A Cross-Venue Data Dictionary for Perpetual Futures: Funding-Rate Mechanisms, Pre-Trade Rules, and Measurement Error Across 21 Exchanges

Hang Liu humarb Research, London, United Kingdom https://humarb.ai · ORCID 0009-0008-1059-3926 humarb Research 03 · Methodological preprint · 3 November 2026 Correspondence: humarb Research, https://humarb.ai/research/

Abstract

Cross-venue perpetual-futures analysis often treats a funding rate as a directly comparable scalar. We test that premise using official documentation and read-only public interfaces for 21 centralized and on-chain venues. We construct a dated data dictionary covering funding formulas, settlement schedules, caps, field semantics, clock and rate units, open-interest and turnover bases, fees, liquidation backstops, and pre-trade rules. Fifteen venues share a clamped interest-minus-premium core, but three use distinct formula structures and three require separate treatment. Instrument-level intervals invalidate an exchange-wide eight-hour assumption: Gate's 1,070 rows on 3 November 2026 contain 420 four-hour contracts, while GRVT and Variational also return multiple intervals. A new sample of 1,840 public history rows shows median absolute timestamp offsets from the nearest hour of 36 ms at Hyperliquid, 937 ms at Extended, and 3,680 ms at Pacifica; all 100 edgeX records are exact. Deterministic examples show 100-fold unit errors, twofold annualization errors, and fee differences exceeding one typical funding payment. The results provide a reproducible normalization layer for funding-rate and execution-cost measurement, not evidence of trading profitability.

Keywords: perpetual futures; funding rates; data dictionary; market microstructure; measurement error; crypto derivatives **JEL classification:** G12; G13; G14; C58

1. Introduction

A perpetual future has no expiry date. Its contract price is anchored to an external reference through funding transfers between long and short positions. When the same underlying asset trades on several venues, the difference between funding streams is an input to cross-venue hedged positioning and to studies of segmentation, price discovery, and arbitrage limits. The empirical operation appears trivial: read two rates, place them on a common annualized basis, and subtract. The premise behind that operation is strong. It requires the two fields to describe the same economic object, over the same interval, in the same numerical unit, at the same phase of the funding cycle, and for positions whose cash-flow bases and trading states are comparable.

Public interfaces do not guarantee those conditions. Binance's `lastFundingRate` behaves like a current forecast despite its name; edgeX publishes the last settled rate as `fundingRate` and places the forecast in a separate field; Deribit's `funding_8h` is trailing

realized accrual. GRVT reports rates in percentage points and timestamps in nanoseconds, while Gate mixes seconds and milliseconds within WebSocket frames. Eight hours is a common default rather than a universal property. Some venues accelerate settlement after cap hits, some calculate an eight-hour rate but pay one eighth each hour, and others accrue continuously. Field names therefore do not identify a common statistical variable.

This paper studies the measurement problem directly. We ask two questions. First, which funding and pre-trade mechanisms differ systematically across the 21 venues in the research universe? Second, how do those differences enter measured rate differentials, settlement cash flows, and entry or exit costs? The object of inference is not a latent pricing coefficient. It is a set of public, dated rules and fields whose meanings can be checked against official text and read-only responses. This distinction matters: a deterministic unit error is not sampling noise, and collecting more observations cannot remove it.

We make four contributions. First, we assemble a venue-level data dictionary for 10 centralized exchanges (CEXs) and 11 on-chain or hybrid venues. The dictionary joins funding formulas, schedules, caps, field semantics, units, open interest (OI), turnover, fees, liquidation backstops, price constraints, quantity rules, status fields, trading hours, and transfer-state availability. Every venue row carries an official source, an evidence date, a refresh outcome, and explicit inference notes. Second, we define a normalization tuple that maps venue-native fields into comparable rates and cash flows. The tuple separates semantic phase, numerical basis, accrual interval, event time, position basis, and trading state. Third, we quantify the consequences using reproducible deterministic examples and two empirical distributions: instrument-level settlement intervals and 1,840 funding-history timestamps. Fourth, we release the public snapshots, extraction scripts, analysis code, CSV/JSON dictionaries, and vector figures needed to reproduce every new numerical table.

The mechanism census yields a clear taxonomy. Fifteen venues use the core form $F=P+\text{clamp}(I-P)$, including implementations that rescale the eight-hour basis to hourly payments. MEXC clamps the premium itself, Deribit applies a premium deadband without an interest term, and dYdX obtains premiums from validator votes. Paradex smooths a cross-venue weighted median, GMX adjusts a state factor with long–short imbalance, and LBank publishes values observed to be bit-identical to Binance with a short lag. These are not interchangeable parameterizations of one formula.

The measurement consequences are economically material even before any statistical model is estimated. Reading a four-hour contract as an eight-hour contract halves its daily funding estimate. Reading a percentage-point value as a decimal multiplies the rate by 100. In an illustrative premium scenario, moving the clamp changes the rate by 0.01 percentage point per period. Base-tier taker fees range from zero to 0.075%, implying a $15 difference for a single $10,000 round trip before spread, impact, transfer, borrowing, or liquidation costs. These examples are not return forecasts. They show why mechanism normalization belongs inside the measurement equation rather than in a footnote.

The new evidence also demonstrates why dated snapshots are necessary. Gate’s public contract list grew from 1,048 rows on 19 October 2026 to 1,070 on 3 November, while its four-hour count rose from 416 to 420. GRVT returned 125 eight-hour and 69 four-hour

perpetual rows in the refresh; Variational returned 261 eight-hour, 304 four-hour, and 1 one-hour rows. In the timestamp sample, edgeX records lie exactly on the hour, whereas Pacifica records have a median offset of 3.680 seconds. These offsets describe public record timestamps, not guaranteed cash-settlement latency. They nevertheless reject the assumption that a nominal hourly label is sufficient to reconstruct event time.

This study differs from the two preceding humarb Research papers. Liu (2026a) defines executable net carry after funding, execution, and liquidation-aware costs; Liu (2026b) frames estimation using settled histories, depth-walked execution, and conservative evaluation. Those papers ask how a carry measure should be constructed and evaluated once inputs are comparable. The present paper asks an antecedent question: when do two public fields represent comparable inputs at all? It supplies the venue-specific dictionary and normalization evidence required before either net-carry framework is applied across heterogeneous venues.

The paper proceeds as follows. Section 2 reviews perpetual-futures pricing, funding design, arbitrage, and risk mechanisms. Section 3 documents the institutional taxonomy. Section 4 describes sources, sample construction, and snapshot coverage. Section 5 defines the normalization model and measurement-error calculations. Section 6 reports mechanism, interval, timestamp, and pre-trade results. Section 7 examines robustness across dates, venue groups, parameter choices, and alternative timestamp estimators. Section 8 states limitations, and Section 9 concludes.

2. Related work

2.1 Perpetual claims and funding design

Shiller (1993) introduces perpetual futures as cash-settled claims designed for assets whose spot values are difficult to observe. Although crypto perpetuals differ in market structure and collateral conventions, the central idea persists: periodic transfers replace expiry-based convergence. Angeris et al. (2022) formalize perpetuals in continuous time and connect funding to replication and no-arbitrage restrictions. He et al. (2022) emphasize that perpetual futures need not converge to spot and study the economic information contained in funding. Ackerer et al. (2026) derive pricing results for linear, inverse, and quanto perpetuals in discrete and continuous time. Kim and Park (2025) study how funding-rate design affects alignment with a target asset.

This literature explains why funding exists and how its functional form affects pricing. Our object is different. We take venue mechanisms as institutional objects and ask whether their published outputs can be compared without measurement distortion. The clamp-position comparison between MEXC and the Binance family, for example, is not a competing pricing model. It is a demonstration that fitting one venue’s formula to another venue’s fields can create a false model-rejection result.

