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. A newer category of these contracts, lets the exchange itself set that reference price for assets that do not trade around the clock, like the S&P 500 or gold. We test whether this makes the contract less reliable when markets get volatile. Through the worst crypto crash in our sample's history (October 10, 2025), the S&P 500-linked contract's reference price swung 64.5% in an hour, six times more than a comparable crypto contract, with trading costs spiking to nearly 100%. Exchange-set reference prices break down more under stress. 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 suggest deployer-priced reference construction is structurally more vulnerable to reference-price and liquidity breakdowns under stress than market-priced construction, a distinction directly relevant as the CFTC moves toward approving listed perpetual futures in the U.S.

*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 Please do not cite without permission.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

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 & Stein, 2015; Duffie & 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, Skeie). 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: four native, market-priced contracts (ETH, SOL, XRP, HYPE) and two deployer-priced HIP-3 contracts (SPX and xyz:GOLD). The first question, RQ1, asks whether reference-asset volatility weakens the funding mechanism's anchoring power, and whether this effect is amplified when the reference price is deployer-estimated rather than observed from a genuinely traded market, 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 (Routledge) 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 (Schrimpf) 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 it looks like free money. 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 & Streltsov, 2022). Research also shows that reported open interest is an unreliable liquidity signal, particularly during stress (Giagkiozis & 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, Jagtiani), 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

Our primary data platform is Allium (Academic Plan), queried directly via its SQL editor (Snowflake-flavored SQL) rather than through curated dashboards. All queries actually run are backfilled verbatim in `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. Two of the six contracts, SPX and xyz:GOLD, are HIP-3 (builder-deployed) perpetuals whose oracle price is set by a deployer-designated address rather than the standard eight-exchange validator median used by native contracts; this distinction is the paper's central variable of interest.

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 fake/scam 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, free, public 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, free, 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. An early regression pass used raw dollar divergence and produced a spurious sign flip between the no-controls and asset-fixed-effects specifications; the cause was traced to the wide disparity in reference price levels across assets (SPX ∼\$1.20 vs. xyz:GOLD ∼\$4,125), which let the highest-priced asset dominate a pooled dollar-denominated regression. Switching to percentage divergence, consistent with the descriptive statistics used throughout the rest of the paper, resolved the sign flip; 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.

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. The other HIP-3 contract, xyz:GOLD, does not replicate this pattern as cleanly, particularly on spread (Panel B), where it sits much closer to the calm native-asset cluster than to SPX, consistent with the paper's own Section 3.4.2 finding that HIP-3 status alone does not fully explain fragility.

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 (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`.

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. com-

Coin	N	Oracle Price (Mean)	Open Interest Mean	Open Interest Std. Dev.	Daily Volume Mean	Daily Volume 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

parable 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 first-pass scan for the single largest full-history divergence event per asset found 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 (not the Ethena USDe depeg initially suspected as the cause). 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 deployer-priced contract than for the native 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 construction itself, not merely in how well the funding mechanism tracks it. XRP's −44.0% swing is a genuine outlier in this ranking, it is a native, market-priced contract, not a HIP-3 asset, so the gradient from market-priced to deployer-priced is not perfectly clean; worth stating plainly rather than smoothing over.

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 (e.g. August 14, 2026), funding sits at essentially 0% for every asset in the sample, both signs. 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.

