HUMARB RESEARCH 10

Methodological preprint · Author handoff · Not peer reviewed

Event-Time Accounting for Two-Leg Perpetual Positions:

Quantity Closure, Asynchronous Settlement and Evidence Levels

Hang Liu · humarb Research, London, United Kingdom · https://humarb.ai · ORCID 0009-0008-1059-3926

Corresponding author: humarb Research · https://humarb.ai · admin@humarb.ai

DOI: 10.5281/zenodo.23932842

Document status. Author handoff dated 4 November 2026. The DOI and Zenodo record are given on the title page.

Research channel. https://humarb.ai/research/. This paper was completed by humarb Research; humarb.ai is its research channel.

License. CC BY 4.0. This document is a methods note with a bounded public-book illustration, not investment advice or a live-trading record.

Abstract

Two-leg perpetual positions are often evaluated by matching quote-currency notionals and extrapolating a displayed funding-rate spread. That practice can leave residual quantity, omit eligible funding events, double-count execution costs, and turn unavailable evidence into zero. This methods note develops an event-time ledger for linear, same-settlement-currency perpetual pairs. It combines common-base-quantity closure, four depth-walked fills, per-event funding eligibility, a closed-position identity, a basis-minus-crossing decomposition, scenario-consistent error attribution, and explicit evidence labels. Deterministic examples show that symmetric enumeration produces 7.20 USDT rather than 0.20 USDT in an eight-hour toy window and that a direction-fixed "oracle" need not be an upper bound. A bounded illustration using three routes and two mutually exclusive arms on one day yields net-before-funding paper results from −24.93 to −89.91 USDT. These are not independent trades or portfolio returns. Official settlement marks, account fee evidence, and real orders are absent; exact account net remains unavailable. The framework separates accounting validity from statistical or economic validation.

Keywords: perpetual futures; funding settlement; event-time accounting; quantity closure; execution costs; evidence levels

JEL classification: G12; G13; G14

Series and access: humarb Research 10 · https://humarb.ai/research/ · DOI 10.5281/zenodo.23932842.

Contents

Abstract **1**

1 Introduction **5**

1.1 Research question and gap 5
1.2 Approach and contributions 5
1.3 Identification, evidence, and what the numbers mean 6
1.4 Robustness, falsification, and scope 7
1.5 Relation to the humarb Research series and roadmap 7

2 Related literature **8**

2.1 Perpetual futures and funding design 8
2.2 Cross-market basis and limits to arbitrage 9
2.3 Implementation shortfall and transaction costs 9
2.4 Forecast evaluation, selection, and research degrees of freedom . . . 10
2.5 Position within the humarb Research sequence 10

3 Institutional background **11**

3.1 Premium, interest, damping, and caps 11
3.2 Eligibility is a snapshot rule, not merely a timestamp 12
3.3 Contract units, fee bases, and public APIs 12
3.4 Implications for a defensible event ledger 13

4 Data and evidence inventory **13**

4.1 Source, date, and unit of observation 13
4.2 Pre-specified quantity and book rules 14
4.3 Fee, funding, and other-cost assumptions 14
4.4 Protocol freezing and the pre-registration limitation 15
4.5 Missing, anomalous, and unavailable records 15

5 Event-time accounting methods **16**

5.1 Scope, signs, and notation 16
5.2 Quantity closure and four fills 16
5.3 Eligible funding events and symmetric enumeration 18
5.4 Event attribution and record completeness 18
5.5 Closed-position identity and the $B - X$ decomposition 19
5.6 Mixed quote and settlement currencies 20

5.7 Current-state scenario and error attribution 21
5.8 Scenario keys and final-objective consistency 22
5.9 Evidence labels and NA semantics 22
5.10 Joint cash risk and covariance 23
5.11 The direction-free oracle requirement 24

6 Results of the bounded paper illustration 24
6.1 Four-fill results in the 150-second arm 24
6.2 Settlement-arm scenarios and funding evidence 25
6.3 Current-state error attribution 26
6.4 What the illustration establishes 28

7 Robustness and audit checks 28
7.1 Holding-arm and asset-direction strata 28
7.2 Alternative crossing estimator 29
7.3 Fee scaling and evidence-state sensitivity 29
7.4 Independent arithmetic and implementation tests 29
7.5 Requirements for a future statistical robustness study 30

8 Limitations 30
8.1 Scope limitations of the formal ledger 30
8.2 Limitations of the public-book illustration 31
8.3 Pre-specification and reproduction limitations 32
8.4 Evaluation and external-validity limitations 32

9 Conclusion 33

Declarations 33

A Algebraic derivations 38
A.1 Basis-minus-crossing identity 38
A.2 Error-budget identity 38
A.3 Joint cash variance 38

B Reference audit algorithms 38
B.1 Quantity-closed book walk 39
B.2 Event reconciliation 39
B.3 Scenario-key validation 39

C Evidence-state dictionary **40**

D Reproduction data dictionary **40**

E Attempted configurations and reporting inventory **41**

F External pre-registration checklist for a confirmation run **41**

1 Introduction

A two-leg perpetual position looks simple only after its quantities, clocks, currencies, and evidence have been aligned. A long contract on one venue and a short contract on another may reference the same asset while using different contract multipliers, lot steps, funding schedules, valuation marks, fee bases, and settlement records. A displayed rate spread supplies none of the reconciliation needed to turn those contracts into one closed cash ledger. If a trader opens 10,000 USDT on each side and later spends another 10,000 USDT to close, the exit quantities need not equal the entry quantities. If one leg settles hourly and the other every eight hours, counting only the first receipt while retaining every payment is not a conservative estimate of the same window; it is a different and incorrectly defined cash flow. If an official settlement mark or an exit book is missing, replacing it with zero converts an absence of evidence into a false observation.

1.1 Research question and gap

This paper asks a deliberately narrow question: what accounting contract must a two-leg perpetual position satisfy before its funding, execution, and price components can be summed or compared without silently changing the position, the event window, or the evidence state? The question precedes prediction. A forecasting model may estimate future rates, and an execution model may optimize orders, but neither can repair a ledger that mixes quantities or scenario keys. Conversely, an internally consistent ledger does not establish that a route is profitable, that a forecast generalizes, or that a backtest is free of selection bias.

The gap is practical as well as conceptual. Perpetual-futures pricing research explains how funding links a contract to its underlying asset (He, Manela, Ross, and von Wachter, 2022; Ackerer, Hugonnier, and Jermann, 2026); execution research separates paper performance from realized implementation (Perold, 1988; Almgren and Chriss, 2001); and cryptocurrency-market research documents why cross-market price differences cannot be equated with costless arbitrage (Makarov and Schoar, 2020; De Blasis and Webb, 2022). Yet a small accounting error can invalidate all downstream economic interpretations before those richer models are even estimated. The required unit of analysis is therefore not a displayed rate or a nominal spread, but a quantity-closed, event-attributed, scenario-keyed ledger with explicit evidence states.

1.2 Approach and contributions

The paper makes four related contributions. First, it defines quantity closure on the common feasible base-asset lattice and attaches four depth-walked fills to that same quantity. Entry and exit are both volume-weighted average price (VWAP) calculations, and an incomplete book produces an incomplete fill rather than an invented execution at the last visible level. The closed-position identity then separates price P&L, funding cash, fees, and other applicable cash costs. A basis-minus-crossing decomposition shows that crossing is already embedded in the four

execution prices and must not be deducted twice.

Second, the paper defines funding at the event level. Each event carries an expected settlement time, a venue-recorded settlement time, and a local receipt time, together with the quantity held immediately before the event, the official valuation mark, and the settled rate. Eligibility is established under the venue's rule rather than inferred from a nominal clock boundary alone. Receipts and payments are enumerated over the same half-open window, and history records are matched one-to-one to expected events. "No scheduled event," "scheduled event but record missing," and "event semantics unverified" remain different states.

Third, the paper separates an unchanged-current-state scenario from a realized or paper-realized outcome. The scenario freezes only inputs known at the scenario time and uses a hypothetical four-fill round trip on current books. An error-attribution identity decomposes the subsequent gap into funding, price, fee, and other-cost differences rather than labelling the entire gap "slippage." Every component carries the same scenario key: route, contract specification, quantity, quote currency, fee basis, holding view, target time, and information version. A deterministic example further shows that selecting a holding horizon with a surrogate that omits a horizon-dependent basis term can reverse the ranking implied by the final objective.

Fourth, the paper introduces a compact evidence vocabulary with non-arithmetic missing-data semantics.

- Cash and snapshot evidence: `official_cash_only` and `snapshot_cost_proxy`.
- Paper and actual execution: `four_book_paper` and `actual_fills`.

Unavailable inputs are reported as not available (NA), never as zero and never as losses. The vocabulary is paired with two audit counterexamples. One demonstrates that truncating receipts can change both cash and preferred exit time. Another demonstrates that an ex post calculation retaining the direction chosen by a last-settled-rate rule is not a perfect-foresight upper bound; freeing direction can reverse the toy conclusion.

1.3 Identification, evidence, and what the numbers mean

The formal conclusions are identified by algebra under explicit assumptions, not by statistical correlation. The main closed-position result assumes verified linear contracts, a common base asset and settlement currency, equal base quantity on both legs, no resizing, and full closure. The basis decomposition additionally requires synchronous or clearly attributed mid-prices on a common quote basis. The event-cash equation requires an eligibility set and valuation marks consistent with venue rules. When these assumptions fail, the paper does not estimate around them: inverse and quanto contracts, continuous accrual, special index adjustments, liquidation, and unresolved mixed currencies require separate adapters or an unsupported/NA status.

The illustrative evidence is intentionally bounded. On 15 October 2026, the source experiment examined three public-book routes linking Gate and BingX:

RSR in both directions and a BTC control. Two arms were defined before entry: exit after 150 seconds and exit at the first qualifying paired books after 11:00:20 UTC, following an 11:00 UTC settlement. The per-leg budget ceiling was 10,000 USDT, but exits used the entry base quantity rather than a renewed dollar budget. All six route-arm net-before-funding paper results were negative, ranging from −24.93 to −89.91 USDT. This is a description of six correlated counterfactual accountings, not a zero-percent win rate. No sum, portfolio return, annualized return, or opportunity frequency is identified.

Funding evidence is narrower still. The 150-second windows contained no scheduled funding event under the pre-entry schedule and post-exit history, so their empty event sets produce zero funding by accounting definition. In the settlement arm, the BTC route also had a verified empty event set. The RSR proxy used a pre-entry frozen mark of 0.001919 and a settled rate of −0.000108, producing proxies of +1.08 and −1.08 USDT for the two directions. The frozen mark was not the official settlement mark. Therefore exact RSR funding cash and exact account-level net are NA; the proxy-inclusive values are scenarios, not realized profits or losses.