2.2 Arbitrage limits, costs, and empirical implementation

Funding differentials are often interpreted through an arbitrage lens, but a quoted differential is not an executable return. Execution requires entry and exit on two venues, collateral management, and survival through settlement and liquidation events. Spread, market impact, transfer delay, borrowing, and account-specific fee schedules intervene between a public rate and realized cash flow. Liu (2026a, 2026b) develops net-carry and evaluation frameworks around this distinction. The present study supplies a lower-level input audit: semantic phase, rate unit, accrual interval, notional basis, status, and book fillability.

This focus also complements empirical work on funding predictability and arbitrage. A forecasting model can be correctly specified in a statistical sense yet receive inputs that mix next-period forecasts with last-period settlements. Likewise, a backtest can include realistic transaction-cost parameters while assigning an eight-hour rate to a four-hour instrument. Our results imply that data-lineage checks should precede sample-out-of-sample evaluation. They do not substitute for such evaluation when a predictive claim is made.

2.3 Liquidation and loss-allocation mechanisms

Auto-deleveraging (ADL), insurance funds, backstop liquidity providers, and socialized losses determine who absorbs residual losses when ordinary liquidation fails. Chitra (2026) formalizes impossibility and optimization boundaries for ADL, while Chitra et al. (2026) study ADL as sequential online learning. Sepper (2026) addresses forward-looking liquidity risk in perpetual markets. Gupta and Polson (2026) examine a stock-linked pre-IPO perpetual, and Nechepurenko (2026) classifies event-linked perpetuals. These extensions make status and contract-type fields economically relevant: a frozen, pre-market, event, or real-world-asset contract should not be treated as an ordinary crypto perpetual merely because its interface publishes a funding field.

Our venue evidence shows that backstop rules are not portable. Gate’s documented ADL ranking uses maintenance-margin multipliers in the opposite direction from Binance, Bitget, and ApeX for profitable positions; dYdX describes randomly selected counterparties; Paradex uses socialized losses rather than ADL; Variational’s pool structure can transfer the liquidation penalty to the user. We record these rules as tail-risk inputs but do not estimate their trigger frequencies, which would require account-level or long-horizon event data.

2.4 Position in the literature

The reviewed literature provides theories of perpetual pricing, funding design, arbitrage, and tail risk. The missing operational link is a cross-venue data dictionary that preserves the meaning of public fields and distinguishes documented facts, measurements, and inference. This paper addresses that link with a census of the 21 venues in scope. It does not claim coverage of every perpetual venue or estimate a pricing model. Its contribution is to make cross-venue measurement auditable before economic hypotheses are tested.

3. Institutional background: one label, several mechanisms

3.1 Formula families

Table 1 summarizes the taxonomy. The shared expression for the largest group is

$$F_t = \bar{P}_t + \mathrm{clamp}(I_t - \bar{P}_t, -c, c),$$

where P_t is an averaged premium index, I_t is an interest component, and c is a small clamp, commonly 0.05%. The outer funding cap is a separate mechanism and must not be confused with c. Fifteen venues use this core form, but their sampling, scaling, interest rates, and premium construction differ.

Table 1. Funding-formula taxonomy across the 21-venue census

Class	Venues	Distinguishing feature
Clamped interest-minus-premium core (15)	Binance, OKX, Bybit, Bitget, Gate, KuCoin, BingX, Hyperliquid, edgeX, Extended, GRVT, Lighter, Pacifica, Variational, ApeX	Equation (1), with material timing and parameter variants
Premium clamp (1)	MEXC	F=I+clamp(P); clamp acts on premium itself
Deadband accrual (1)	Deribit	No interest term; deadband around zero; continuous accrual
Validator-voted premium (1)	dYdX	Median of validator funding-premium votes
Cross-venue EWMA (1)	Paradex	Weighted median of external impact premiums with smoothing
Imbalance-driven state (1)	GMX	Per-second adjustment with long–short imbalance
Published mirror (1)	LBank	Values observed to match Binance with lag; publication method undocumented

Note: KuCoin’s family membership is an inference from isomorphic public parameters because formula text was not obtained. Variational shares the core family but follows external settlement windows. LBank’s mirroring is a measured statement, not a claim about its internal method. Sources are the official venue documents listed in the references, accessed 19–20 October 2026.

Interval scaling creates the first major within-family division. OKX, Gate, Bitget, and GRVT normalize an eight-hour basis to N-hour settlement. Hyperliquid, Extended, Pacifica, and Lighter calculate on an eight-hour basis and pay one eighth hourly. ApeX uses N=1 without the same scaling. Bybit’s documentation does not normalize the per-period absolute rate in the same way; its interest component scales with the number of periods. Thus, two numerically identical rates need not imply the same daily transfer.

Interest components also vary. A common crypto value is 0.01% per eight hours, but Gate documents zero for TradFi contracts, BingX documents 0.004% per eight hours for foreign-exchange contracts, Extended uses 0.0035% for real-world-asset order-book markets, and edgeX publishes a daily rate divided by the number of settlement periods.

Premium construction ranges from order-book impact prices to GRVT's exponentially weighted fair price, Variational's request-for-quote (RFQ) prices, and Hyperliquid's distinct builder-market formula. ApeX includes the current interval's funding rate in its premium definition. These details preclude a universal parser based only on field names.

MEXC applies the clamp at a different location:

$$F_t^M = I_t + \mathrm{clamp}(P_t, -c, c).$$

For P=0.1\%, I=0.01\%, and c=0.05\%, Equation (2) gives 0.06%, whereas Equation (1) gives 0.05%. MEXC's interest value is an inference from measured pinning because the consulted official page did not state the number; the example is therefore labeled as using that inferred value. Deribit instead computes an uncapped rate as max(0.025\%,P)+min(-0.025\%,P), producing zero inside the deadband. dYdX uses the median of validators' per-block premium votes, divided by eight for hourly payment, plus an applicable interest component.

Paradex combines a weighted median of impact premiums from six venues with clipping and exponential smoothing. GMX has no premium-index formula: its funding factor evolves with long–short imbalance and keeper-set bounds. LBank documents the family formula, but repeated observations found values identical to Binance with an approximately 30–100 second lag. We preserve the observation and the unknown mechanism separately.

3.2 Settlement schedules and event time

Discrete eight-hour settlement at 00:00, 08:00, and 16:00 UTC is only a baseline. Binance, OKX, Gate, and Bybit can accelerate after cap-related conditions; the step sizes and de-escalation conditions differ. Bitget, MEXC, BingX, and KuCoin allow contract-specific schedules or announcement-driven changes. Hyperliquid, ApeX, Lighter, Pacifica, Extended, and dYdX pay hourly. edgeX's measured field says 240 minutes even though two official pages stated eight hours and one hour. Deribit accrues continuously and nets daily at 08:00 UTC; GMX accrues per second; Paradex recomputes per second and accrues nothing across gaps longer than 30 seconds.

Table 2. Settlement models and the required interval evidence

Settlement model	Venues	Required treatment
Contract-level discrete intervals	Binance, OKX, Bybit, Bitget, Gate, KuCoin, MEXC, LBank, BingX, GRVT, Variational	Read interval per instrument; never impute eight hours
Fixed hourly	Hyperliquid, dYdX, edgeX (4 h measured), Extended, Lighter, Pacifica, ApeX	Use the documented or measured interval; edgeX is an exception to the hourly grouping
Continuous or per-second accrual	Deribit, GMX, Paradex	Integrate accrual; do not invent a discrete next-settlement timestamp
Externally timed windows	Variational	Read `interval_seconds`; zero early-window rates need not imply zero pressure
Pre-market and frozen states	Binance, OKX, Bybit, KuCoin, ApeX, dYdX and others	Filter by status before interpreting the rate as live

Note: edgeX is included in the second row for source grouping but its measured BTC interval is four hours. Official documentation and baseline measurements are dated 19–20 October 2026; refresh distributions appear in Section 6.

Clock fields add another layer. Most venues use milliseconds, but Gate REST uses seconds and some Gate WebSocket funding fields remain in seconds while frame clocks use milliseconds. ApeX WebSocket and Paradex SBE frames use microseconds; GRVT uses nanoseconds; Variational uses seconds. KuCoin’s `nextFundingRateTime` was observed to be a countdown rather than a timestamp. A parser that infers units from digit length alone is fragile because values can be null, omitted, or change precision.