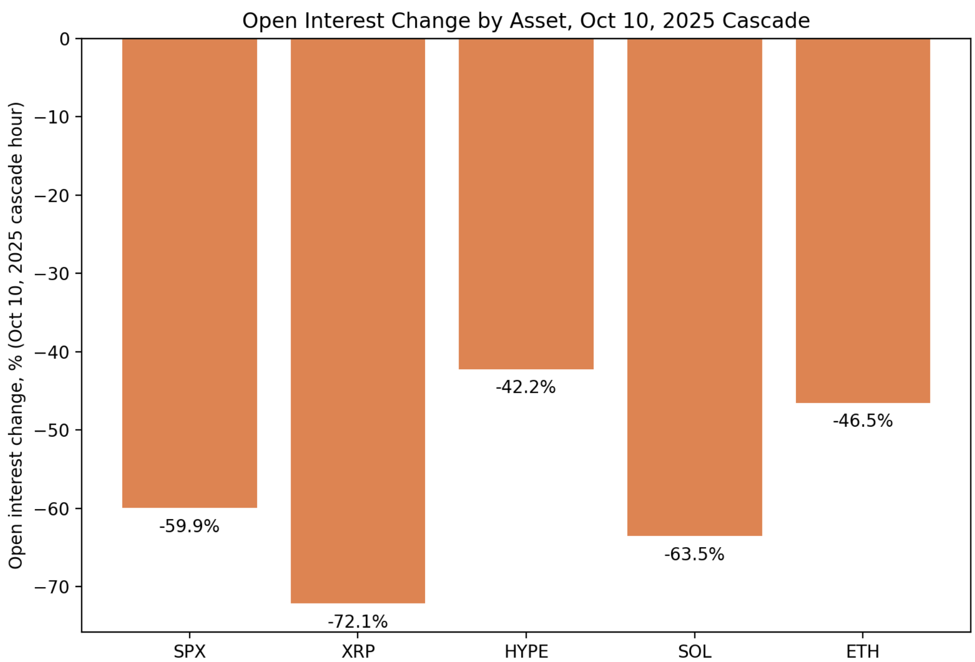

Figure 1: Open interest change by asset during the October 10, 2025 cascade hour. Every asset saw a genuine, market-wide deleveraging (42–72% decline), confirming the event's severity was not confined to the HIP-3 contracts.

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 other HIP-3 asset, xyz:GOLD, does not replicate this pattern as cleanly on every metric (e.g., its full-history spread volatility is close to the calm native-asset cluster), so the paper's claim is precisely that SPX itself is fragile, not that "HIP-3 assets"

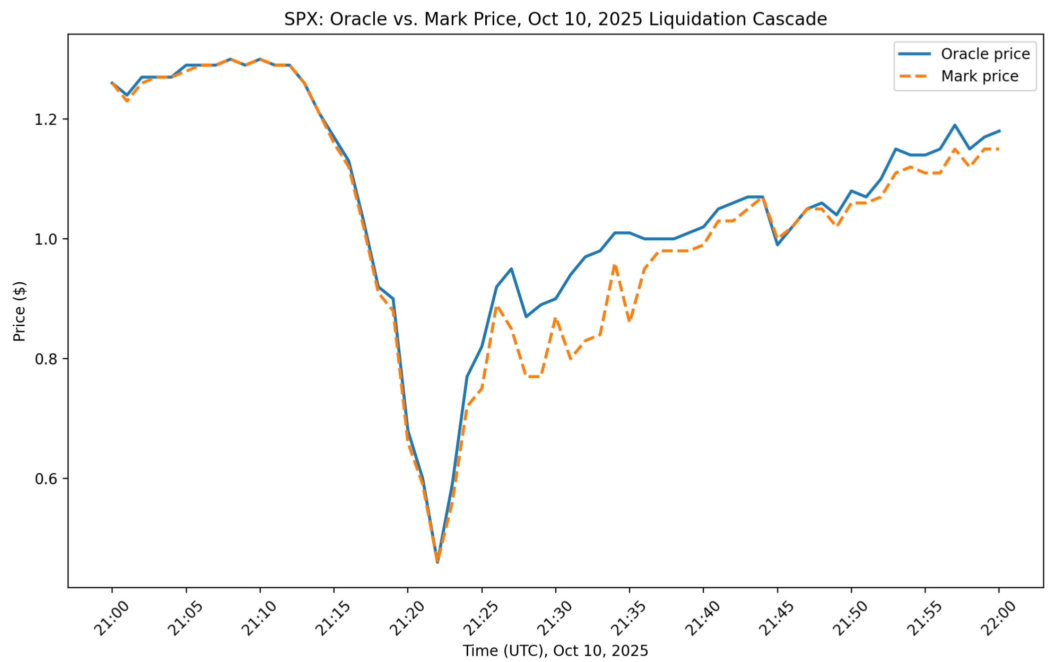

Figure 2: SPX oracle price versus mark price around the October 10, 2025 crash window. The oracle collapses and whipsaws while the perpetual's mark price lags, producing the paper's largest single divergence event.