1.4 Robustness, falsification, and scope

The paper evaluates robustness at the level justified by the evidence. Four specification checks vary the holding arm, asset and direction stratum, crossing estimator, and fee scaling; a fifth check varies funding evidence from a verified empty set to a labelled proxy and then to NA. Independent recomputation of the supplied summary tables verifies arithmetic within disclosed display rounding. These checks test accounting sensitivity, not market-regime stability. There is no multi-day subsample analysis, third-venue comparison, predictive out-of-sample exercise, multiple-testing adjustment, or econometric confidence interval. Those omissions are limitations rather than results to be implied through language.

The central identities are nevertheless falsifiable. A reviewer can supply a worked fill set and test whether quantities close, whether price P&L equals basis minus crossing, and whether the scenario-error residual vanishes. A reviewer can duplicate or omit a settlement record and test whether event matching rejects it. A reviewer can remove an official mark and test whether the implementation returns NA rather than zero. The accompanying clean-room scripts expose these checks without revealing production rules or private implementation details.

1.5 Relation to the humarb Research series and roadmap

humarb Research 01 asks how executable net carry should be measured from funding estimates and depth-dependent costs (Liu, 2026a). humarb Research 02 asks how reference and conservative net-carry cases should be framed under settlement, execution, and liquidation-aware assumptions (Liu, 2026b). This complementary methodological layer does not repeat their settlement-block grids, annualization discussion, or liquidation framework. It asks whether the underlying two-leg ledger is closed and attributable in the first place. The three papers are

complementary: measurement, framing, and event-time accounting address different failure modes.

The remainder follows the journal structure specified for this methods note. Section 2 positions the contribution in five literatures. Section 3 summarizes time-varying venue mechanics. Section 4 documents the public-book illustration and evidence inventory. Section 5 gives the complete accounting contract. Section 6 reports each route-arm result without aggregation. Section 7 presents specification and audit checks. Section 8 states what the paper cannot support, and Section 9 concludes without adding results. Appendices provide derivations, audit algorithms, a notation table, and the reproduction dictionary.

2 Related literature

This paper sits between contract design, execution measurement, and research audit. It does not propose a pricing kernel, an order-scheduling optimizer, or a forecasting estimator. Its closest intellectual antecedent is implementation shortfall: a result can be wrong not because the economic idea is incoherent, but because the object measured on paper differs from the object that would have been traded (Perold, 1988). Event-time accounting narrows that concern to the point at which two perpetual legs are first represented as one position.

2.1 Perpetual futures and funding design

Perpetual futures replace terminal convergence with a sequence of transfers intended to anchor the contract price to an underlying reference. He et al. (2022) derive frictionless no-arbitrage prices and bounds with trading costs, emphasizing that perpetuals do not guarantee convergence to spot. Ackerer et al. (2026) characterize linear, inverse, and quanto contracts under alternative funding specifications and show how the funding design affects pricing. These models make clear that a perpetual is not simply a futures contract with an unusually distant expiry. Contract type and cash-flow specification are primitives, not implementation details.

The present paper takes a different viewpoint. It does not derive the price that a funding rule should imply under no-arbitrage assumptions. It asks how cash already associated with a specified position should be recorded under the rule actually in force at each event. That distinction matters for linearity and currency. The pricing literature can compare the economics of inverse and quanto contracts; the accounting framework here deliberately excludes them until a suitable adapter is supplied. A pricing analogy is not a unit conversion.

Funding design also creates a measurement distinction between a rate and a cash amount. A rate is defined by the venue's premium, interest, damping, cap, and interval conventions. Cash requires the eligible quantity and the valuation mark at the settlement event, together with a sign convention. The same settled rate can therefore coexist with different cash amounts across contract specifications and quantities. This paper's per-event equation makes those inputs visible

rather than embedding them in a single annualized spread.

2.2 Cross-market basis and limits to arbitrage

Cross-market cryptocurrency research provides the economic reason not to read a price or funding discrepancy as free income. Makarov and Schoar (2020) document large but constrained arbitrage across cryptocurrency exchanges, while De Blasis and Webb (2022) connect Bitcoin futures arbitrage to contract design and market structure. The broader limits-to-arbitrage literature shows why capital, funding, and intermediation frictions can prevent apparent price discrepancies from disappearing (Shleifer and Vishny, 1997). For a two-leg perpetual position, these frictions appear separately on each venue: collateral is not automatically shared, one leg can be liquidated while the other remains open, and a transfer cannot be assumed to occur at the moment it is needed.

The accounting framework does not estimate those economic constraints. It prevents them from being hidden by an inconsistent identity. A spread-reversion term can be positive or negative; it is not interchangeable with funding cash. Margin is committed capital rather than a transaction fee; treating it as a fee would distort net P&L, while ignoring the capital constraint would distort economic interpretation. Transfer and borrow costs enter only when applicable and measured. This separation allows the later economic analysis to ask the right question, but it does not answer that question in this paper.

2.3 Implementation shortfall and transaction costs

Implementation shortfall distinguishes the performance of a paper portfolio from that of an actual portfolio (Perold, 1988). Optimal-execution models subsequently formalize the trade-off between market impact and timing risk (Almgren and Chriss, 2001). The four-book exercise here is much narrower than either research program. It walks displayed books at one fixed base quantity, producing a paper VWAP rather than a forecast of queue position, child-order execution, or market response. No optimal order schedule is inferred.

Still, the implementation-shortfall distinction disciplines the evidence labels. A current book is not a future fill. Four books are not order confirmations. An account's actual fills are stronger execution evidence, but even they do not remove the need to reconcile funding and fees. The basis-minus-crossing decomposition provides a precise bookkeeping bridge between mids and VWAPs: crossing is measured relative to a mid, while price P&L is measured from executions. Subtracting both price P&L and a second copy of crossing would overstate cost. This is an accounting correction, not a claim that crossing is the only implementation shortfall.

The fixed-quantity choice is equally important. A notional-matched position is not necessarily quantity-matched after prices move. Execution research often takes the order quantity as given and asks how to trade it; a renewed-notional exit changes that given quantity between entry and exit. The resulting residual exposure is a position-management outcome, not a slippage estimate. The common-lattice construction in this paper ensures that the execution question is

asked about the same base quantity on all four walks.

2.4 Forecast evaluation, selection, and research degrees of freedom

A large literature warns that selecting the best result from many trials can create post-selection optimism. The optimizer's curse formalizes the upward bias of optimized expected values under uncertainty (Smith and Winkler, 2006). White (2000) and Hansen (2005) develop reality-check and superior-predictive-ability approaches for data snooping, while Harvey, Liu, and Zhu (2016) argue for stronger multiple-testing discipline in financial factor research. Bailey, Borwein, López de Prado, and Zhu (2017) address backtest overfitting through combinatorially symmetric cross-validation.

Those procedures are not decorative requirements to attach to every numerical table. They require a defined estimator, loss, candidate set, evaluation sample, and dependence structure. The six route-arm rows in this paper do not constitute such a design. They were not assembled to estimate a forecasting rule, and the two RSR directions and two arms are related counterfactuals. Running a nominal significance test on them would create false precision. Accordingly, the paper reports no Diebold–Mariano test, coverage test, Newey–West standard error, reality check, SPA test, or probability of backtest overfitting.

The literature nevertheless shapes the audit protocol. Any future selection over routes, horizons, or parameters must archive the full candidate set and evaluate the final objective on genuinely held-out observations. It must not add a horizon-dependent basis term only after a surrogate has selected the horizon. It must also preserve unmeasured candidates in an unknown category rather than moving them into the loss denominator. The deterministic objective example translates these research-design lessons into a simple algebraic failure that can be detected without market data.

2.5 Position within the humarb Research sequence

The two companion preprints address adjacent but distinct layers. Liu (2026a) frame executable net carry using robust funding estimates, depth-walked costs, freshness controls, and explicit unmeasured inputs. Liu (2026b) distinguish reference and conservative accounting cases and discuss asynchronous settlement and liquidation-aware stress. Neither paper's snapshot estimate should be read as a realized return. This paper supplies a more elementary control: before net carry is framed or forecast, the position must be quantity-closed, the funding events must be attributed, and the evidence state must be explicit.

The contribution is therefore complementary rather than substitutive. A net-carry estimate built on a closed ledger can use the earlier measurement and framing methods. A net-carry estimate built on mixed scenario keys cannot be repaired by choosing a more sophisticated funding median or a more conservative liquidation buffer. Event-time accounting is the audit layer that determines whether those later calculations refer to one well-defined position at all.

3 Institutional background

Venue documentation is part of the data-generating process for a perpetual ledger. It determines when a position is eligible for funding, which price values the position, how the rate is transformed, and when the resulting record becomes visible. These rules can vary by contract and through time. The discussion below is therefore a documentation checklist rather than a claim that every perpetual shares one permanent formula. Official Gate and BingX materials were accessed on 4 November 2026; the paper experiment used the rules captured in its 15 October 2026 protocol. The companion net-carry paper discusses settlement-grid construction and annualization in detail (Liu, 2026b); they are not repeated here.

3.1 Premium, interest, damping, and caps

A common funding construction begins with a premium index comparing depth-weighted perpetual prices with an external index. In schematic form,

$$PI_t = \frac{\max(0, B_t - I_t) - \max(0, I_t - A_t)}{I_t}, \tag{1}$$

where I_t is the reference index, B_t is the designated impact or depth-weighted bid, and A_t is the corresponding ask. The venues define those designated prices and their sampling schedules. A second component represents interest, and a clamp creates a dead zone around the interest term. Gate's published formula averages premium observations with increasing weights toward settlement, scales the result for the contract's settlement interval, and applies contract-specific limits (Gate, 2026d). BingX's January 2026 notice describes $F = P + \text{Clamp}(I - P, \beta, -\beta)$, where P is its weighted average premium and β is the damping parameter (BingX, 2026c).

The economic labels "premium" and "interest" should not obscure the implementation differences. Gate documents minute-level premium and funding calculations, whereas the cited BingX notice documents five-minute premium sampling. Gate's documentation specifies an outer limit and an interval adjustment; the BingX notice describes a base interest rate that is proportionally adjusted for the settlement cycle. Older help pages can display different formulas or defaults. The correct response is not to choose whichever page gives a convenient constant. It is to version the rule by contract and effective time and to retain the official settled rate for reconciliation.

A cap or dead zone matters at the settlement scale. Averaging uncapped displayed rates and applying a cap only to the final average is not generally equivalent to applying the venue's transformation at each event. The same warning applies to a missing interval change. If a contract moves from an eight-hour schedule to hourly settlement, an eight-hour assumption omits seven potentially eligible events. For this reason, the ledger consumes the settled event history and the effective contract specification; it does not infer the schedule from a venue-level habit.