3.3 Caps, position bases, and loss allocation

Outer caps determine extreme funding transfers. Binance and Bitget link caps to first-tier maintenance margin; Bybit uses a formula involving initial and maintenance margin with an adjustable coefficient; Hyperliquid and Pacifica document ±4% per hour; Variational documents ±2% per hour; GMX uses keeper-set APR bounds; dYdX links its eight-hour cap to the gap between initial and maintenance margin fractions. LBank’s cap depends on tiered maintenance margin and is not a public per-contract constant. ApeX states that amounts beyond ±0.05% per hour are absorbed by the platform rather than passed to users. These rules imply different tail bounds even when normal-state formulas coincide.

Funding cash flow also depends on the conversion price and contract size. Public documents variously specify mark, index, oracle, or fair price; others leave the basis undocumented. KuCoin, Gate, MEXC, and OKX use contract multipliers, so a raw contract count is not a coin quantity. OI can be base coin, quote notional, one-sided, two-sided, or per-side. Lighter exposes base-denominated OI in one view and quote-valued OI under the same field name in another. Variational publishes per-side USDC values that must not be summed or halved without an explicit economic justification.

Base-tier taker fees range from zero at Lighter and Variational to 0.075% at Gate in the baseline evidence. Maker fees range from a 0.01% rebate at Gate to 0.04% at edgeX. GMX instead charges position fees and price impact. Liquidation and backstop mechanisms include insurance-fund–ADL sequences, liquidity-provider pools, socialized

losses, random deleveraging, and RFQ-pool absorption. These rules belong to the cost and tail-risk data dictionary, but account-level realized charges remain outside the public-read-only scope.

4. Data

4.1 Venue universe and evidence unit

The study is a census of the 21 venues covered by the research project, not a random sample of the global perpetual-futures industry. It contains 10 CEXs—Binance, OKX, Bybit, Bitget, Gate, KuCoin, MEXC, LBank, BingX, and Deribit—and 11 on-chain or hybrid venues—Hyperliquid, dYdX, edgeX, Extended, GMX, GRVT, Lighter, Pacifica, Paradex, Variational, and ApeX. Coinbase International, Kraken, Crypto.com, HTX, BitMEX, Aster, and other venues are outside scope. Consequently, counts are descriptive of this census and are not extrapolated to the industry.

The baseline evidence unit is a venue-by-checklist item. For each venue, researchers reviewed 12 mechanism categories: funding formula; settlement interval; settlement instant; mark and index prices; contract specifications; trading fees; order-book interfaces; OI and turnover; liquidation and maintenance margin; listing, delisting, halt, and pre-market states; rate limits; and special handling. A second checklist records price constraints, quantity rules, status fields, trading hours, and deposit/withdrawal state. The resulting venue dictionary has 29 columns, each defined in the replication package.

Three evidence labels are preserved. **Documented** means that official text states the rule. **Measured** means that a read-only public response contains the field or value at a recorded time. **Inference** means that the interpretation follows from consistency checks rather than an explicit official sentence. Examples of inference include Binance’s current-forecast reading of `lastFundingRate`, KuCoin’s family membership from isomorphic parameters, and MEXC’s interest value inferred from measured pinning. A failed request is stored as a refresh failure, never as a missing rule.

4.2 Baseline documents and public measurements

Baseline documents and public responses were collected on 19–20 October 2026. Official mechanism pages supply formulas, schedules, fees, and liquidation rules. Public endpoints supply instrument rows, funding parameters, status flags, and market-data fields. No credentials, private endpoints, account permissions, or orders were used. A separate researcher independently reviewed pre-trade fields for 18 venues, and Binance, Bybit, and OKX price-limit rules received an additional comparison. This process corrected an OKX endpoint-path error and distinguished limit-order filters from market-order protection at Binance.

The instrument universe includes active, inactive, pre-market, reduce-only, settling, and delisted rows when the public endpoint returns them. This retention rule reduces survivorship within the returned universe, but cannot recover instruments already absent from an endpoint. It also creates an explicit need for status filters before a rate is interpreted as tradable. Counts in Figure 2 describe all returned rows and are not presented as counts of currently tradable assets.

4.3 Refresh snapshot and history sample

We attempted a second low-frequency public snapshot on 3 November 2026. Eleven venues produced complete raw JSON responses through the reproduction script. Bybit, MEXC, and Deribit were additionally verified through public web-interface samples. Seven venues could not be refreshed in this environment because of network unreachability, timeout, access restriction, or a non-JSON response. The manifest stores each URL, UTC start and finish time, HTTP status where available, elapsed time, raw-file path, and error. It does not silently convert errors into empty datasets.

Table 3. Evidence coverage by stage

Evidence component	Coverage	Date or period
Venue mechanism census	21/21 venues	19–20 Oct 2026
Pre-trade field maps	21/21 venues	20 Oct 2026
Complete raw refresh responses	11 venues	3 Nov 2026
Additional public web-interface samples	3 venues	3 Nov 2026
Venues not refreshed in this environment	7 venues	3 Nov 2026
Funding-history venues in expanded timestamp sample	4 venues	29 Sep–3 Nov 2026, varying by venue
Funding-history records	1,840 rows	Retrieved 3 Nov 2026

Note: Refresh success describes access from the research environment, not venue uptime or document existence. Full request-level details are in `public-snapshot-2026-11-03/manifest.json`.

The history sample contains 500 Hyperliquid rows from 29 September to 20 October, 400 Pacifica rows from 17 October to 3 November, 840 Extended rows from 29 September to 3 November, and 100 edgeX rows from 17 October to 3 November. Hyperliquid’s endpoint returned at most 500 rows, so its coverage ends earlier than the retrieval date. Pacifica and Extended use hourly records; edgeX uses four-hour settlement records. We do not pool these rows as if they were identically distributed.

4.4 Missing data, anomalies, and reproducibility

“Not documented” and “not located” are distinct from zero. LBank has no public settled-history endpoint in the baseline evidence, so a settled series is not backfilled. Bybit’s dynamic price-limit coefficient is not exposed in the examined public instrument response, so it is not assumed to be observable. GMX’s documented paying-side rule conflicted with one measurement; both are retained and the cause remains unresolved. MEXC fee responses previously differed across hosts and dates; the refresh BTC row again reports 0.02% taker and zero maker, but this single row does not resolve the earlier host-wide discrepancy.

The reproduction package contains the curated CSV/JSON dictionaries, raw responses for the 11 complete refresh venues, history responses for four venues, the request manifests, extraction scripts, and the analysis script that regenerates Table 5 and Figures 1–4. Public responses are retained for auditability, not redistributed as a

substitute for official documentation. Venue terms and rate limits apply to any new collection; the scripts are deliberately sequential and low frequency.

5. Methods

5.1 Mechanism normalization tuple

For venue v, instrument i, and observation time t, define the mechanism tuple

$$\mathcal{M}_{vit} = (S_{vit}, U_{vit}, h_{vit}, \tau_{vit}, Q_{vit}, Z_{vit}).$$

Here S is semantic phase (next-period forecast, last settled, trailing realized, or current accrual), U is numerical unit, h is accrual or settlement interval, τ is event-time convention, Q is position and cash-flow basis, and Z is trading and settlement status. A raw rate r_{vit} is not treated as comparable until the relevant tuple elements are known.

Let k(U) convert the published number to a decimal rate: k=1 for decimal units and k=0.01 for percentage points. For a discrete rate measured over h hours, the linear hourly-equivalent rate is

$$\tilde{r}_{vit} = \frac{k(U_{vit})\, r_{vit}}{h_{vit}}.$$

A linear annualized display uses $8{,}760\tilde{r}_{vit}$. We use linear scaling for deterministic error examples because it makes the interval mistake transparent; compound annualization would answer a different question and does not cure a wrong h. Rates from different S categories are not subtracted merely because Equation (4) places them on an hourly scale.