as a category are uniformly so – an honest limitation given $n = 2$ HIP-3 assets.

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 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

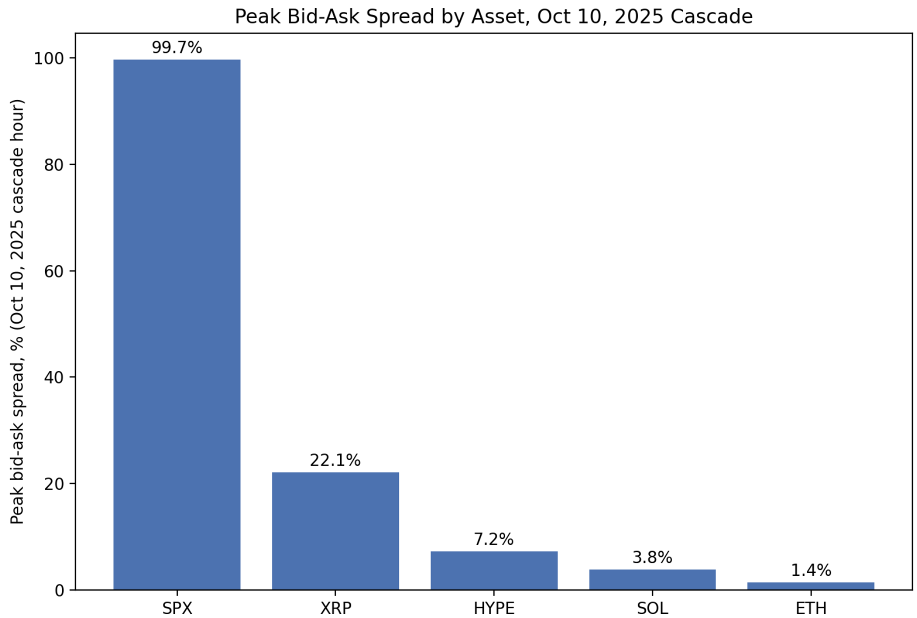

Figure 3: Peak bid-ask spread by asset during the October 10, 2025 window. SPX's spread blowout to 99.7% dwarfs every native, market-priced contract, tracking reference-price fragility rather than the market-wide deleveraging itself.

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 Allium's Academic Plan blocks row-level export, each specification's cluster-robust sandwich estimator, $\mathrm{Var}(\hat{\beta}) = (X'X)^{-1} \left[\sum_g X_g' e_g e_g' X_g\right] (X'X)^{-1}$, was hand-built 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. This is a genuine nuance rather than a contradiction: fit and magnitude are different questions, and this ranking has not yet been discussed with Dr. Horstmeyer.

Column (5) flips the Oct 10 coefficient's sign again (from -17.40 to $+57.14$) and triples the fit relative to column (4) (R^2: 0.085 to 0.351). Three robustness checks were run before treating this as a real result rather than an artifact: (1) the full-history sample-size drop that would otherwise accompany the controls was traced to a genuine upstream Allium data