3.2 Eligibility is a snapshot rule, not merely a timestamp

Both venues describe funding as a transfer between long and short position holders rather than as an exchange trading fee. Positive rates ordinarily mean that longs pay shorts; negative rates reverse the direction. Gate states that eligible position value for a USDT-margined linear contract is mark price times position quantity times the contract multiplier (Gate, 2026d). BingX similarly describes funding as nominal position value times the settlement rate, with nominal value based on position amount and mark price (BingX, 2026c).

Eligibility requires more than seeing a whole hour inside a holding window. Gate describes an asynchronous process in which a snapshot is taken at the scheduled boundary and settlement can complete later. BingX's fee documentation warns that calculation and settlement can be delayed and that openings or closures near the boundary can receive different treatment depending on the precise processing sequence (BingX, 2026b). A backtest that assigns eligibility solely from a rounded local timestamp can therefore be wrong even if the expected event time is correctly identified.

This paper responds with three distinct clocks. The expected event clock answers when settlement was scheduled. The venue-recorded clock answers when the venue says the cash or settlement record was created. The local receipt clock answers when the research system observed the record. All three can be correct while differing. A look-ahead-free replay must use information that would actually have been available at the simulated decision time; an event-time ledger must attribute the cash to the correct scheduled event. Those are related but not identical requirements.

3.3 Contract units, fee bases, and public APIs

Contract identity includes the base asset, quote and settlement currencies, linear or inverse type, contract multiplier, quantity step, minimum size, and price precision. Gate exposes public contract fields such as multiplier, funding interval, and per-contract taker fee rate through its futures API (Gate, 2026a,b). BingX's official public market-data reference lists contract, depth, premium-index, and funding-rate endpoints and documents base-denominated quantities in its current public interface (BingX, 2026a). These public responses are suitable for reproducible market-data collection; private account endpoints are neither needed nor used in the paper illustration.

Fee evidence has its own hierarchy. A public per-contract taker rate is a transparent scenario assumption, but it is not proof of the rate that a particular account would pay. Gate's fee documentation describes account tiers and point-based deductions (Gate, 2026c). BingX publishes a standard taker schedule and also describes VIP discounts (BingX, 2026b). The source experiment therefore applies its recorded public per-contract rates—0.00075 for Gate and 0.0005 for BingX—to all four paper fill notionals, while refusing to label the result an actual account fee. A rebate, if evidenced, would enter with a negative fee rate rather than being silently discarded.

The distinction between a budget and a quantity is especially important for Gate-style contracts. If one contract represents m_i base units, then n_i contracts produce $Q_i = n_i m_i$. Rounding n_i without converting through m_i can make an apparently small unit discrepancy dominate the funding calculation. BingX's coin-denominated book quantities remove that particular multiplier conversion but do not remove lot-step or minimum-size checks. Both legs must be placed on a common base-asset lattice before entry.

3.4 Implications for a defensible event ledger

Four documentation rules follow. First, store the effective contract specification alongside the raw book, not in a separate unversioned lookup. Second, store scheduled and recorded settlement times separately and treat ambiguous records as unresolved rather than merging them. Third, store the official valuation mark whenever exact funding cash is claimed; if only a proxy is available, retain a proxy label. Fourth, record the public fee basis and the absence of account-specific evidence. These requirements are deliberately mundane. They prevent a sophisticated calculation from resting on an undocumented unit or clock.

The methods developed below assume that these mechanics have been verified for linear, stablecoin-settled pairs. They do not silently extend to inverse, quanto, continuously accruing, or specially adjusted contracts. Unsupported mechanics produce NA/UNSUPPORTED, which is a valid research output. An unavailable answer is preferable to a numerically precise answer for the wrong instrument.

4 Data and evidence inventory

4.1 Source, date, and unit of observation

The numerical illustration comes from a desensitized research-team note documenting a public-book paper experiment on 15 October 2026. Two perpetual venues, Gate and BingX, were queried through public market-data endpoints. No orders were submitted, no private account was read, and no account funding or fee record was available. The writer task package contains the experiment's rounded summary tables and protocol description, but not its archived raw responses, raw-response hashes, or independent re-verification script. This distinction defines what the present reproduction package can and cannot verify.

The unit of observation is a route-arm accounting, not a trade. A route specifies an asset and the venue and direction of each leg. An arm specifies a holding endpoint. There are three routes and two arms, yielding six rows. Route A is long Gate and short BingX RSR; route B reverses those RSR directions; route C is long Gate and short BingX BTC. The two RSR directions are mutually exclusive counterfactuals, and the two arms reuse the same entry state. The rows are consequently related by construction and cannot be treated as six independent trades.

The experiment's universe scan looked for a common USDT perpetual asset whose settlement schedule covered the 11:00 UTC event. It selected one lower-

priced asset, RSR, in both directions, and one liquid control, BTC. This is a targeted methodological sample, not a random or comprehensive sample of perpetual contracts. Delistings, suspensions, and assets outside the observable common-asset scan are outside its scope. Survivorship and coverage biases cannot be assessed from the supplied summary.

Table 1: Descriptive sample inventory and funding evidence by route

Route	Asset	Long / short venues	Q (base)	Arms	Funding evidence in settlement arm
A	RSR	Gate / BingX	5,212,100	150 s; 11:00Z	Frozen-mark proxy: +1.08 USDT
B	RSR	BingX / Gate	5,212,000	150 s; 11:00Z	Frozen-mark proxy: −1.08 USDT
C	BTC	Gate / BingX	0.1235	150 s; 11:00Z	Verified empty event set: 0 USDT

Notes: All rows use a 10,000 USDT per-leg entry budget ceiling and common-quantity exit. The 150-second arm has a verified empty funding event set for all routes. Route identifiers are used only to keep the three routes distinct; no cross-route aggregation is performed. Exact account net is NA for every row because no real orders or account evidence exist.

4.2 Pre-specified quantity and book rules

The protocol fixed the quantity before entry. Each leg had an entry budget ceiling of 10,000 USDT. Both entry books had to be fully measurable at the chosen quantity; the common Q was the smaller budget-walk quantity rounded down to the common lot step. Gate contract counts were converted using the published multiplier, while BingX depth used coin-denominated quantities. The exit then bought or sold that same Q. If the feasible common grid had contained no positive quantity, the appropriate output would have been "grid infeasible," not a fabricated minimum fill.

Entry required the first qualifying paired books with source age no greater than 10 seconds and cross-venue source-clock difference no greater than 5 seconds. These are experimental parameters, not production thresholds. They address input freshness only; they do not guarantee simultaneity, queue priority, or execution. Each of the four walks produces a VWAP and a fee notional. If displayed depth had been insufficient, the result would have been an incomplete fill rather than a VWAP padded with a synthetic last-level quantity.

The two exit arms provide a short-horizon and a settlement-spanning comparison. The first targets entry plus 150 seconds. The second uses the first qualifying paired books after 11:00:20 UTC, with both exit source times at or after the target. The exit books of the two RSR directions were captured about 1.1 seconds apart. Their basis changes therefore need not mirror one another, even though the routes share an asset and an entry design.

4.3 Fee, funding, and other-cost assumptions

The protocol applies the recorded public per-contract taker rates of 0.00075 on Gate and 0.0005 on BingX to each of the four actual paper-fill notionals. Four-

fill fee totals are approximately 25 USDT per row. No VIP tier, points deduction, rebate, or other account-specific fee evidence is assumed. The fee calculation is thus a public-rate scenario, not an account fee reconstruction.

Funding is kept separate from fills. The public funding history available to the source experiment contained rate and time but not an official settlement valuation mark. In the RSR settlement arm, the proxy uses the entry quantity, a pre-entry frozen mark of 0.001919 captured at about 10:39 UTC, and a settled rate of -0.000108. The sign depends on route direction. The proxy is $+1.08$ USDT for route A and -1.08 USDT for route B. Exact funding cash remains unavailable because the frozen mark is not the official settlement mark. The BTC settlement arm and all 150-second rows have verified empty event sets, for which zero is an accounting result rather than a missing-data substitute.

Other cost O is set to zero only as an explicit isolated-accounting assumption. Queueing, order-book reaction, partial fills, financing, transfers, liquidation, borrow costs, and capital constraints are not simulated. Setting $O = 0$ therefore means that these channels are outside the toy ledger's measured boundary, not that they are economically irrelevant or known to be zero in live trading. Margin is not subtracted as a fee.

4.4 Protocol freezing and the pre-registration limitation

The source note describes a locally frozen protocol with a SHA-256 hash before entry. It also records an important exception. The first protocol version was aborted after a venue returned HTTP 400 for an excessive depth request; no qualifying paired books and no entry or exit results were produced. Version 2 changed only the requested depth. An external public code-hosting receipt arrived after the version-2 freeze and after the short-arm result. It therefore cannot independently establish that the full protocol was publicly timestamped before all results.

This paper accordingly describes the exercise as pre-specified and locally hashed, not as independently pre-registered before every result. The local hash chain and raw request clocks are evidence of process, but the external timing limitation remains. A future confirmation run should obtain an independently timestamped protocol receipt before any route outcome is computed, preserve the complete candidate list, and execute without changing quantity, arms, or fee basis after seeing results.

4.5 Missing, anomalous, and unavailable records

The data protocol distinguishes three conditions that are often collapsed. A verified empty event set means that the schedule and history establish no eligible event in the window. A scheduled event with no matching history record means that evidence is incomplete. Two records near one event may represent distinct cash flows, a correction, or an update; without venue semantics they remain ambiguous and are not merged. A latest history timestamp beyond the exit does not prove that every earlier event is present.

The same semantics apply to price and fee inputs. An absent exit book precludes the `four_book_paper` label. An absent official mark precludes exact funding cash. An absent account fee tier precludes an actual-fee claim. NA is not zero and does not enter a loss bucket. With these rules, Table 1 is both a sample description and an evidence inventory: it shows not only what was measured, but also which stronger measurements remain unavailable.

5 Event-time accounting methods

5.1 Scope, signs, and notation

The baseline ledger covers two verified linear perpetual contracts on the same base asset, settled in the same explicitly named currency. Leg $i \in \{L, S\}$ denotes the long and short legs, with direction $d_L = +1$ and $d_S = -1$. One contract on leg i represents m_i base units, so n_i contracts produce $Q_i = n_i m_i$. Prices are quoted in the common quote currency per base unit. Funding rates and fee rates are fractions rather than percentages. All cash amounts are signed from the combined position's perspective: receipts are positive and payments are negative.

The baseline assumes full closure and no resizing. Inverse, quanto, continuously accruing, and specially adjusted contracts require their own algebra. Liquidation interrupts the full-closure assumption and therefore lies outside the baseline identity. Mixed USDT/USDC settlement requires event-time conversion evidence or an explicitly labelled stablecoin-parity scenario. Unsupported mechanics produce NA/UNSUPPORTED, never a zero-filled ledger.