For a linear perpetual, a signed funding cash-flow illustration is

$$C_{vit} = -s_{vit}\, q_{vit}\, m_{vi}\, p_{vit}\, k(U_{vit})\, r_{vit},$$

where s=+1 for a long and -1 for a short, q is position quantity in the venue’s native unit, m is the contract multiplier, and p is the documented conversion price. Inverse, quanto, pool-based, and per-side designs require their own formulas; Equation (5) is not imposed on them. The tuple therefore defines when a common equation is valid and when a venue-specific equation is required.

5.2 Deterministic error measures

We evaluate four deterministic mismatches. The interval error from treating true interval h as $\hat{h}$ is

$$e_h = kr\left(\frac{1}{\hat{h}} - \frac{1}{h}\right),$$

in hourly-equivalent units. The unit error from using $\hat{k}$ instead of k is $e_U=(\hat{k}-k)r$. Formula error is the difference between two documented formulas evaluated at the same premium and interest assumptions. Semantic error is the difference between fields associated with different funding events; its sign and magnitude are time-varying, so we do not assign a universal bound.

For economic scale, we report deterministic examples per $10,000 of position notional. A 0.01% rate corresponds to $1 per period. Reading 0.01 percentage points as decimal 0.01 produces $100 rather than $1, an absolute difference of $99. Treating a 0.01% four-hour rate as eight-hour changes the daily estimate from $6 to $3. The clamp example changes one period's payment by $1. A zero–0.075% taker-fee range corresponds to a $15 difference for one $10,000 entry-and-exit leg. These are arithmetic illustrations, not observed profits or losses.

5.3 Timestamp-offset estimator

For a public record timestamp t_j in milliseconds, define the signed offset to the nearest nominal hour as

$$\delta_j = ((t_j + 1{,}800{,}000) \bmod 3{,}600{,}000) - 1{,}800{,}000.$$

We report n, exact-zero count, mean, median, 95th percentile, and maximum of $|\delta_j|$ by venue. The nearest-hour convention accommodates records shortly before or after a boundary. edgeX uses four-hour settlement, but hourly modulo remains valid because each four-hour boundary is also an hourly boundary.

The estimand is a property of the public record timestamp. It is not a direct estimate of the exchange's internal settlement instant or of cash availability. Pacifica's `created_at` and Extended's calculation/application timestamp are especially important to distinguish from guaranteed ledger time. Paradex rate snapshots are excluded because they are not settlement events.

5.4 Evaluation and robustness design

The main conclusions are census descriptions and deterministic calculations, so sampling standard errors, Newey–West corrections, predictive sample-out-of-sample tests, Diebold–Mariano tests, and multiple-hypothesis adjustments are not applicable to them. No forecasting model, trading rule, or significance claim is estimated. We therefore use four robustness checks appropriate to the evidence: (i) repeat instrument distributions on a second date; (ii) compare CEX and on-chain subgroups without pooling their mechanism classes; (iii) vary interval, unit, and formula assumptions in the error calculations; and (iv) compare mean, median, tail, and exact-zero timestamp summaries. An alternative coding of Variational tests whether the core count depends on the most ambiguous venue.

Table 4. Principal symbols

Symbol	Definition
r_{vit}	Venue-native published funding-rate field
$\mathcal{M}_{vit}$	Mechanism tuple: semantics, units, interval, time, position basis, status
S,U,h,τ,Q,Z	Components of the mechanism tuple
k(U)	Conversion from published unit to decimal
$\check{r}$	Hourly-equivalent decimal funding rate
q,m,p	Native position quantity, contract multiplier, conversion price
δ_j	Public record timestamp offset from the nearest nominal hour

6. Results

6.1 Mechanism heterogeneity is structural

Figure 1 summarizes the measurement chain. Venue-native fields must pass through mechanism normalization before they become inputs to a rate differential, cash-flow estimate, or execution-cost model. The largest formula group shares Equation (1), but this does not create a single homogeneous sample: interval scaling, interest terms, premium construction, caps, and event time still vary. The three singleton formula structures and three additional outliers demonstrate why formula family must be part of the dictionary rather than inferred from the rate’s magnitude.

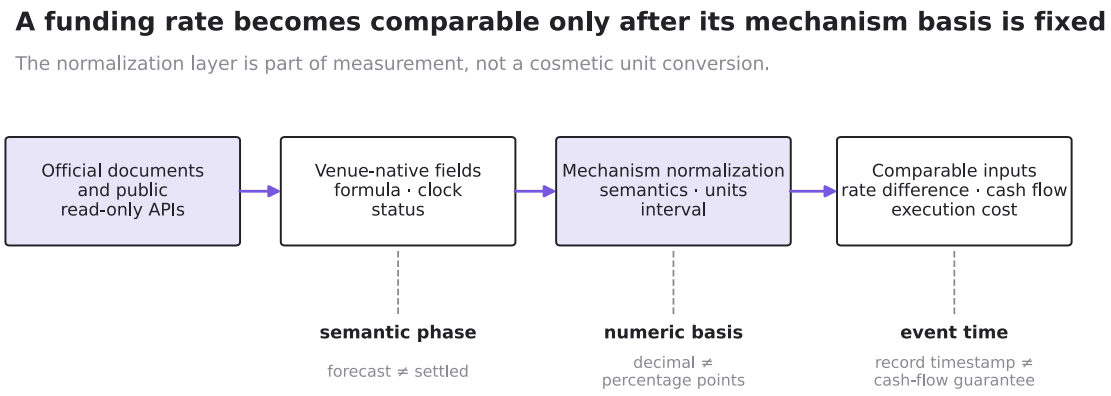


Figure 1. Mechanism normalization pipeline. Source: authors' framework.

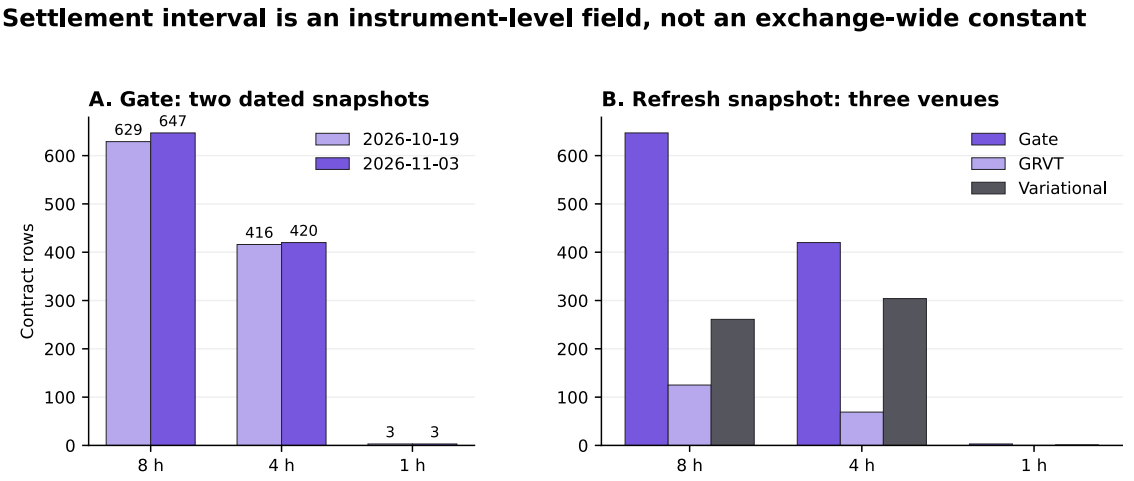
The MEXC comparison isolates formula position from market noise. At the stated premium, interest, and clamp assumptions, the difference is 0.01 percentage point per period. Deribit's deadband creates a different normal-state rate from the interest-floor behavior of Equation (1). Paradex's smoothing and pause rule mean that multiplying a current value by elapsed window time can overstate accrual when oracle gaps occur. GMX's imbalance-driven factor and separate borrowing fee make a single funding field insufficient for cost measurement.