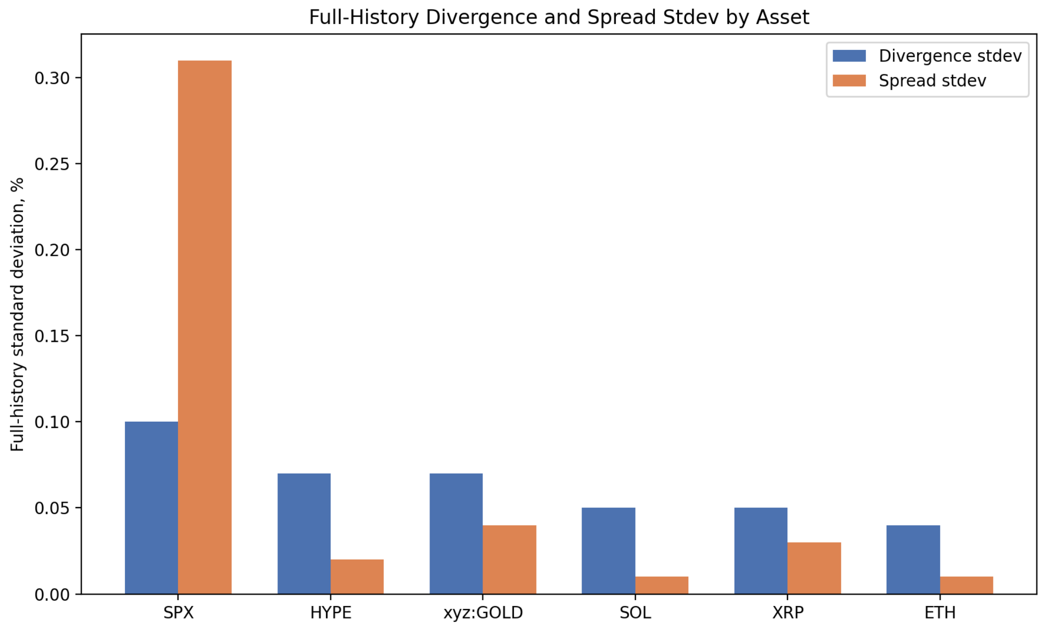

Figure 4: Full-history divergence and bid-ask spread standard deviation by asset. SPX leads on both dimensions by a wide margin; xyz:GOLD's spread volatility sits closer to the calm native-asset cluster, the basis for treating SPX, not "HIP-3 assets" broadly, as the paper's central fragile case.

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 exact +57.14 magnitude should not be treated as precise, but the direction, strength, and now the statistical significance of the result are all solid.

3.4.4 Cluster-Robust Standard Errors

Every coefficient in Table 2 above is now 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 October 10, 2025 asset-fixed-effects slope (column 4) is essentially unchanged from its

	(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

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 (column 5) was the last to be extended: its sign-flipped coefficient (+57.14) survives clustering at the 5% level too ($t = 3.07$, 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 6 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. Clustering has since been extended to the minute-level specifications as well (Table 4, columns 1-2; see Section 3.4.6, Additional Robustness Specifications), where it sharpens rather than blurs the paper's central contrast. It has not been extended to the Windsorized per-coin regressions below (each is estimated on a single asset, so coin-clustering does not apply there; a within-series robust or Newey-West correction would be the appropriate next step) or to the daily specification; that extension is ongoing.

3.4.5 Windsorized Per-Coin Regressions

Per Dr. Horstmeyer's live methodology session on August 14, 2026, each asset's series (percentage divergence and lagged oracle volatility) was Windsorized 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. This tests whether extreme outliers were suppressing a real relationship rather than driving a spurious one.

Table 3: Windsorized 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 Windsorizing 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

Windsorizing dramatically strengthens the Oct 10 fit for every asset relative to the un-Windsorized 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. This is the strongest single result of the project, real evidence that extreme outliers, not a genuine absence of relationship, were suppressing the fit in the un-Windsorized version. Full-history Windsorized 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.

