exchange and underlying asset fixed and vary only the settlement currency, isolating the depeg's mechanical effect on quoted price.

calculation to a hard peg to BTC-USD in July 2025, more than two years after the event studied here, ruling out a pricing-convention artifact) and against BitMEX's independent order-book data, which confirms the same crash timing from a third, unrelated venue.

4 Conclusion

This paper set out to answer two questions about how a perpetual future's reference price is constructed and what that construction means for market quality under stress. The first, and stronger, finding is that reference-price design matters a great deal once markets are actually stressed, but not along the axis one might expect. During the October 10, 2025 crypto liquidation cascade, SPX showed by far the largest oracle-price collapse, the largest

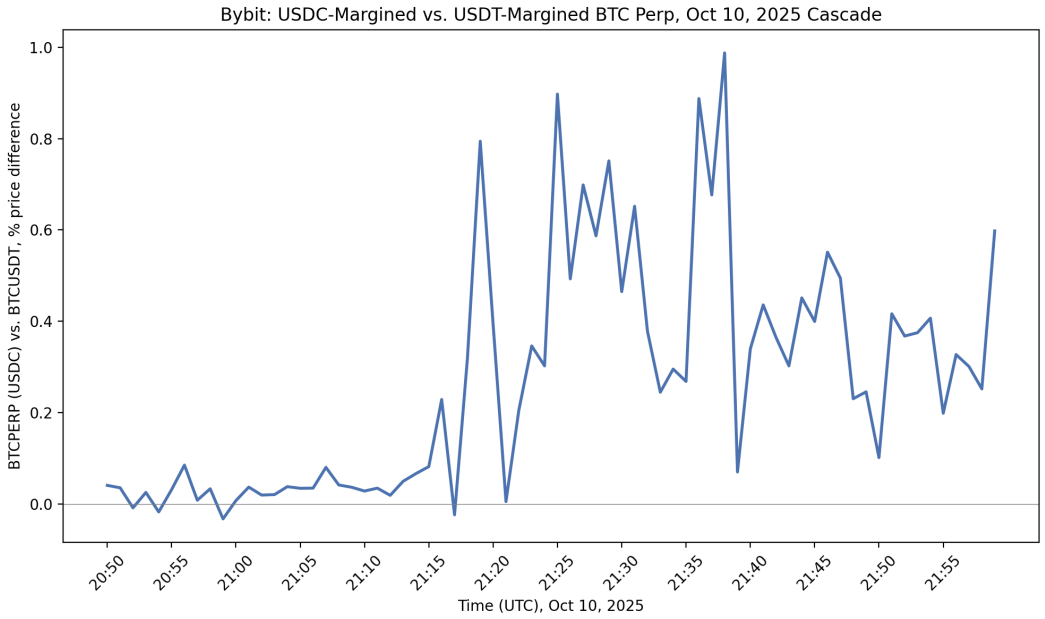

Figure 8: Bybit's USDC-margined vs. USDT-margined BTC perpetual during the October 10, 2025 cascade. The settlement-currency premium widens from a near-zero baseline to nearly 1%, the same directional effect as the Deribit/March 2023 comparison, but far smaller given BTC's much deeper liquidity.

peak divergence from its own oracle, and the widest bid-ask spread of any contract studied, while funding rates stayed flat throughout, ruling out a funding-mechanism failure as the cause. Full-history structural statistics confirm SPX ranks worst on every fragility metric across roughly 408 days of data, not just on the single stress day. What makes this informative is that SPX is priced by the same eight-exchange validator median as the four calm native contracts it is measured against, so its fragility cannot be an artifact of oracle construction; what distinguishes it is the thinness of the market its reference price is drawn from. The sample's one deployer-priced contract, xyz:GOLD, completes the picture from the other direction: it references gold, one of the deepest markets in existence, and stays calm on every metric despite an oracle set directly by its deployer. Deployer pricing alone is therefore not sufficient to produce fragility in the six-asset sample. A matched-pair design (Section 3.4.8) extends what that sample cannot: across 20 pairs of contracts on the identical underlying asset, the venue-deployed leg is persistently further from fair value than its validator-priced twin, in 19 of 20 pairs and by roughly an order of magnitude at the median ($p = 2.0 \times 10^{-5}$). The magnitude of that gap is closely tied to liquidity, which explains about four fifths of its cross-sectional variation. Oracle-level decomposition (Section 3.4.9) then resolves the attribution: for 19 of the 20 pairs the two oracles agree to within 0.005% at the median, and oracle disagreement contributes only 3.4–9.9% of the observed