LBank illustrates a different problem: dependence between venues. Bit-identical values with a short lag indicate that two displays need not represent independent price discovery or an executable spread. We do not infer an internal replication method. The measurement implication is narrower: same-instant subtraction can create a transient nonzero difference caused by update timing, while the economic independence of the two series should not be assumed.

6.2 Instrument-level intervals dominate the eight-hour default

The baseline already rejects a universal eight-hour assumption: Binance, Gate, Bitget, MEXC, and LBank each returned hundreds of four-hour contracts in their respective snapshots. The refresh confirms that this is not a one-day artifact. Gate returned 647 eight-hour, 420 four-hour, and 3 one-hour rows on 3 November, compared with 629, 416, and 3 on 19 October. The four-hour share was 39.25% in the refresh and 39.69% in the baseline. Total rows increased by 22, while the interval mix remained similar.

GRVT's refresh response contained 125 eight-hour and 69 four-hour perpetual rows. Variational contained 261 eight-hour, 304 four-hour, and 1 one-hour perpetual rows. Thus, heterogeneity appears in both a CEX and on-chain or hybrid venues. Figure 2 plots the distributions. Because counts include all returned rows, they measure interface composition rather than the active-trading universe.



Source: public contract endpoints; baseline 2026-10-19 and refresh 2026-11-03. Counts include all returned rows, not only active BTC markets.

Figure 2. Settlement-interval distributions. Source: public Gate, GRVT, and Variational contract endpoints, 19 October and 3 November 2026.

Interval mistakes affect both level and timing. At a constant per-period rate, treating four hours as eight halves the daily estimate. Cap-hit acceleration can additionally move the cash-flow event itself. Hourly payment of an eight-hour-computed rate has the opposite trap: comparing the displayed hourly value directly with an unscaled eight-hour value overstates the differential by a factor of eight. Correct normalization therefore requires both the formula's realization basis and the actual settlement interval.

6.3 Semantics, units, and status determine whether a rate is live

Semantic phase is the first hidden source of differential error. Subtracting edgeX's last-settled field from Binance's current-forecast field creates a next-period-minus-previous-period comparison. Deribit's trailing eight-hour value is one full phase away from a next-period forecast. dYdX's `nextFundingRate`, by contrast, refers to the next hourly event and therefore need not imply an additional phase gap. Extended's `nextFundingRate` is a next-settlement timestamp rather than a rate. These examples show why a schema based on names alone is not a semantic normalization layer.

Unit errors are more mechanical but potentially larger. GRVT and Lighter publish percentage-point rates, while many CEXs publish decimals. A value of 0.01 in percentage points is 0.01%, not 1%. Timestamp units range from seconds to nanoseconds, and KuCoin's countdown field can be mistaken for a 1970 timestamp. Unlike sampling variation, these errors persist with higher-frequency collection.

Status fields determine whether a number should enter the live universe at all. Baseline measurements found frozen rates for settling Binance contracts, funding-disabled ApeX rows, and a large number of dYdX markets in final settlement. Pre-market rules can force zero or fixed rates that are not comparable with ordinary market-driven funding. The dictionary therefore treats Z as a required field. Missing status is not permission to assume active trading.

6.4 Pre-trade rules require a separate dictionary

Funding normalization does not make execution costs comparable. Price bands, quantity increments, contract multipliers, market-order protection, trading sessions, and transfer state determine whether a calculated position can be entered and exited. Table 5 groups the main pre-trade patterns.

Table 5. Pre-trade rule classes in the 21-venue census

Rule class	Representative venues	Measurement consequence
Public relative price bands	Binance, Gate, Bitget, MEXC, LBank, edgeX, Extended, Paradex	Store reference price, ratio, and order type separately
Public absolute price limits	OKX, KuCoin	Timestamp the absolute bound; do not treat it as a fixed parameter
Protocol or oracle protection without a per-market band	Hyperliquid, dYdX, Lighter, GMX	Use venue-specific rejection or price-impact logic
RFQ quotes rather than a conventional fillable book	Variational; many Extended markets	Indicative depth cannot enter depth-walked cost as executable liquidity
Contract-count quantity systems	OKX, Gate, MEXC, KuCoin	Multiply by contract value before computing notional
Public trading-hours fields	Extended	Preserve `tradingHours` and `isOffHours` in event alignment
Documented off-hours funding continuation	Pacifica	Do not suspend funding accrual merely because the underlying is closed

Note: “Representative” does not imply that other venues lack rules. It indicates the distinct public-field patterns documented on 19–20 October and refreshed where available.

The refresh provides a concrete example of schema drift. Extended’s current BTC row exposes `minOrderSize`, `minOrderSizeChange`, `limitPriceCap`, `limitPriceFloor`, and `maxMarketOrderValue` under `tradingConfig`. Earlier evidence had not consistently exposed these fields in the same public view. This change supports storing raw responses and field-level timestamps rather than overwriting a static schema.

Book semantics are equally important. Extended’s baseline default books for RFQ markets were indicative rather than directly fillable. In the refresh, 340 of 416 perpetual rows are marked RFQ, although not all are active. Variational’s RFQ quote sizes define the available liquidity schedule. Binance documents that retail-price-improvement orders are excluded from ordinary depth responses, so visible depth need not exhaust executable liquidity. The direction of the resulting cost bias differs across these cases: indicative depth can invent liquidity, while hidden eligible orders can conceal it.

6.5 Funding timestamps are not homogeneous event markers

Table 6 reports the expanded timestamp analysis. edgeX provides the sharpest contrast: all 100 settlement records are exactly on the nominal hour. Hyperliquid’s median absolute offset is 36 ms and its 95th percentile is 89 ms. Extended’s corresponding values are 937 ms and 1,611 ms; Pacifica’s are 3,680 ms and 4,046 ms. Mean absolute offsets are 39 ms, 995 ms, and 3,507 ms, respectively, so the main ordering is not driven by a single extreme row.

Table 6. Absolute offsets of public funding timestamps from the nearest nominal hour

Venue	Records	Exact zero	Mean (ms)	Median (ms)	95th pct. (ms)	Maximum (ms)
Hyperliquid	500	14	39.2	36	89	139
Pacifica	400	0	3,507.2	3,680	4,046	4,046
Extended	840	0	995.1	937	1,611	2,094
edgeX	100	100	0	0	0	0

Note: Retrieved 3 November 2026. Coverage differs by venue and is listed in Section 4.3. The estimand is record-timestamp offset, not internal settlement or cash-availability latency. Computed by `analyze_snapshots.py`.