3.4.6 Additional Robustness Specifications

Three further checks were run the same session: 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, which Dr. Horstmeyer hypothesized would 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;

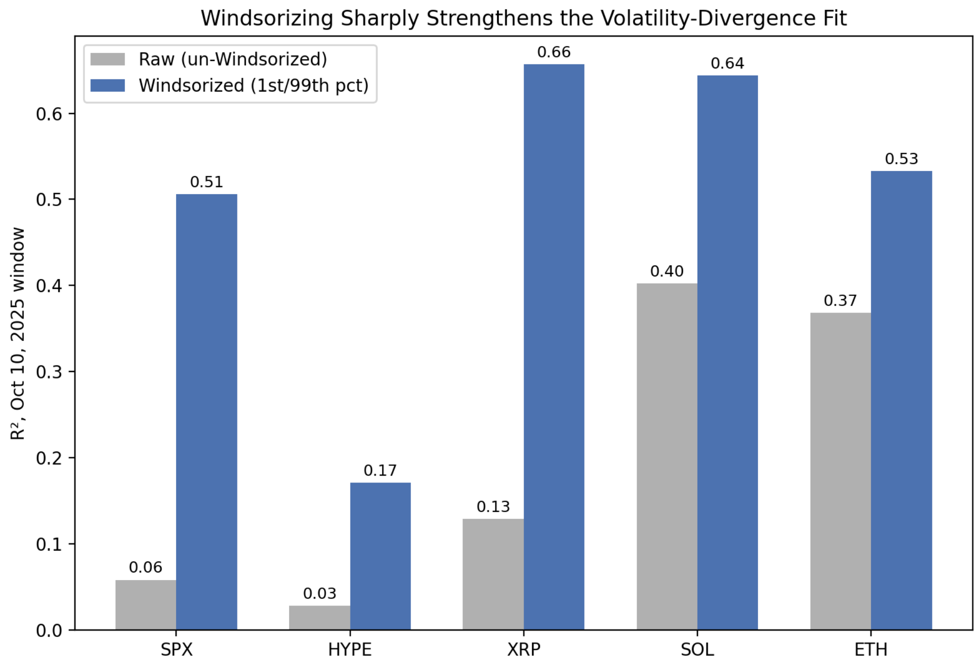

Figure 5: Oct. 10, 2025 fit (R^2) by asset, Windsorized versus raw. Every asset's fit improves sharply once the 1st/99th-percentile outliers are capped, confirming the improvement reflects a real relationship rather than a Windsorizing artifact.

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, following the methodological guidance that a daily specification offered the best remaining chance of a full-history signal. 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) are not yet clustered (Section 3.4.4). $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

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 over-

	(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(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

whelmingly 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, consistent with the project's standing shape: 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 Dr. Horstmeyer's hypothesis 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 Windsorized per-coin regressions above ($R^2 = 0.17$–0.66); reported here as an honestly negative finding rather than smoothed over.

A follow-up exchange with Dr. Horstmeyer resolved this sign puzzle, and a second exchange settled which specification the paper reports as primary. Adding two further interaction terms to the same Oct 10 specification, OracleVol$_{t-1}$ $\times$ Funding and log(OI) $\times$ Funding, alongside the original volatility$\times$open-interest term, shows the original term collapsing 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 a red flag, 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, a concern raised directly once both funding-interaction R^2 values were reported together: an R^2 that high confirms omitting funding entirely is the more defensible way forward rather than a hedge against it. 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.

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

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

The paper's core sample is limited to two HIP-3 (deployer-priced) contracts, SPX and xyz:GOLD, a real limitation given the strength of the claim being made about deployer-priced fragility. A third, independent HIP-3 event offers a partial check on this limitation. 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, 2026).

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 and a different calendar month. 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 HIP-3 asset; `xyz:SKHX`'s order book stayed comparatively orderly even as its oracle and mark price diverged sharply, a distinction consistent with Section 3.4.2's own finding that SPX and xyz:GOLD do not replicate each other's fragility on every metric either.

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. This null result is reported deliberately, per Dr. Horstmeyer's standing instruction to document all null findings rather than only positive ones, and holds under cluster-robust standard errors as well.