Table 2: Notation used in the event-time ledger

Symbol	Definition	Unit
i, d_i	Leg index and direction: $d_L = +1$, $d_S = -1$	dimensionless
n_i, m_i	Contract count and base units per contract on leg i	contracts; base/contract
Q, Q_i	Common closed base quantity and leg-specific base quantity	base
N	User-set per-leg entry notional ceiling; not a closure invariant	quote
$p_{i,k,j}, q_{i,k,j}$	Price and base quantity at depth level j of action k on leg i	quote/base; base
P_i^e, P_i^x	Entry and exit VWAPs on leg i	quote/base
m_i^e, m_i^x	Attributed mid-prices at entry and exit on leg i	quote/base
$f_{i,k}$	Fee rate for action $k \in \{e, x\}$; a rebate is negative	fraction
Φ	Sum of all four execution fees	quote
τ	A scheduled funding event, with quantity evaluated at τ^-	UTC time
E_i	Eligible funding-event set for leg i under the effective venue rule	event set
$Q_i(\tau^-)$	Actual quantity held immediately before event τ	base
$M_{i,\tau}$	Official funding valuation mark for leg i at event τ	quote/base
$r_{i,\tau}$	Settled funding rate for leg i at event τ	fraction
C	Signed funding cash over the declared window	quote
O	Other applicable measured cash costs not already in prices or fees	quote
Π	Closed-position net P&L	quote
B, X	Mid-price basis change and round-trip crossing cost	quote
$S_0(T)$	Unchanged-current-state scenario for target T	quote

5.2 Quantity closure and four fills

A notional ceiling determines how large a position may be opened; it does not determine how much quantity must be closed. The common feasible lattice com-

bines both venues' lot steps and minimum sizes. A candidate Q is feasible only if it can be represented on both legs, both entry notionals are at most N, and both entry books can fill it. The experiment chooses the smaller budget-walk quantity rounded down to that common grid. The short entry sells Q, and its exit buys back the same Q; spending N again at exit would buy a price-dependent quantity and leave residual exposure.

For each leg i and action $k \in \{e, x\}$, a complete depth walk satisfies

$$\sum_j q_{i,k,j} = Q, \qquad P_{i,k} = \frac{\sum_j p_{i,k,j} q_{i,k,j}}{Q}. \tag{2}$$

The VWAP is undefined if the book cannot supply Q. An implementation must return the unfilled remainder and an incomplete-fill state rather than assuming that the remainder trades at the last visible level. The two entry and two exit VWAPs are therefore measurements at fixed quantity, not top-of-book approximations and not predictions of order-book replenishment.

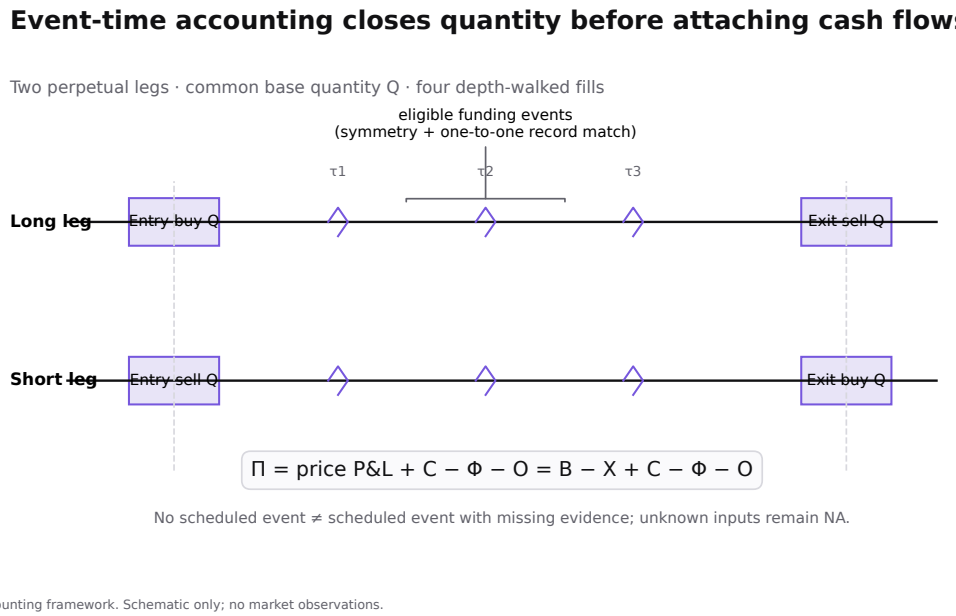


Figure 1: Event-time accounting for a quantity-closed two-leg position. Four fills use one common base quantity, while eligible funding events are matched one-to-one to settlement records. Source: authors' accounting framework; schematic only.

With action-specific fee rates, total trading fees are

$$\Phi = \sum_{i \in \{L,S\}} \sum_{k \in \{e,x\}} \left(\sum_j p_{i,k,j} q_{i,k,j} f_{i,k} \right) = \sum_{i \in \{L,S\}} \sum_{k \in \{e,x\}} Q P_{i,k} f_{i,k}. \tag{3}$$

The second equality follows from Equation (2) when the same fee rate applies to every level of an action. A rebate enters as $f_{i,k} < 0$. The shortcut $\Phi = 2N(f_L + f_S)$ holds only if all four notionals equal N, an assumption that is generally false after depth walking and price movement. Unknown required fee inputs remain

NA rather than defaulting to a venue headline rate.

5.3 Eligible funding events and symmetric enumeration

For each leg, E_i is the set of events for which the actual position is eligible under the effective snapshot/payment rule. The signed funding cash is

$$C = -\sum_{i\in\{L,S\}}\sum_{\tau\in E_i} d_i Q_i(\tau^-) M_{i,\tau} r_{i,\tau}. \tag{4}$$

With a positive rate, a long pays and a short receives; with a negative rate, the signs reverse. Partial fills or resizing enter through $Q_i(\tau^-)$, not through a constant notional assumption. Account-level funding cash, when available, is authoritative; Equation (4) then serves as a reconciliation. If the official valuation mark is unavailable, a disclosed proxy can be used only with a proxy label, and exact funding cash remains unavailable.

The eligible set is constructed over one half-open window for both legs. For a discrete leg whose first eligible settlement after `start` is `first` and whose interval is `step`, the scheduled count inside $(\texttt{start}, t]$ is

$$n(t) = \begin{cases} 0, & \texttt{first} > t, \\ 1 + \lfloor (\, t - \texttt{first}\,)/\texttt{step} \rfloor, & \texttt{first} \le t. \end{cases} \tag{5}$$

Spot legs contribute no perpetual funding event. Continuously accruing instruments require a different adapter and no settlement count. Equation (5) enumerates scheduled opportunities; eligibility still depends on the venue's snapshot and processing rules.

Symmetric enumeration means that every eligible receipt and every eligible payment inside that same window enters once. Suppose, purely as a hand-checkable example, that the fast leg receives 1.00 USDT hourly and the slow leg pays 0.80 USDT once in an eight-hour window. Full enumeration gives $8(1.00) - 0.80 = 7.20$ USDT. Capping receipts at the first event while retaining the payment gives $1.00 - 0.80 = 0.20$ USDT, only 2.8% of the full figure. The latter is not a conservative estimate of the former; it is a misspecified cash flow.

The truncation can also alter exit selection. In a second deterministic example, receipts of 3.77 arrive before hours 2 and 6, while a payment of 3.53 occurs at hour 6. At hour 2 both ledgers report 3.77. At hour 6 the truncated score is $3.77 + 3.77 - 3.77 - 3.53 = 0.24$, whereas the full ledger is $3.77 + 3.77 - 3.53 = 4.01$. The truncated rule prefers the earlier endpoint; the full ledger prefers the later one. Figure 2 visualizes the first example.

5.4 Event attribution and record completeness

Three timestamps must coexist without being conflated: the expected event time τ^{sched}, the venue's recorded settlement time τ^{venue}, and the local receipt time τ^{local}. The first attributes cash to the holding window. The second documents when the

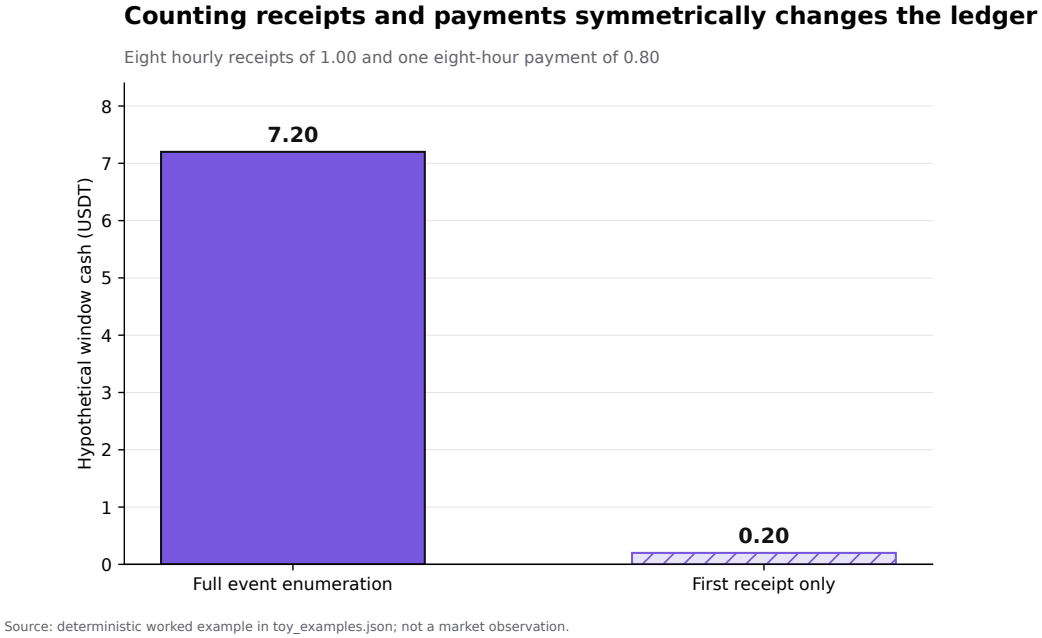


Figure 2: Symmetric versus truncated event enumeration in a deterministic eight-hour example. Full enumeration includes all eight hourly receipts and the single eight-hour payment. Source: deterministic worked example; no market observations.

venue recorded settlement. The third determines when a replaying system could have known the record. Rounding either recorded time to the minute can move information across a nominal boundary and invalidate a claim that a replay uses only previously available information.

Expected events and history records are matched one-to-one with a documented venue-specific tolerance. One record cannot satisfy two expected events. Conversely, two same-minute records are not automatically duplicates: they may be separate cash flows, a correction, or an update. Their semantics must be established before aggregation. "No scheduled event" and "scheduled event but record missing" receive different states. A maximum history timestamp beyond the planned exit proves only that the latest observed record is late enough; it does not prove completeness in the middle.