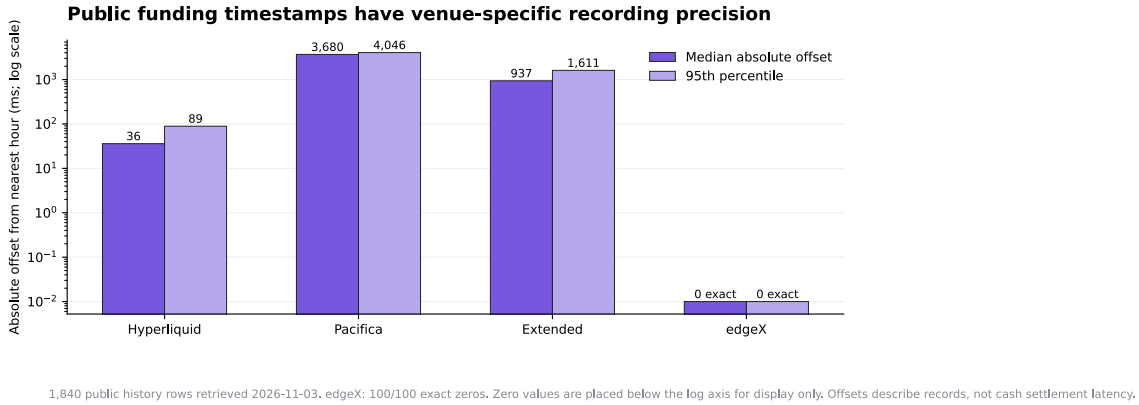
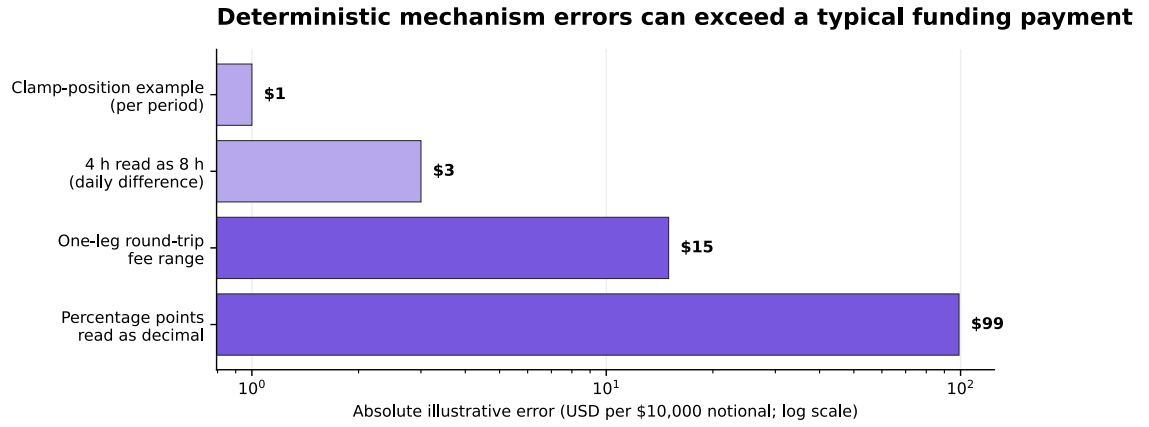


Figure 4. Timestamp-offset summaries. Source: 1,840 public funding-history records, retrieved 3 November 2026.

The practical implication is not that every venue needs the same tolerance. It is that tolerance is a venue- and field-specific parameter. A single zero-tolerance rule would discard every Pacifica and Extended record in this sample; a multi-second rule would be unnecessarily coarse for Hyperliquid and edgeX. Any alignment across venues should preserve the raw timestamp and document whether it denotes calculation, application, snapshot, or record creation.

6.6 Deterministic errors are economically non-trivial

Figure 3 places four illustrative errors on a common $10,000-notional scale. The unit misread dominates at $99 per period, followed by the maximum base-tier one-leg round-trip taker-fee difference of $15, the daily interval error of $3, and the clamp-position difference of $1. These quantities are not additive by construction and should not be read as a complete cost budget. Spread, depth-walked impact, collateral transfer, borrowing, and liquidation losses are additional.



Illustrations, not realized P&L: 0.01% per-period clamp difference; 0.01% per 4 h annualization; 0–0.075% taker fee range; 0.01% read as 1%.

Figure 3. Illustrative measurement-error magnitudes. Source: deterministic calculations using public rules and base-tier fee evidence.

The appropriate conclusion is limited but useful. Mechanism errors can be larger than the funding payment a researcher is trying to measure, and they do not average away. A high-resolution time series with a wrong unit or interval is precisely wrong. The dictionary and normalization tuple are therefore prerequisites for economically meaningful cross-venue empirics.

7. Robustness

7.1 Repeated-date instrument distributions

The principal interval result survives a repeated snapshot. Gate’s four-hour count rises from 416 to 420, while the total row count rises from 1,048 to 1,070. The four-hour share changes by only −0.44 percentage points. The eight-hour count rises from 629 to 647, and the one-hour count remains three. Thus, the conclusion does not depend on a particular collection day or on interpreting one unusual listing batch as a permanent rule. GRVT and Variational provide independent same-day evidence that mixed four- and eight-hour schedules are not unique to Gate.

This check addresses composition stability, not rule permanence. A venue can change intervals by announcement, governance, or automatic cap response. Accordingly, the paper treats the dictionary as a dated schema plus dated observations rather than a permanent parameter file. A production or research collector should refresh instrument fields before using them.

7.2 Venue groups and alternative taxonomy coding

Separating CEX and on-chain or hybrid venues does not remove heterogeneity. Five CEXs have measured non-eight-hour instrument populations; on-chain or hybrid venues include hourly settlement, four-hour settlement, continuous accrual, external-window following, and imbalance-driven designs. The shared clamped core appears in both groups, as do instrument-level schedules. Venue architecture alone is therefore not a sufficient proxy for funding mechanics.

We also test the effect of ambiguous class boundaries. Variational is coded in the main taxonomy as a clamped-core venue with an external-window timing subclass. Moving it

out of the 15-venue core would change that count to 14 but leave the formula variation and measurement conclusions unchanged. LBank remains a separate published-mirror class despite its documented core formula because the dependence of its published series is the measurement-relevant feature. KuCoin remains explicitly inferential. The results therefore do not rely on silently resolving borderline venues in the direction of a cleaner taxonomy.

7.3 Parameter and formula sensitivity

The deterministic examples are recomputed under alternative plausible assumptions rather than treated as universal constants. For the formula comparison, P=0.0005, I=0.0001, and c=0.0005 yield 0.0005 under Equation (1) and 0.0006 under Equation (2). The 0.0001 difference follows from clamp location. It is not a claim that every MEXC period differs by that amount: when the relevant clamp is inactive, the formulas can coincide.

For interval scaling, a 0.01% per-period rate produces six daily payments at four hours and three at eight hours. Holding the displayed per-period value fixed while changing only the interval assumption produces the $3 daily difference shown in Figure 3. If the venue instead publishes an interval-normalized rate, the correct conversion follows the documented realization basis, not a generic multiplier. This alternative illustrates why the dictionary stores both `default_interval` and `interval_scaling`.

For units, the ratio between percentage-point and decimal interpretation is 100 regardless of notional. For timestamps, changing the summary from the mean to the median reduces the influence of large observations but leaves the venue ordering intact. These checks vary the assumptions that drive the paper’s conclusions; they do not add a statistical robustness label to deterministic arithmetic.

7.4 Alternative timestamp summaries and coverage

Table 6 reports six summaries per venue, and Figure 4 emphasizes medians and 95th percentiles. Hyperliquid has 14 exact-zero records, so its mean exceeds its median only modestly: 39.2 versus 36 ms. Pacifica’s mean and median are similarly close, at 3,507 and 3,680 ms. Extended has a wider upper tail, with a 2,094 ms maximum versus a 937 ms median. edgeX’s complete exact-zero pattern is invariant to every reported location or tail measure. The central cross-venue contrast is therefore stable across estimators.

Coverage is less easily interchangeable. Hyperliquid’s 500-row cap gives it a shorter history window than Extended’s 840 rows, and edgeX’s 100 four-hour records cover a different event frequency from the hourly venues. We therefore do not rank venues with a pooled significance test or claim population equality. The defensible conclusion is descriptive: within these observed public records, timestamp conventions differ materially and should not share a blind zero-tolerance filter.

8. Limitations

First, the venue census is bounded by the project’s 21-venue scope. It is neither a random sample nor an exhaustive map of all perpetual-futures venues. The taxonomy can be extended, but counts should not be extrapolated to the global industry. New contract types, particularly real-world-asset and event-linked perpetuals, may introduce mechanisms absent from this census.

Second, public documents and interfaces are moving targets. The baseline was collected on 19–20 October 2026 and the refresh on 3 November 2026. Seven venues were not refreshed because of access or network conditions. Even for successful venues, a current response can differ from an older official help page, as the Extended schema and edgeX interval evidence illustrate. Every numerical field should therefore be read with its access timestamp. The paper does not assert that a measured fee, cap, or interval remains current after the snapshot.

Third, formula and field semantics are not always fully documented. KuCoin’s family membership, some position-unit interpretations, and LBank’s publication dependence require inference. We preserve these labels rather than presenting inferred rules as official text. Unrecognized semantic drift remains possible. Account-specific fees, actual charged amounts, private rate limits, and liquidation or ADL notifications cannot be observed through the public read-only design.