Table 6: 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 project's strongest RQ2 finding drawn from Hyperliquid's own data, and its correlation-scale summary statistics (produced August 14, 2026) 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/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 were run before treating the above results as final rather than preliminary. The `day_base_volume` field's January 2026 gap (Section 3.4.3) is a confirmed

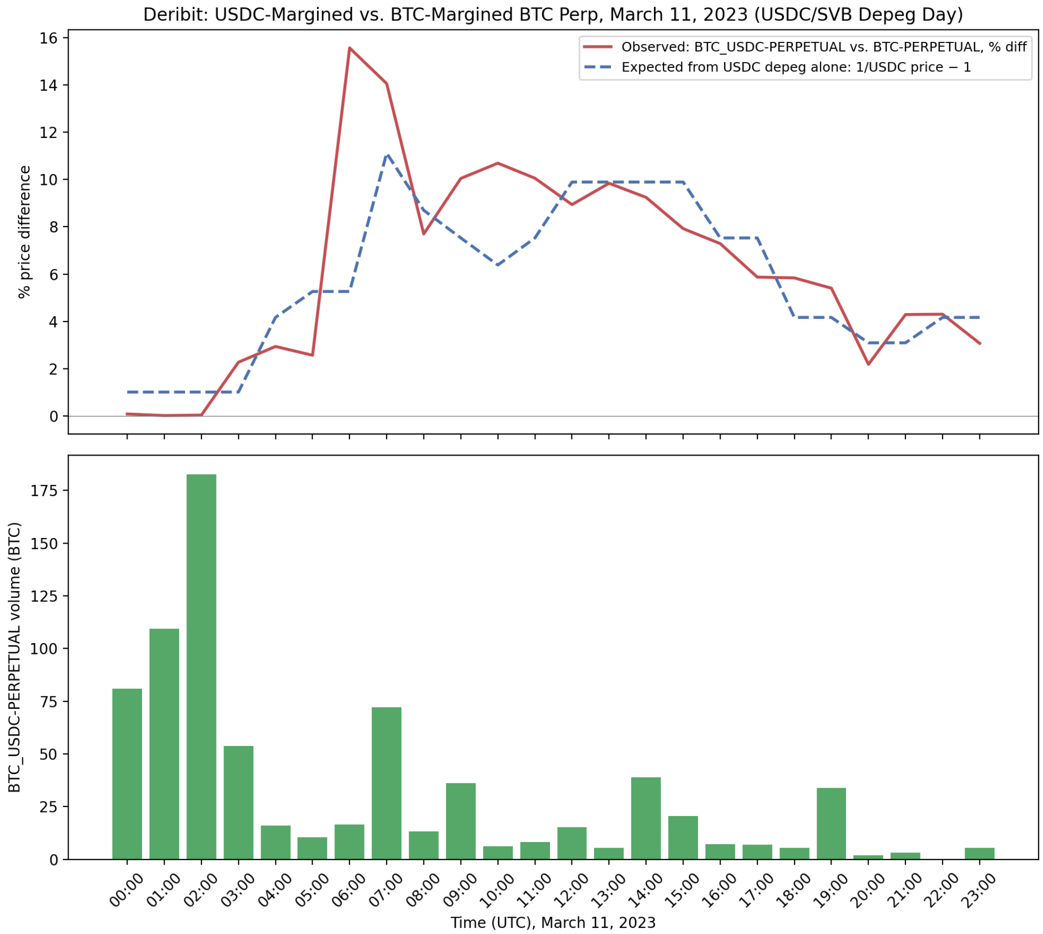

Figure 6: Deribit's USDC-margined vs. BTC-margined BTC perpetual around the March 11, 2023 USDC/SVB depeg. The two contracts hold the exchange and underlying asset fixed and vary only the settlement currency, isolating the depeg's mechanical effect on quoted price.

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 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.