A defensible coverage check therefore starts from the expected schedule and reconciles forward to records, rather than starting from the largest and smallest records and assuming that the interior is complete. Ambiguous coverage is labelled `coverage unverified`. If exact account cash is required, unresolved ambiguity blocks an exact total rather than being absorbed into an apparent zero.

5.5 Closed-position identity and the $B - X$ decomposition

Under one settlement currency, common base quantity Q, no resizing, and full closure, net P&L is

$$\Pi = Q(P_L^x - P_L^e) + Q(P_S^e - P_S^x) + C - \Phi - O. \tag{6}$$

The first two terms are four-fill price P&L. O contains only applicable cash costs not already represented in execution prices or fees, such as measured borrow interest or a required transfer. Unknown required costs are not set to zero. Margin is committed capital, not a fee, and does not appear as a subtraction in Equation (6).

Let m_i^e and m_i^x be attributed mid-prices on the same quote basis. Round-trip crossing is

$$X = Q\left[(P_L^e - m_L^e) + (m_L^x - P_L^x) + (m_S^e - P_S^e) + (P_S^x - m_S^x)\right], \tag{7}$$

and the mid-price basis term is

$$B = Q\left[(m_S^e - m_L^e) - (m_S^x - m_L^x)\right]. \tag{8}$$

Direct substitution gives

$$\Pi_{\text{price}} = Q(P_L^x - P_L^e) + Q(P_S^e - P_S^x) = B - X, \qquad \Pi = B - X + C - \Phi - O. \tag{9}$$

Equation (9) is an alternative decomposition of Equation (6), not a justification for deducting X again. The four VWAPs already contain spread crossing and displayed-depth impact. Any additional market-response or opportunity-cost model must be separately defined; it cannot be smuggled into the identity as a second copy of crossing.

A numerical check uses $Q = 10$, entry and exit long VWAPs of 100.10 and 100.40, and entry and exit short VWAPs of 100.80 and 100.55. The corresponding mids are 100.00 and 100.50 for the long and 100.90 and 100.45 for the short. Price P&L is 5.50, $B = 9.50$, and $X = 4.00$, so $B - X = 5.50$. This identity is exact with decimal inputs and is included in the reproduction tests.

Without exit books, setting $B = 0$ and equating exit crossing with entry crossing are scenario assumptions, not observations. They may be useful for a labelled current-state calculation, but they cannot upgrade a snapshot proxy to `four_book_paper`. Similarly, a positive B is not funding income and a negative B is not necessarily slippage; the terms have different reference prices and economic meanings.

5.6 Mixed quote and settlement currencies

Currency conversion applies to cash flows when those flows are economically realized, not to every nominal trade amount as if it were a spot transaction. Opening futures notional is not a spot cash outflow. Converting four nominal trade amounts at four different stablecoin/USD rates would invent spot cash flows and could dominate an apparent funding result.

Let $Y_{i,\tau}$ denote a conversion rate from the contract's settlement currency to the reporting currency at the relevant cash-settlement time. In a mixed-currency extension, realised price P&L is converted when cash-settled, funding at each fund-

ing event, and fees at each fee event. Schematically,

$$\begin{aligned}\Pi^R = &\sum_i Y_{i,x}\Pi^{c_i}_{i,\text{price}} + \sum_i \sum_{\tau\in E_i} Y_{i,\tau}C^{c_i}_{i,\tau} \\ &- \sum_i \sum_{k\in\{e,x\}} Y_{i,k}\Phi^{c_i}_{i,k} - O^R,\end{aligned} \tag{10}$$

where c_i is the leg's settlement currency and R is the reporting currency. The precise timing convention must match the venue's cash-realization rules. If conversion evidence is missing, currency-level ledgers remain valid but aggregate reporting-currency P&L is NA. A stablecoin-parity scenario may set the relevant conversion rates to one only when explicitly labelled; it is not realised USD profit.

5.7 Current-state scenario and error attribution

An unchanged-current-state scenario answers a restricted counterfactual: what would the ledger look like if only the schedule, rates, marks, and books known at t_0 were used? For fixed Q, define

$$C_0(T) = -\sum_i \sum_{\tau\in E_i(t_0,T)} d_i Q M_i(t_0) r_i(t_0) \tag{11}$$

and

$$S_0(T) = C_0(T) + \Pi_{\text{price},0}(T) - \Phi_0 - O_0. \tag{12}$$

"Same books" means a hypothetical four-fill round trip on the current snapshot at the same Q; it is not an observation of future books. Fees are charged once for the complete holding scenario, not once per funding event. Different T values represent different strategies and must not share cash or exit-basis terms computed under another horizon.

With complete outcome evidence and matched scenario keys,

$$\begin{aligned}\Pi - S_0 = &(C - C_0) + (\Pi_{\text{price}} - \Pi_{\text{price},0}) \\ &- (\Phi - \Phi_0) - (O - O_0).\end{aligned} \tag{13}$$

The cash difference can be further separated into settled-rate changes, valuation-mark changes, and added or cancelled events. The price difference separates into exit-basis and book changes; the fee difference separates into notional, fee-tier, and fee-currency effects. This attribution identifies which assumption changed instead of describing every gap as slippage.

For the ex-funding experiment comparison, write $\Delta B = B - B_0$, $\Delta X = X - X_0$, and $\Delta\Phi = \Phi - \Phi_0$, with $O = O_0 = 0$. The signed execution overestimate is

$$D_{\text{exec}} = S_0^{\text{ex}} - \Pi^{\text{ex}} = -\Delta B + \Delta X + \Delta\Phi. \tag{14}$$

A positive D_{exec} means that the scenario was too favorable; a negative value means that the realized paper result was better than the scenario. The sign is preserved throughout the tables and figures.

5.8 Scenario keys and final-objective consistency

Every summed or ranked amount carries a key

$$\kappa = (\text{route, contract specification, } Q, \text{ quote, fee basis,}$$
$$\text{holding view, } T, \text{ information version}).$$

Funding, fees, crossing, and exit basis must share κ. Combining funding at one horizon with exit basis at another is a type error, not a conservative approximation. Implementations should reject the sum when keys differ. The same rule applies to a change in contract multiplier, quote currency, fee basis, or information version.

Objective consistency is equally strict. If a holding horizon is selected by $\arg\max_T J(T)$, then J must be the final objective evaluated at every candidate T. A deterministic counterexample makes the failure visible. Candidate A has cash 8, basis −3.07, crossing 2, fees 2, and other cost 0. Candidate B has cash 12, basis −12.64, and the same costs. The surrogate that omits basis scores A at 4 and B at 8, selecting B. The final net scores A at 0.93 and B at −4.64, selecting A. Figure 3 shows the reversal. This example demonstrates objective inconsistency only; it is not an observed trade.

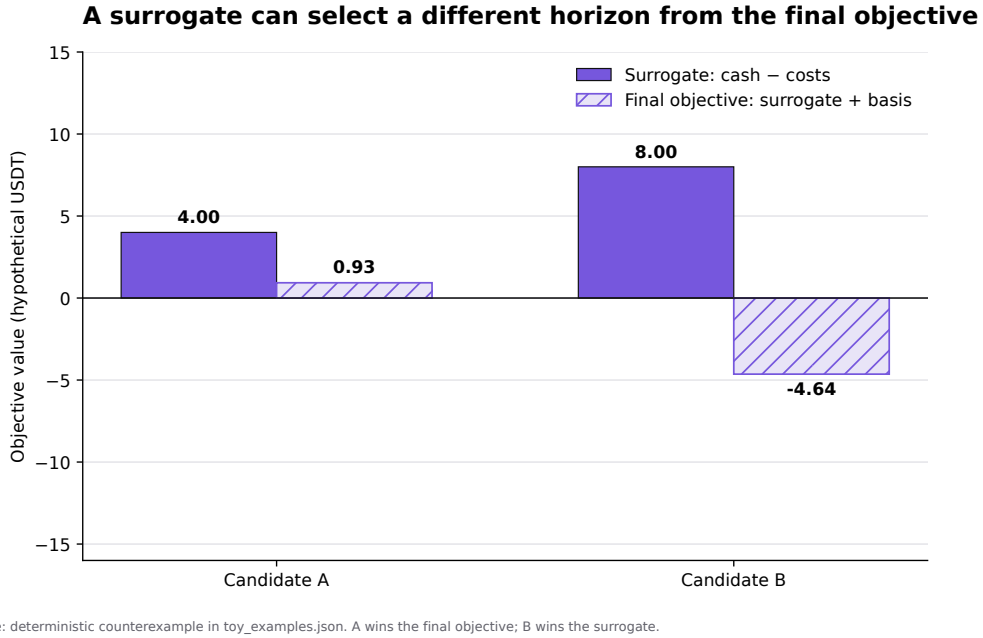


Figure 3: Surrogate and final objectives select different candidates in a deterministic example. The final objective includes the horizon-dependent basis term before selection. Source: deterministic worked example; no market observations.

5.9 Evidence labels and NA semantics

Evidence labels describe how a number was obtained and which stronger claim remains unsupported. `official_cash_only` means that account cash is authoritative for funding but does not by itself establish execution evidence. `snapshot_cost_proxy` means that a current snapshot and disclosed proxies

are used; missing exit books or official marks remain explicit. `four_book_paper` requires entry and exit books at fixed quantity, with no claim of actual orders. `actual_fills` requires real order/fill evidence, while fees and funding still require reconciliation. The labels are not a universal ordinal ranking: official cash can be stronger for one component while four public books are stronger for another.

NA has no arithmetic value. It is not zero, a loss, or an observation that can be silently removed from a denominator. Candidates with unmeasured outcomes belong to an unknown state. A measured zero requires a positive reason, such as a verified empty eligible-event set. A missing settlement mark, an unavailable exit book, and an absent account fee record are all NA, but they imply different restrictions on downstream claims. Figure 4 summarizes the vocabulary.