Fourth, timestamp analysis is limited to four venues and unequal coverage windows. Public timestamps can denote calculation, application, record creation, or publication. They do not identify the exchange’s internal settlement process or guarantee when funds become withdrawable. No causal latency ranking or cash-settlement service-level claim is made. A more complete study would collect longer, time-synchronized histories and obtain explicit ledger-time definitions.

Fifth, the economic examples are deterministic illustrations, not realized execution results. No orders were placed, no account credentials were used, and no executable bid/ask or liquidation sequence was reconstructed. RFQ and indicative-book findings establish a measurement warning, not a complete estimate of available liquidity. We do not estimate funding-return predictability, arbitrage profitability, Sharpe ratios, or failure frequencies.

Sixth, the dictionary is deliberately selective rather than exhaustive. It prioritizes fields that can change the sign, scale, timing, or implementability of a cross-venue measurement. It does not encode every contract specification, jurisdiction-specific eligibility rule, portfolio-margin interaction, insurance-fund accounting entry, or venue-specific error code. Consequently, a complete row in the dictionary means coverage of the stated 29-field schema, not exhaustive due diligence on a venue. Researchers adding a new economic object—such as options on perpetual contracts, unified-margin collateral, or structured products—should extend the schema rather than force the new object into an existing field. This boundary also limits cross-study comparisons: another dictionary with fewer fields may appear complete while omitting precisely the mechanism distinction that drives a measurement error.

Finally, reproducibility depends on public access remaining available and permitted. Collection scripts use sequential, low-frequency requests, but researchers must review

current venue terms and rate limits before running them. Raw-response schemas can change without backward compatibility. The package mitigates this risk by preserving dated raw files and manifests, not by promising a permanent live connector.

9. Conclusion

A funding rate becomes a comparable financial measurement only after its mechanism is known. Across the 21 venues in this study, the common product label conceals seven formula classes, instrument-specific settlement intervals, multiple rate phases and numerical units, continuous and discrete accrual, distinct position bases, and incompatible trading-status conventions. Pre-trade rules add another layer: contract multipliers, dynamic price limits, RFQ liquidity, trading hours, and transfer availability determine whether a calculated opportunity can be implemented.

The empirical results make the problem concrete. Gate’s refresh contains 420 four-hour contracts among 1,070 rows; GRVT and Variational also contain mixed schedules. Among 1,840 public history records, median absolute timestamp offsets range from exact zero at edgeX to 3,680 ms at Pacifica. Deterministic examples show that unit, interval, formula, and fee mismatches can be large relative to the funding payment being measured. These are measurement findings, not claims of trading profitability.

The resulting data dictionary and normalization tuple provide a practical sequence: establish semantic phase, unit, interval, event time, position basis, and status before constructing a cross-venue rate or cash-flow variable. This sequence complements the net-carry frameworks in humarb Research 01 and 02 by addressing their antecedent input-comparability problem. Better downstream models begin with correctly defined inputs.

Declarations

Author contributions. Hang Liu conceived and designed the study, curated the venue evidence, developed the normalization framework, interpreted the results, and approved the manuscript.

Competing interests. The author is affiliated with humarb Research. No other competing interests are declared.

Funding. No external grant funding is declared for this work.

Data availability. The curated 21-venue dictionaries, successful public refresh responses, funding-history extracts, timestamp-derived data, and request manifests are included in the accompanying reproduction package. Failed refreshes are identified in the manifests. Public venue rules remain subject to their original terms and change over time. An archival DOI for this paper and package will be added after publication.

Code availability. The package contains scripts for public snapshot collection, history collection, dictionary construction, descriptive analysis, and figure generation. The scripts require no credentials and submit no orders. They are provided for low-frequency research use, subject to venue terms and rate limits.

Ethics. The study uses public market documentation and read-only public market data. It involves no human participants, private user records, or animal subjects.

AI-assisted research disclosure. AI-based tools assisted with code drafting, language editing, and literature discovery. The author is responsible for source verification, calculations, interpretation, and the final manuscript. No confidential code, credentials, private endpoints, user data, or undisclosed production parameters are included.

Research status and non-advice statement. This is a working research paper, not a representation of peer review. It is provided for research and methodological purposes and is not investment advice, a solicitation, or a claim of guaranteed returns.

References

Ackerer, D., Hugonnier, J., and Jermann, U. J. (2026). Perpetual futures pricing. *Mathematical Finance*. https://doi.org/10.1111/mafi.70018 (earlier version: arXiv:2310.11771).

Angeris, G., Chitra, T., Evans, A., and Lorig, M. (2022). A primer on perpetuals. arXiv:2209.03307. Published version (2023), *SIAM Journal on Financial Mathematics*. https://doi.org/10.1137/22M1520931.

Chitra, T. (2026). Autodeleveraging: Impossibilities and optimization. arXiv:2512.01112, version 3.

Chitra, T., Thogiti, N., Trujillo Ramirez, M. J. P., and Xu, V. (2026). Autodeleveraging as online learning. arXiv:2602.15182.

Gupta, A., and Polson, N. (2026). Perpetual futures for stocks: The SpaceX pre-IPO market. arXiv:2609.05433.

He, S., Manela, A., Ross, O., and von Wachter, V. (2022). Fundamentals of perpetual futures. arXiv:2212.06888 (version consulted: 2026).

Kim, J., and Park, H. (2025). Designing funding rates for perpetual futures in cryptocurrency markets. arXiv:2506.08573.

Liu, H. (2026a). *Measuring Executable Net Carry in Cross-Venue Perpetual Funding-Rate Arbitrage*. humarb Research 01. https://doi.org/10.5281/zenodo.23868361.

Liu, H. (2026b). *Framing Net Carry for Cross-Venue Perpetual Funding Positions: A Settled-History, Depth-Walked, and Liquidation-Aware Method*. humarb Research 02. https://doi.org/10.5281/zenodo.23897013.

Nechepurenko, M. (2026). A taxonomy of event-linked perpetual futures. arXiv:2605.10428.

Sepper, O. (2026). Slippage-at-Risk (SaR): A forward-looking liquidity risk framework for perpetual futures exchanges. arXiv:2603.09164.

Shiller, R. J. (1993). Measuring asset values for cash settlement in derivative markets: Hedonic repeated measures indices and perpetual futures. *The Journal of Finance*, 48(3), 911–931. https://doi.org/10.1111/j.1540-6261.1993.tb04024.x.

Official venue documentation

The following official documentation and public API references were accessed on 19–20 October 2026 unless a refresh is explicitly stated. Article titles and URLs are listed here; the reproduction package supplies the dated public responses rather than reproducing documentation pages.

ApeX. *ApeX Omni: Trading and Funding Fees; Margin, Risk Management and Liquidations; ADL*. https://apex.exchange/ and https://api-docs.omni.apex.exchange/.

Binance. *Introduction to Binance Futures Funding Rates; What Are Mark Price and Price Index in USD□-M Futures; USD□-M Futures API*. https://www.binance.com/en/support/faq/360033525031 and https://developers.binance.com/.

BingX. *Perpetual Futures API Reference; Funding Rate and Trading Rules*. https://bingx.com/en-us/ and https://github.com/bingx-exchange/open-api-sdk-v2.

Bitget. *What Is the Funding Rate in Bitget Futures Trading?; Contract API*. https://www.bitget.com/support/articles/12560603817108 and https://www.bitget.com/api-doc/contract/intro.

Bybit. *Introduction to Funding Rate; V5 Market API*. https://www.bybit.com/en/help-center/article/Introduction-to-Funding-Rate and https://bybit-exchange.github.io/docs/v5/intro. BTC instrument fields refreshed 3 November 2026.