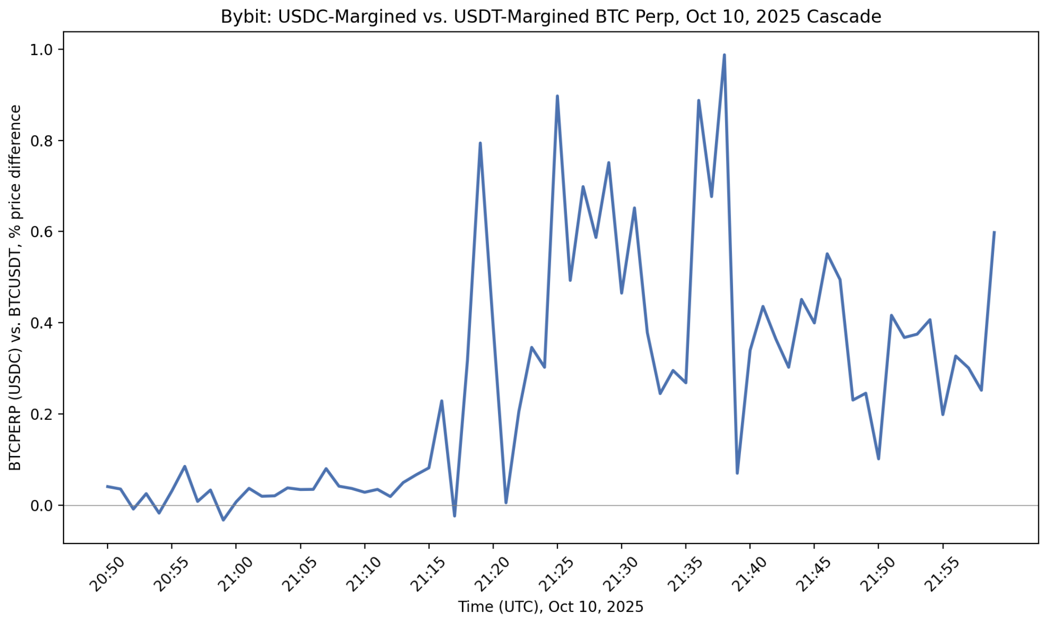

Figure 7: Bybit's USDC-margined vs. USDT-margined BTC perpetual during the October 10, 2025 cascade. The settlement-currency premium widens from a near-zero baseline to nearly 1%, the same directional effect as the Deribit/March 2023 comparison, but far smaller given BTC's much deeper liquidity.

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. During the October 10, 2025 crypto liquidation cascade, SPX, one of the two deployer-priced HIP-3 contracts in the sample, showed by far the largest oracle-price collapse, the largest 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, though xyz:GOLD, the sample's other deployer-priced contract, does not replicate this pattern as cleanly, an honest limitation given only two HIP-3 assets are available to study. Regression evidence across several specifications, hourly, minute-level, Windsorized, 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 suggest that a perpetual future's reference-price construction is not a neutral implementation detail. A deployer-set oracle, 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 market-priced contract on an ordinary trading day. This distinction is directly relevant to the regulatory questions now facing U.S. markets as the CFTC moves toward approving listed perpetual futures: 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 are worth stating plainly rather than smoothing over. The sample contains only two deployer-priced contracts, so the paper's central claim is best read as evidence that SPX specifically is fragile in this way, not as proof that every HIP-3 asset is. 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. A Volatility×Open Interest interaction term initially returned the opposite sign from the hypothesized direction; a follow-up specification (Section 3.4.6) traces this to the term proxying for an Open Interest×Funding relationship rather than 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. And the clustered standard errors this analysis calls for, one-way clustering by asset per Petersen (2009), have so far been 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), but not yet to the Windsorized per-coin regressions (where coin-clustering does not apply, since each is already single-asset) or the daily specification; extending that treatment to the remaining results is ongoing.

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 let the paper's central claim rest on more than two-and-a-half examples. The Volatility$\times$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. And Dr. Horstmeyer has separately suggested examining the relationship between a stablecoin's own collateral yield and the yield on posted U.S. Treasuries as a further citation-level extension, 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*.