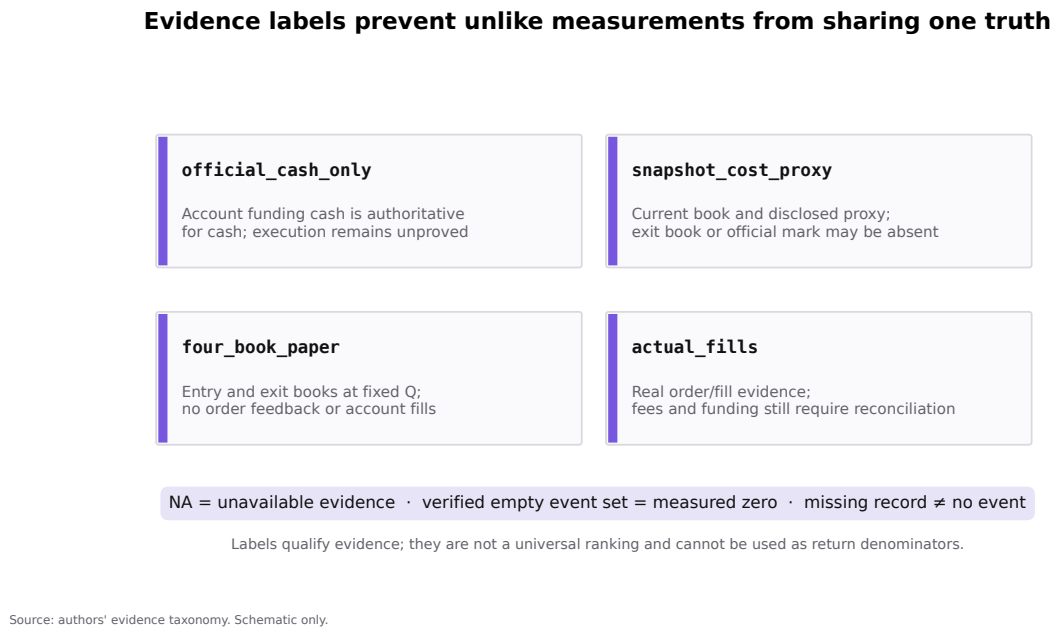


Figure 4: Evidence labels and missing-data semantics. Labels qualify components rather than supplying a universal ranking. Source: authors' evidence taxonomy.

5.10 Joint cash risk and covariance

A signed ledger is not a risk distribution. If dollar components form a vector Z and a contains their signs, then

$$\mathbb{E}[a^\top Z] = a^\top \mathbb{E}[Z], \qquad \mathrm{Var}(a^\top Z) = a^\top \mathrm{Cov}(Z) a. \tag{15}$$

For outgoing long-leg cash L and incoming short-leg cash S, the net is $S - L$, with variance $\mathrm{Var}(S) + \mathrm{Var}(L) - 2\,\mathrm{Cov}(S, L)$. Setting the covariance to zero is an independence assumption, not a conservative default.

With unit standard deviations and correlation 0.9, the difference has standard deviation $\sqrt{0.2} = 0.447$; independence would give $\sqrt{2} = 1.414$, approximately 3.16 times larger. Neither number supplies a probability of profit without an estimated distribution and out-of-sample calibration. A "mean minus z standard deviations" display is therefore a scenario statistic unless its probabilistic calibration has been

separately established. Unknown covariance is NA, not zero covariance.

5.11 The direction-free oracle requirement

An ex post upper bound must optimize over every decision that the claimed oracle is allowed to know. Suppose a strategy chooses direction from the last settled gap. If a companion "oracle" uses realised funding but retains that fixed direction, it evaluates the strategy with perfect funding knowledge, not the best direction with perfect foresight. Zero-gap points that the strategy skips must also be considered if either direction is feasible to the oracle.

A toy example makes the logical error explicit. Let the cost proxy be 28. The fixed direction realises funding of −34, while the opposite direction realises +34. The fixed-direction calculation reports −62 after cost and "cannot cover cost." A direction-free bound selects +34 and reports $34 - 28 = +6$. The latter does not establish a tradable profit: reverse-direction books, basis, units, quantity, and costs must all be recomputed. It establishes only that the fixed-direction column is not an upper bound. The appropriate label is "ex post funding coverage of the fixed direction."

6 Results of the bounded paper illustration

The results answer accounting questions rather than market-performance questions. How large are the four-fill results under the two pre-specified arms? How does a labelled funding proxy change the settlement-arm scenarios? And which components explain the gap between an unchanged-current-state baseline and the subsequent paper outcome? Each table addresses one of these questions. No row is used as an observation in a predictive regression or a portfolio return calculation.

6.1 Four-fill results in the 150-second arm

Table 3 reports price P&L, four fees, and net before funding for the short arm. All three values of net before funding are negative: −76.99 USDT for route A, −89.91 USDT for route B, and −24.93 USDT for route C. The BTC control's price P&L is +0.049 USDT before fees, illustrating why a small positive price component cannot be interpreted as net carry when four taker fees are approximately 24.98 USDT.

The table does not show that the BTC route is "better" in an economic or statistical sense. It has a different asset, quantity, price scale, and basis path. Nor do the two RSR rows provide independent replications of one strategy. They are opposite directions selected as mutually exclusive counterfactuals. The proper reading is mechanical: after fixed-quantity closure and four public-rate fee calculations, each of these three short-window paper ledgers is below zero before funding.

Table 3: 150-second arm: four-fill paper results before funding

Route	Direction	Q (base)	Price P&L (USDT)	Four fees (USDT)	Net before funding (USDT)
A	RSR long Gate / short BingX	5,212,100	−52.03	24.97	−76.99
B	RSR long BingX / short Gate	5,212,000	−64.93	24.98	−89.91
C	BTC long Gate / short BingX	0.1235	+0.049	24.98	−24.93

Notes: The 150-second windows contain no scheduled funding event according to the pre-entry schedule and post-exit history. Other cost $O = 0$ is an isolated-accounting assumption. Source values are independently rounded; for route A, displayed price P&L minus displayed fees differs from reported net by 0.01 USDT.

6.2 Settlement-arm scenarios and funding evidence

Table 4 adds the funding evidence for the settlement arm. Net before funding is −88.15 USDT for route A, −76.05 USDT for route B, and −25.40 USDT for route C. The RSR proxy increases route A's scenario value by 1.08 USDT and decreases route B's by the same magnitude, yielding proxy-inclusive scenario values of −87.07 and −77.13 USDT. Route C has no eligible funding event, so its value remains −25.40 USDT.

Table 4: 11:00Z settlement arm: paper results and funding evidence

Route	Price P&L (USDT)	Fees (USDT)	Net before funding (USDT)	Funding amount and evidence	Scenario incl. funding (USDT)	Exact account net (USDT)
A	−63.20	24.95	−88.15	+1.08, proxy	−87.07	NA
B	−51.10	24.95	−76.05	−1.08, proxy	−77.13	NA
C	−0.41	24.99	−25.40	0, empty set	−25.40	NA

Notes: RSR funding uses Q, frozen mark 0.001919, and settled rate −0.000108. The mark was captured before entry and is not the official settlement mark. Exact RSR funding is therefore NA. BTC's zero is the result of a verified empty event set. Exact account net is NA for every route because there are no real orders, account fee proof, or complete official cash records.

The NA column is essential. Replacing it with the proxy-inclusive values would imply a degree of official cash knowledge that the experiment did not possess. Replacing route C's verified zero with NA would make the opposite error: it would discard positive evidence that no event was scheduled. Evidence-sensitive accounting requires both restraints—do not invent unknown cash, and do not erase a measured empty set.

Across the two arms, the six net-before-funding values range from −24.93 to −89.91 USDT. Figure 5 displays every value without summing them. All six are negative, but the paper deliberately does not convert that description into a win rate. The numerator would combine different assets and related counterfactuals, while the denominator would pretend that the six rows form a defined sample of independent opportunities. Neither operation is justified.

There is also no annualization. A 150-second accounting and a settlement-

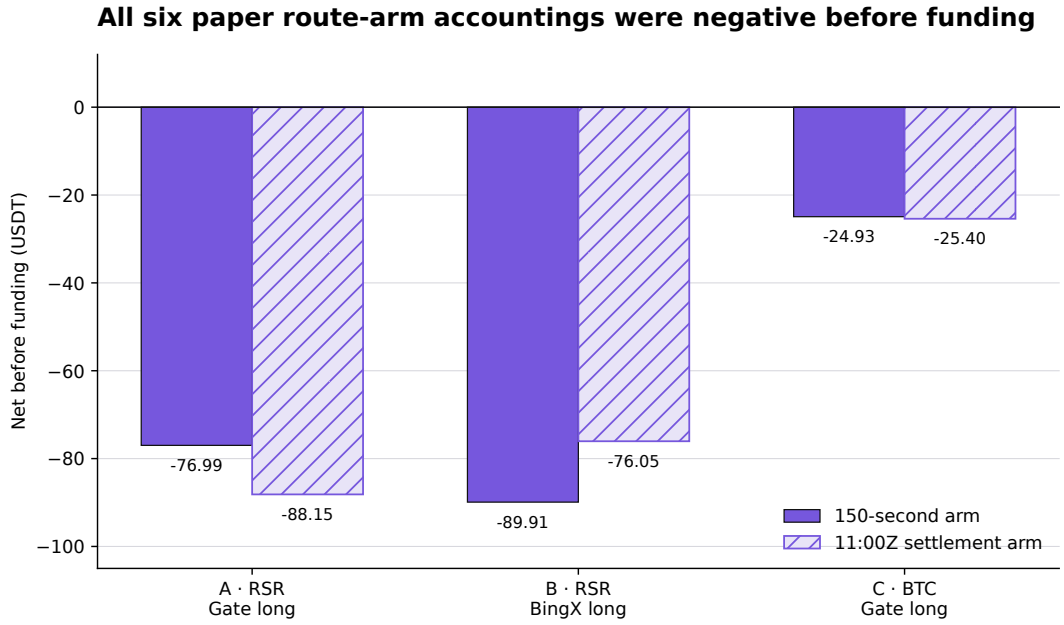


Figure 5: Net before funding for all six route-arm paper accountings. Bars show individual correlated counterfactuals, not independent trades or a portfolio. Source: supplied public-book paper-execution summary for 15 October 2026.

spanning accounting have different holding paths, event exposure, and basis risk. Multiplying either result by the number of identical intervals in a year would assume stable rates, books, fees, and eligibility that the data do not measure. The reported economic magnitude is the actual USDT amount under the experiment's 10,000 USDT per-leg budget ceiling, not a return on margin or a promised annual rate.

6.3 Current-state error attribution

The third results question concerns the source of scenario error. For route A, the same-book ex-funding baseline is −76.90 USDT and the subsequent settlement-arm paper result is −88.15 USDT, so $D_{\text{exec}} = 11.25$ USDT. Basis deterioration contributes 10.42 USDT, increased crossing contributes 0.84 USDT, and the fee delta is −0.016 USDT. The negative fee delta means that the realized paper fee was slightly lower, partially offsetting the other effects.

For route B, the baseline is −76.89 USDT and the paper result is −76.05 USDT, giving $D_{\text{exec}} = -0.84$ USDT. For route C, the baseline is −25.05 USDT and the paper result is −25.40 USDT, giving $D_{\text{exec}} = +0.35$ USDT. Route A's funding-scenario error is −0.35 USDT, while route B's is +0.35 USDT. The opposite signs are consistent with opposite funding directions but do not imply that every execution component must mirror.