Deribit. *Funding Specifications; Fees; Public API*. https://support.deribit.com/hc/en-us/articles/31424939178397 and https://docs.deribit.com/. BTC instrument and fee fields refreshed 3 November 2026.

dYdX. *Funding; Margining; Liquidations; Contract Loss Mechanisms; Oracle Prices*. https://docs.dydx.xyz/concepts/trading/.

edgeX. *Funding Fees; Funding Rate API*. https://edgex-1.gitbook.io/edgeX-documentation/trading/funding-fees.md and https://pro.edgex.exchange/. Instrument and history responses refreshed 3 November 2026.

Extended. *Funding Payments; Trading Fees and Rebates; Liquidation Logic; Public API*. https://docs.extended.exchange/ and https://api.starknet.extended.exchange/. Instrument and history responses refreshed 3 November 2026.

Gate. *Futures Funding Rate and Funding Fee Overview; Futures API*. https://www.gate.com/help/futures/futures-logic/27569 and https://www.gate.com/docs/developers/apiv4/en/. Instrument response refreshed 3 November 2026.

GMX. *Fees; Liquidations; Oracle Prices*. https://docs.gmx.io/docs/trading/. Public market information refreshed 3 November 2026.

GRVT. *Funding Rate Mechanism; Fee Model; Public API and SDK*. https://help.grvt.io/en/articles/12555039 and https://api-docs.grvt.io/. Instrument response refreshed 3 November 2026.

Hyperliquid. *Funding; Contract Specifications; Liquidations; Auto-Deleveraging; Info API*. https://hyperliquid.gitbook.io/hyperliquid-docs. Instrument and history responses refreshed 3 November 2026.

KuCoin. *Funding Rate Tutorial; Liquidation and Liquidation Price; Insurance Fund; ADL; Futures API*. https://www.kucoin.com/support/26695933047705 and https://www.kucoin.com/docs-new/rest/futures-trading/market-data/get-symbols-list.

LBank. *Funding Rate: Definition, Components, and Uses; Futures API*. https://www.lbank.com/support/articles/21422922435353 and https://www.lbank.com/en-US/docs/futures.html. Public responses refreshed 3 November 2026.

Lighter. *Funding; Trading Fees; Liquidations and LLP; Public API*. https://docs.lighter.xyz/trading/ and https://apidocs.lighter.xyz/. Public responses refreshed 3 November 2026.

MEXC. *Funding Rate Information; Contract API*. https://www.mexc.com/funding-rate and https://mexcdevelop.github.io/apidocs/contract_v1_en/. BTC contract fields refreshed 3 November 2026.

OKX. *Introduction to Perpetual Swap Funding Fee; Pre-Market Product Rule; Public Instruments API*. https://www.okx.com/help/iv-introduction-to-perpetual-swap-funding-fee and https://www.okx.com/docs-v5/en/.

Pacifica. *Funding Rates; Liquidations; Margin and Leverage; Public API*. https://docs.pacifica.fi/ and https://api.pacifica.fi/api/v1/info. Instrument and history responses refreshed 3 November 2026.

Paradex. *Funding Mechanism; Trading Fees; Liquidations; Public API*. https://docs.paradex.trade/ and https://api.prod.paradex.trade/. Public responses refreshed 3 November 2026.

Variational. *Funding Rates; Fees; Automatic Deleveraging and Counterparty Liquidation; Public API*. https://docs.variational.io/ and https://omni.variational.io/api/. Public responses refreshed 3 November 2026.

Appendix A. Venue coverage and mechanism map

Table A1 is a compact navigation layer for the full 29-column CSV/JSON dictionary. “Default interval” is not a substitute for the per-instrument `interval_source`. Continuous and external-window rules are written in words where a numeric hour value would be misleading. `full_raw` denotes a complete response saved by the refresh script; `web_sample` denotes a smaller public interface verification; all other statuses are refresh failures and do not imply missing venue rules.

Table A1. Compact 21-venue mechanism and refresh map

Venue	Type	Formula class	Default interval (h)	Refresh status
Binance	CEX	Clamped core	8	failed
OKX	CEX	Clamped core	8	failed
Bybit	CEX	Clamped core	8	web_sample
Bitget	CEX	Clamped core	8	failed
Gate	CEX	Clamped core	8	full_raw
KuCoin	CEX	Clamped core (inference)	8	failed_non_json
MEXC	CEX	Premium clamp	8	web_sample
LBank	CEX	Published mirror	8	full_raw
BingX	CEX	Clamped core	8	http_403
Deribit	CEX	Deadband accrual	Continuous; daily net 08:00 UTC	web_sample
Hyperliquid	On-chain	Clamped core	1	full_raw
dYdX	On-chain	Validator-voted premium	1	failed
edgeX	On-chain	Clamped core	4 measured	full_raw
Extended	On-chain	Clamped core	1	full_raw
GMX	On-chain	Imbalance-driven state	Per second	full_raw
GRVT	On-chain	Clamped core	8	full_raw
Lighter	On-chain	Clamped core	1	full_raw
Pacifica	On-chain	Clamped core	1	full_raw
Paradex	On-chain	Cross-venue EWMA	Per second; pause rule	full_raw
Variational	On-chain	Clamped core; external-window timing	1–8	full_raw
ApeX	On-chain	Clamped core	1	failed

The dictionary additionally records formula text, interest defaults, interval scaling, semantic phase, numerical and clock units, caps, settlement price bases, OI and turnover conventions, base-tier fees, liquidation backstops, price and quantity rules, status fields, trading hours, transfer state, settled-history availability, primary official URL, access dates, and evidence qualifications. Empty or undocumented cells are not silently imputed.

Appendix B. Minimum normalization checklist

The following eight questions operationalize the paper’s central argument. They should be answered for each venue-instrument-time observation before a cross-venue dataset is released.

Table B1. Field-level audit before cross-venue comparison

Check	Required action	Failure prevented
Formula	Record the venue's own formula, clamp location, interest term, and premium construction	Applying a Binance-family formula to MEXC, Deribit, dYdX, Paradex, or GMX
Semantic phase	Classify forecast, last settled, trailing realized, and current accrual fields	Subtracting rates associated with different funding events
Numeric unit	Convert decimal and percentage-point fields explicitly	A 100-fold rate and cash-flow error
Accrual and settlement	Read interval per instrument and preserve formula scaling	Incorrect annualization and incorrect event counts
Clock	Preserve raw timestamp unit and event meaning; convert to UTC after validation	Millisecond/microsecond/nanoseconds errors and false 1970 dates
Position basis	Store multiplier, settlement price, OI convention, and turnover convention	Wrong notional, one/two-sided OI, and volume aggregation
Status	Filter frozen, pre-market, delisted, reduce-only, settling, and funding-disabled rows	Treating frozen or indicative values as live market rates
Execution	Store price limits, quantity increments, minimum notional, book fillability, and fee tier	Confusing a quoted differential with implementable net carry

For hourly comparisons, rates should be converted only after these checks. For event studies, raw event timestamps should remain alongside normalized UTC fields. For cash-flow calculations, venue-specific inverse, quanto, pool, and per-side formulas should be used where Equation (5) does not apply. These requirements are deliberately conservative: an unavailable normalization field produces a comparability warning, not an assumed default.

Appendix C. Reproduction notes

The included scripts have three separable stages. `refresh_public_snapshots.py` collects public instrument and rule responses and writes a UTC request manifest. `refresh_settlement_histories.py` collects the four BTC history samples. `build_data_dictionary.py` writes the curated venue and field dictionaries. `analyze_snapshots.py` reads the saved responses and regenerates the descriptive summaries and Figures 1–4. Network access is needed only for the first two stages; analysis of the included responses is offline.

The historical files used for the reported paper results are immutable within the package. Running a refresh with a new output directory creates a new dated sample rather than overwriting them. Researchers should compare any new interval distribution or timestamp summary with the reported dates before treating changes as contradictions. Network errors, HTTP errors, and JSON-decoding errors remain in the manifest. A successful HTTP status without parseable JSON is not counted as a successful response.

Figure 2 uses all returned rows; Figure 4 applies Equation (7) to raw timestamps; Figure 3 uses hypothetical $10,000 notional. No figure estimates returns.