gap, with no amplification under stress. The dislocation is therefore a stale-price effect on sparsely traded contracts; deployer-set reference construction is not shown to be systematically inferior once the confound is removed. The oracle data does, however, support a different and better-identified claim, developed in Section 3.4.10. Several underlyings are listed by multiple independent deployers at once, and those feeds disagree with one another by more than an order of magnitude even on Bitcoin, while oracle update frequency ranges from 21% to 56% frozen across the four verified S&P 500 listings and 26% to 55% across the four verified gold listings, all of which face the same trading calendar. Deployer-set pricing at its best is indistinguishable from the validator median; what the category introduces is variance across operators, undisclosed to the trader choosing between them. Regression evidence across several specifications, hourly, minute-level, Winsorized, and with continuous controls for volume and open interest, consistently locates the real signal in acute stress windows rather than in full-history averages, which are uniformly close to null. The second question, whether settlement-stablecoin volatility spills over into perpetual market quality, returns a more measured answer: everyday USDC fluctuations do not move Hyperliquid perpetual markets in any full-history specification, but during the small number of hours where USDC's own deviation from its peg exceeds two standard deviations, divergence volatility rises sharply for five of the six contracts studied. The March 2023 USDC/Silicon Valley Bank depeg, examined outside Hyperliquid's own data window through a same-exchange Deribit comparison, shows this same effect directly and close to mechanically: a Bitcoin perpetual quoted in a depegged currency inflated in price by roughly the size of the depeg itself.

Taken together, these results relocate rather than confirm the underlying concern. A venue-deployed perpetual's reference price is not systematically worse than a validator-median one. What is true, and arguably more consequential for market design, is that reference-price quality varies widely across operators listing the same asset on the same venue, with no interface-level signal distinguishing a feed that tracks the validator median to within 0.01% from one that diverges twenty times as far, or a contract whose oracle updates continuously from one carrying millions in daily volume on a price that is unchanged more than half the time. A perpetual referencing a thin, sparsely traded market, lacking any equivalent of the physical delivery mechanism that forces a traditional futures contract's price back to reality, appears structurally more vulnerable to both reference-price and liquidity breakdowns once markets come under genuine stress, even though it may perform indistinguishably from a deeply-referenced contract on an ordinary trading day. This has a direct implication for the regulatory questions now facing U.S. markets as the CFTC moves toward approving listed perpetual futures: the question worth asking is not who computes a contract's reference price, but whether per-contract reference-price quality is disclosed, mon-

itored, and bounded. On the evidence here, two contracts sitting side by side on the same venue, presented identically to the trader, can differ in reference-price reliability by more than an order of magnitude. A reference price that looks authoritative in calm conditions, the same lesson LIBOR and the WM/Reuters FX fix each taught in their own eras, is not the same thing as a reference price that has actually been tested against stress.

Several limitations qualify these results. The six-asset core sample contains only one thinly-referenced contract and only one deployer-priced contract. That limitation motivated the matched-pair design in Section 3.4.8, which raises the deployer-priced sample to 20 validated pairs and removes the asset-class confound, but the matched pairs carry their own limitations: all but one are listed by a single deployer (`hyna`), with `flx:XMR` the sole independent replication, so the dispersion result in Section 3.4.9 rests on four failing contracts drawn from two deployers and should be treated as a documented pattern rather than a population estimate. The proposed mechanism for those failures, thin or fragmented spot markets in the underlying, is inferred from the fact that Monero fails on both deployers while Bitcoin fails on neither, and is not directly tested here. XRP's 44% oracle swing during the October 10 cascade is the clearest data point cutting against a pure depth story. The magnitude of the controls-adjusted regression coefficient, while robust in direction and strength across a leave-one-asset-out check, is imprecisely estimated because of real multicollinearity between the control variables and the main regressor. The Volatility×Open Interest specification (Section 3.4.6) produces the opposite sign from the hypothesized direction. Additional specifications indicate that this term is highly collinear with an Open Interest×Funding relationship rather than reflecting a genuine volatility/open-interest effect, robust to swapping Open Interest for Volume and to dropping the funding rate entirely, though the interaction terms themselves are severely multicollinear, so this family of specifications should be read for direction and fit rather than precise coefficient magnitudes. The clustered standard errors this analysis calls for, one-way clustering by asset per Petersen (2009), are applied to all seven headline RQ1/RQ2 specifications reported in Section 3.4.4, including the controls-adjusted specification (Table 2, column 5), and to both minute-level specifications (Table 4, columns 1-2). Coin-clustering does not apply to the Winsorized per-coin regressions, since each is estimated on a single asset, and is not extended to the daily specification.