The attribution shows why "slippage" is an inadequate summary. Most of route A's execution overestimate is assigned to basis deterioration, with a smaller crossing component and a tiny offsetting fee change. Calling the entire 11.25 USDT

Table 5: Settlement-arm error attribution against the same-book baseline

Route	Baseline ex-funding (USDT)	Paper ex-funding (USDT)	D_{exec} (USDT)	Funding-scenario error (USDT)
A	−76.90	−88.15	11.25	−0.35
B	−76.89	−76.05	−0.84	+0.35
C	−25.05	−25.40	0.35	NA

Notes: $D_{\text{exec}} = S_0^{\text{ex}} - \Pi^{\text{ex}}$, so a positive value denotes an overestimate. Blank or unavailable components are not zero. Route C has a verified empty funding event set; the source note does not report a separate funding-scenario error decomposition.

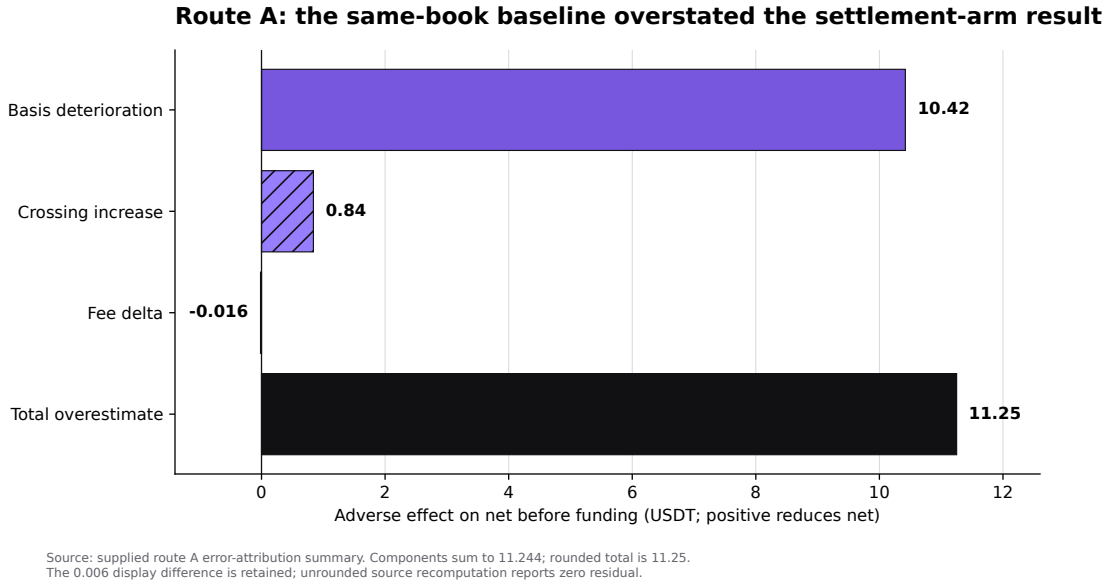


Figure 6: Route A execution-error attribution. Positive values reduce net before funding. Displayed components sum to 11.244 USDT, while the independently rounded total is 11.25 USDT; the 0.006 USDT display difference is retained. Source: supplied error-attribution summary.

"slippage" would mix a mid-price spread change with execution-price movement and fee variation. The identity instead tells a reviewer which assumption requires further investigation.

The source note reports zero identity residuals for all rows using unrounded values. The public tables are rounded independently. Accordingly, route A's detailed components sum to 11.244 USDT while the reported total is 11.25 USDT. The reproduction package preserves that 0.006 USDT display difference rather than changing a source value to force cosmetic agreement. Similarly, the 150-second route A table retains its 0.01 USDT display difference. These are transparent rounding boundaries, not hidden model residuals.

6.4 What the illustration establishes

The illustration supports three limited conclusions. First, common-quantity closure and four-fill accounting are operationally feasible on public books for the three selected routes. Second, fee and basis components can dominate small funding or price components in these particular windows. Third, the current-state error identity provides a useful decomposition of the gap between scenario and subsequent paper outcome. These conclusions concern measurement. They do not establish that negative results are typical, that positive routes do not exist, or that the method improves trading performance.

7 Robustness and audit checks

Robustness has two different meanings in this paper. Algebraic robustness asks whether an identity survives exact recomputation and whether an implementation rejects invalid states. Empirical robustness asks whether an economic estimate survives new periods, venues, and specifications. The supplied evidence supports the first and only a limited form of the second. The checks below are therefore labelled specification checks rather than market-regime validation.

7.1 Holding-arm and asset-direction strata

The first check compares the pre-specified 150-second and settlement arms. Changing the arm changes route A's net before funding from −76.99 to −88.15 USDT, route B's from −89.91 to −76.05 USDT, and route C's from −24.93 to −25.40 USDT. The directions of change are not uniform across assets or directions. This is consistent with the framework's requirement that funding, exit basis, and crossing share the same holding key; a result from one arm cannot be inserted into another without recomputation.

The second check reports RSR in both directions and BTC separately. Route A and route B do not mirror in price P&L or net before funding, and their exit books were captured about 1.1 seconds apart. The BTC control has a much smaller absolute loss than either RSR route but similar four-fee totals. These strata reveal why contract-specific quantity and book evidence matter. They do not identify an

asset effect: one day and one BTC control cannot separate asset, time, direction, and venue-liquidity explanations.

A genuine subsample-period and exchange-group robustness exercise would require additional days and venues with the same pre-specified protocol. No such observations are present. The paper does not relabel the two holding arms as independent time subsamples or the two venues as replicated exchange samples.

7.2 Alternative crossing estimator

The third check compares exact common-quantity crossing with the older fixed-dollar crossing estimate described in the source note. The fixed-dollar estimate is only 0.08–0.13 USDT higher for RSR and 0.00003 USDT higher for BTC than the exact same-quantity crossing. Thus, in these particular books, renewed-notional rounding does not explain the tens-of-dollars differences in net results.

This is a useful negative control. Quantity closure is a correctness requirement even when its numerical effect is small. The magnitude of an accounting correction is case-specific and should be measured rather than assumed to be large. A small discrepancy in one experiment does not license renewed-notional exits in general, because other prices and book shapes can produce larger residual quantities.

7.3 Fee scaling and evidence-state sensitivity

The fourth check increases every displayed four-fee total by 10% and subtracts the increment from net before funding. The downward changes range from 2.495 to 2.499 USDT across the six rows, and every row remains negative. This arithmetic stress shows that the sign descriptions are not fragile to a modest uniform fee perturbation. It does not estimate an account's actual fee distribution, the probability of a tier change, or the effect of alternative execution styles.

The fifth check varies evidence rather than a numerical parameter. Removing the frozen-mark proxy changes the RSR exact funding amount from a labelled proxy to NA; it does not change it to zero. For BTC and all 150-second rows, the verified empty event set supports exact zero within the ledger's scope. The distinction demonstrates why an evidence-state robustness check can be more informative than a spurious confidence interval. A model can be numerically stable while its claim becomes stronger than the underlying evidence permits.

7.4 Independent arithmetic and implementation tests

The accompanying clean-room package independently reads the public summary CSV files and recomputes price-minus-fee arithmetic, proxy-inclusive scenario values, baseline-minus-realized identities, the fee stress, and deterministic counterexamples. It preserves the 0.01 USDT and 0.006 USDT display differences described above. Six unit tests cover book walking and insufficient depth, funding signs and missing marks, half-open event counts, signed error budgets, NA propagation, and covariance-sensitive risk. All six pass in the delivered environment.

These tests do not reproduce the source experiment's raw-book replay. The raw

Table 6: Specification checks and their evidentiary limits

Check	Observed result	What the check does not establish
Holding-arm comparison	Each route is recomputed at 150 s and at the settlement arm; changes are not uniform	Independent time subsamples or stable horizon effects
Asset and direction strata	RSR directions and BTC remain separate; no mirrored execution terms are imposed	Estimated asset, venue, or direction effects
Alternative crossing estimator	Fixed-dollar minus same-quantity crossing: 0.08–0.13 USDT (RSR); 0.00003 USDT (BTC)	That quantity mismatch is immaterial in other books
Uniform fee stress	A 10% fee increase lowers each row by 2.495–2.499 USDT; signs remain negative	Account fee probabilities or live execution robustness
Funding evidence sensitivity	RSR proxy becomes NA without an official mark; verified empty sets remain zero	Official account cash or a statistical interval

responses and independent re-verification script were not included in the writer package, so such a claim would be false. The source note reports that a separate script re-walked archived responses and matched all amounts; that remains part of the source experiment's provenance, not an independent verification newly performed here.

7.5 Requirements for a future statistical robustness study

A future study should freeze and externally timestamp the full candidate set before observing outcomes. It should archive raw books, contract versions, all three event clocks, official marks, and account evidence where available. Evaluation should separate model-selection and final-assessment periods, retain every attempted configuration, and use dependence-aware inference by asset and time block. If predictive statements are made, rolling or expanding out-of-sample evaluation and appropriate multiple-testing controls are mandatory; the methods reviewed in Section 2 provide candidate procedures rather than results for this paper.

The strongest robustness statement supported now is narrower: the accounting identities hold in deterministic tests, the reported rounded tables can be read without hidden aggregation, and the key implementation failures—missing evidence, insufficient books, scenario mismatch, asymmetric event counting, and fixed-direction oracles—have explicit rejection rules. Statistical and economic robustness remain open empirical questions.

8 Limitations

8.1 Scope limitations of the formal ledger

The baseline identity applies to verified linear, same-base, same-settlement-currency perpetual pairs with full closure and no resizing. It does not cover

inverse or quanto contracts, continuous accrual, special index cash adjustments, forced liquidation, or unresolved cross-currency conversion. These are not cosmetic omissions. Each can change the quantity, cash sign, valuation, or closure identity. The framework returns unsupported/NA until an appropriate adapter and evidence are available.

Even within scope, the identity is only as good as its inputs. A mid-price basis decomposition requires a defensible time attribution for mids. An asynchronous mid can move basis between entry and exit without any executable implication. Funding requires the official event rate, valuation mark, eligible quantity, and eligibility rule. A public funding-rate history lacking a settlement mark cannot supply exact funding cash. An account record can resolve some of these gaps but may not expose every contractual input required for independent reconstruction.

The paper also separates accounting from solvency and risk management. It does not model margin calls, liquidation waterfalls, insurance funds, auto-deleveraging, collateral haircuts, transfer delays, counterparty failure, or venue outage. A quantity-closed position can still lose one leg through liquidation and become directionally exposed. Joint variance formulas include covariance, but they do not turn an accounting vector into a calibrated return distribution. No value-at-risk, expected-shortfall, survival probability, or probability of profit is reported.

8.2 Limitations of the public-book illustration

The experiment covers three routes, two venues, two assets, and one day. The two RSR directions and two holding arms are related counterfactuals. This design cannot identify a win rate, expected return, portfolio Sharpe ratio, opportunity frequency, or cross-sectional effect. All-negative net-before-funding rows do not imply that positive opportunities are absent elsewhere. Conversely, a future positive row would not establish a repeatable source of income.