Future work has several natural directions. The `xyz:SKHX` event (Section 3.4.7) was treated here as corroborating evidence rather than a formal regression specification; a full event-study treatment matching Section 3.4.1's Oct 10 analysis, and a systematic search for further HIP-3 stress events beyond the three studied here, would allow the central finding to be evaluated across a broader set of independent stress events. The Volatility×Open Interest interaction term's sign puzzle now has a plausible resolution (Section 3.4.6), but

the severe multicollinearity among its component interaction terms means a more careful specification, ideally one able to separate volatility, liquidity, and funding effects without relying on multiplicative interactions of the same few underlying variables, would strengthen the finding considerably. A further natural extension would examine the relationship between a stablecoin's own collateral yield and the yield on posted U.S. Treasuries, a natural complement to the stablecoin-stress findings reported here. As perpetual futures continue moving from an offshore, crypto-native product into a regulated instrument traded on U.S. exchanges, understanding which reference-price designs hold up under stress, and which do not, will only become more directly useful to the people building and regulating that market.

References

Shiller, R. (1993). Measuring Asset Values for Cash Settlement in Derivative Markets: Hedonic Repeated Measures Indices and Perpetual Futures. *Journal of Finance*.

Angeris, G., Chitra, T., Evans, A., & Lorig, M. (2023). Short Communication: A Primer on Perpetuals. *SIAM Journal on Financial Mathematics*, 14(1), SC17–SC30.

Ackerer, D., Hugonnier, J., & Jermann, U. J. (2024). Perpetual Futures Pricing. *SSRN Electronic Journal*.

Alexander, C., Choi, J., Park, H., & Sohn, S. (2020). BitMEX Bitcoin Derivatives: Price Discovery, Informational Efficiency, and Hedging Effectiveness. *Journal of Futures Markets*, 40(1), 23–43.

He, S., Manela, A., Ross, O., & von Wachter, V. (2022). Fundamentals of Perpetual Futures. *SSRN Electronic Journal*.

Ruan, Q., & Streltsov, A. (2022). Perpetual Futures Contracts and Cryptocurrency Market Quality. *SSRN Electronic Journal*.

Giagkiozis, I., & Said, E. (2024). Reconciling Open Interest with Traded Volume in Perpetual Swaps. *Ledger*, 9.

Eichengreen, B., Nguyen, M. T., & Viswanath-Natraj, G. (2025). Stablecoin devaluation risk. *The European Journal of Finance*, 31(11), 1469–1496.

Duffie, D., & Stein, J. C. (2015). Reforming LIBOR and Other Financial Market Benchmarks. *Journal of Economic Perspectives*, 29(2), 191–212.

Duffie, D., & Dworczak, P. (2021). Robust Benchmark Design. *Journal of Financial Economics*, 142(2).

Benenchia, M., Galati, L., & Lepone, A. (2024). To fix or not to fix: The representativeness of the WM/R methodology that underpins the FX benchmark rates. A pre-registered report. *Pacific-Basin Finance Journal*, 84, 102311.

Petersen, M. A. (2009). Estimating Standard Errors in Finance Panel Data Sets: Comparing Approaches. *Review of Financial Studies*, 22(1), 435–480.

Shehdula, E. (2026). A 24/7 Equity Perp on Hyperliquid Liquidated $59M in Minutes. *Allium Research*, July 29, 2026.

Shehdula, E. (2026). When Wall Street Sleeps: Onchain Price Discovery for US Equities During Overnight and Weekend Gaps. *Allium Research*, May 2026.

Chernoff, A., Jagtiani, J., & Yoshida, N. (2026). Flight to Safety: Evaluating Stablecoin's Role as a Safe-Haven Asset in DeFi Markets. *Federal Reserve Bank of Philadelphia Working Paper*, No. 26-24.

Hayes, A. (2016, May 13). Why XBTUSD is a Superior Trading Product. *BitMEX Blog*.

Hayes, A. (2018, August 14). Why Quanto? *BitMEX Blog*.

CME Sues U.S. Regulator to Stop Kalshi From Offering Popular Perp Futures. (2026, June). *The Wall Street Journal*.

Jakab, S. (2026). Perps Are the Risky New Derivatives That Could Amplify Stock Blowups. *The Wall Street Journal*.

Duffie, D., Skeie, D., & Vickery, J. (2013). A Sampling-Window Approach to Transactions-Based Libor Fixing. *Federal Reserve Bank of New York Staff Reports*, No. 596.

Christin, N., Routledge, B. R., Soska, K., & Zetlin-Jones, A. (2023). The Crypto Carry Trade. *Working Paper, Carnegie Mellon University*.

Schmeling, M., Schrimpf, A., & Todorov, K. (2025). Crypto Carry. *SSRN Electronic Journal*.