The route scan targeted an observable common asset whose schedule covered a specific settlement. That is appropriate for testing settlement accounting but not for estimating the performance of a ranking function. The sample excludes assets without qualifying common contracts and does not provide a complete history of delisted or suspended instruments. Coverage, survivorship, and selection effects cannot be quantified from the supplied material.

Public books are not fills. The exercise does not model queue position, order-book reaction, hidden liquidity, cancellations, partial fills, or market impact after order submission. Source-age and clock-difference checks reduce stale-input risk but do not establish atomic execution across venues. The two RSR exit snapshots were about 1.1 seconds apart, so even their basis terms need not be directly comparable as simultaneous outcomes. A live implementation would require order and fill records, rejection handling, and venue-specific timestamp reconciliation.

Other costs are excluded by an explicit $O = 0$ isolated-accounting assumption. Financing, borrow, transfers, capital constraints, and liquidation are not measured. The public per-contract taker rates are not actual account fee evidence. Therefore

"net before funding" means net only within the stated toy ledger, not all-in live net. The proxy-inclusive RSR values are likewise scenarios, and exact account net is NA for every route.

8.3 Pre-specification and reproduction limitations

The protocol was frozen locally with a hash before entry, but the external public receipt arrived after the version-2 freeze and after the short-arm result. It is not independent pre-result timestamp evidence. The aborted version-1 request and the depth-limit correction are disclosed, but only a new externally timestamped run can remove the pre-registration limitation. The present paper does not use "preregistered" as an unqualified validation label.

The writer package contains rounded summary tables rather than raw archived responses. The delivered reproduction package verifies summary arithmetic and deterministic contracts only. It cannot verify source-clock thresholds, raw-response hashes, book units at each level, the original independent re-verification, or the claim that no values changed when later provenance arrived. Those checks remain part of the source experiment's documentation and should be included in a future public replication bundle.

Display rounding creates small visible differences. Route A's 150-second net differs by 0.01 USDT from displayed price P&L minus displayed fees, and its detailed error components differ from the rounded total by 0.006 USDT. The source reports zero residuals on unrounded values. The public package preserves the display differences, but exact unrounded recomputation is not possible from the rounded tables alone.

8.4 Evaluation and external-validity limitations

No forecasting or estimation claim is made, so no predictive out-of-sample test, Diebold–Mariano comparison, interval-coverage test, proper scoring rule, Newey–West standard error, panel-clustered confidence interval, or multiple-testing adjustment is reported. The fee stress and alternative crossing estimator are accounting checks, not substitutes for those methods. A future predictive study must define the estimator and candidate set, preserve all trials, use genuinely held-out data, and account for dependence by asset and time.

Finally, venue documentation changes. Funding intervals, caps, formulas, fees, and API fields can differ across contracts and effective dates. The official pages accessed on 4 November 2026 are evidence for the institutional description, not a promise that the same defaults will remain available. A replication must archive the documentation and public responses effective at its own experiment time. The methods note offers an accounting contract and audit vocabulary, not an investment recommendation, a live system specification, or a guarantee of execution or income.

9 Conclusion

A two-leg perpetual position becomes a coherent accounting object only after quantity, time, currency, and evidence are aligned. This paper has formalized that alignment. Common-base-quantity closure ensures that four fills enter and exit the same position. Event-level funding ensures that every eligible receipt and payment is counted once under a common window and venue rule. The closed-position identity and $B-X$ decomposition separate funding, basis, crossing, fees, and other measured costs without double counting. Scenario keys and error attribution keep current-state assumptions distinct from subsequent outcomes.

The framework also treats missing evidence as a substantive state. NA is not zero, a verified empty event set is not missing data, public books are not actual fills, and a proxy mark is not an official settlement mark. The deterministic counterexamples show why asymmetric event enumeration, mixed objectives, and fixed-direction "oracles" can reverse conclusions before any statistical model is estimated. The small public-book illustration demonstrates the calculations while remaining explicitly unsuitable for return, win-rate, or portfolio claims.

The appropriate next step is not stronger language but stronger evidence: an externally timestamped protocol, archived raw books and contract versions, official settlement marks and account records where available, and a genuinely held-out evaluation design. Until then, the contribution is a reproducible accounting contract and audit method. It clarifies what a two-leg ledger means, what it does not mean, and which conclusions must wait for additional data.

Declarations

Corresponding author and research channel

Correspondence: humarb Research, London, United Kingdom · https://humarb.ai · admin@humarb.ai. This paper was completed by humarb Research; humarb.ai is its research channel. The named author is Hang Liu, ORCID 0009-0008-1059-3926.

Competing interests

The author reports no competing financial or non-financial interests directly related to this study. humarb Research conducts research on digital-asset market mechanisms. This affiliation does not alter the evidence boundaries or the negative empirical findings reported here.

Funding

This research received no external funding; it was conducted by humarb Research.

Data availability

The institutional description uses official Gate and BingX documentation and public API references accessed on 4 November 2026. The route-arm illustration uses rounded values from a desensitized research-team summary dated 15 October 2026. Those summary values, deterministic examples, a data dictionary, and acquisition instructions are included in the reproduction package. Raw archived order-book responses, account records, private implementation details, and production parameters are not included. The absence of raw responses limits reproduction to summary arithmetic and deterministic tests.

Code availability

A clean-room Python reference implementation and unit tests accompany the paper. They perform fixed-quantity book walking, funding-cash calculations, closed-position and error-budget identities, event counts, NA propagation, and figure generation. They contain no order submission, credentials, private endpoints, production rules, or model fitting. The scripts and their output manifest are included in the handoff package.

Generative AI statement

Generative AI tools assisted with literature discovery, language drafting and editing, code drafting, and multi-role technical review. Source selection, numerical provenance, confidentiality boundaries, corrections, and final responsibility rest with Hang Liu and humarb Research. No AI system is listed as an author. The multi-role review included in the handoff is an internal AI-assisted quality check, not external peer review.

Originality and review status

The manuscript was newly structured around the event-time accounting contribution and checked locally for textual overlap with the supplied companion materials. The local screen is not iThenticate, Turnitin, or a comprehensive internet plagiarism certification. The paper is an author-handoff preprint and has not undergone external peer review. The DOI and Zenodo record are given on the title page.

Research and investment disclaimer

This paper is for research and methodological information only. It does not provide investment, legal, accounting, or tax advice; does not recommend a trade, venue, leverage level, or position size; and does not promise income, execution, capital protection, or protection from liquidation. Public-book paper values are not live-trading outcomes. Digital-asset perpetual positions can involve substantial market, liquidity, operational, counterparty, and legal risks.

Acknowledgements

The research and manuscript were completed within humarb Research's collaborative research process. humarb.ai is the research channel used to document and distribute the work. The author retains responsibility for the final manuscript and its evidence boundaries.

References

Damien Ackerer, Julien Hugonnier, and Urban Jermann. Perpetual futures pricing. Mathematical Finance, 36(3):481–499, 2026. doi:10.1111/mafi.70018. https://doi.org/10.1111/mafi.70018.

Robert Almgren and Neil Chriss. Optimal execution of portfolio transactions. The Journal of Risk, 3(2):5–39, 2001. doi:10.21314/jor.2001.041. https://doi.org/10.21314/jor.2001.041.

David H. Bailey, Jonathan M. Borwein, Marcos López de Prado, and Qiji Jim Zhu. The probability of backtest overfitting. The Journal of Computational Finance, 20(4):39–69, 2017. doi:10.21314/JCF.2016.322. https://doi.org/10.21314/JCF.2016.322.

BingX. Swap market data: Public API reference, 2026a. URL https://github.com/BingX-API/api-ai-skills/blob/main/skills/swap-market/api-reference.md. Official BingX-API repository; accessed 4 November 2026.

BingX. Perpetual futures: Fee schedule, 2026b. URL https://bingx.com/en/support/articles/360046487573. Official venue documentation; accessed 4 November 2026.

BingX. Perpetual futures funding rate mechanism explained, 2026c. URL https://bingx.com/en/support/articles/14857605906575. Official venue notice, 12 January 2026; accessed 4 November 2026.

Riccardo De Blasis and Alexander Webb. Arbitrage, contract design, and market structure in Bitcoin futures markets. Journal of Futures Markets, 42(3):492–524, 2022. doi:10.1002/fut.22305. https://doi.org/10.1002/fut.22305.

Gate. Gate API v4 documentation, 2026a. URL https://www.gate.com/docs/developers/apiv4/. Official public API documentation; accessed 4 November 2026.

Gate. Public perpetual contract specifications: BTC_USDT and RSR_USDT, 2026b. URL https://api.gateio.ws/api/v4/futures/usdt/contracts/BTC_USDT. Public contract endpoint; RSR endpoint uses the same path with RSR_USDT; accessed 4 November 2026.

Gate. Futures trading fee calculation, 2026c. URL https://www.gate.com/help/futures/futures_logic/22079. Official venue documentation; accessed 4 November 2026.

Gate. Futures funding rate and funding fee overview, 2026d. URL https://www.gate.com/help/futures/futures-logic/27569/funding-rate-and-funding-fee. Official venue documentation; accessed 4 November 2026.

Peter Reinhard Hansen. A test for superior predictive ability. Journal of Business & Economic Statistics, 23(4):365–380, 2005. doi:10.1198/073500105000000063. https://doi.org/10.1198/073500105000000063.

Campbell R. Harvey, Yan Liu, and Heqing Zhu. …and the cross-section of expected returns. Review of Financial Studies, 29(1):5–68, 2016. doi:10.1093/rfs/hhv059. https://doi.org/10.1093/rfs/hhv059.

Songrun He, Asaf Manela, Omri Ross, and Victor von Wachter. Fundamentals of perpetual futures, 2022. URL https://arxiv.org/abs/2212.06888. arXiv:2212.06888; version 7 consulted, 17 September 2026; https://doi.org/10.48550/arXiv.2212.06888.

Hang Liu. Measuring executable net carry in cross-venue perpetual funding-rate arbitrage, 2026a. URL https://doi.org/10.5281/zenodo.23868361. humarb Research 01; Zenodo preprint.

Hang Liu. Framing net carry for cross-venue perpetual funding positions: A settled-history, depth-walked, and liquidation-aware method, 2026b. URL https://doi.org/10.5281/zenodo.23897013. humarb Research 02; Zenodo preprint.

Igor Makarov and Antoinette Schoar. Trading and arbitrage in cryptocurrency markets. Journal of Financial Economics, 135(2):293–319, 2020. doi:10.1016/j.jfineco.2019.07.001. https://doi.org/10.1016/j.jfineco.2019.07.001.

André F. Perold. The implementation shortfall. The Journal of Portfolio Management, 14(3):4–9, 1988. doi:10.3905/jpm.1988.409150. https://doi.org/10.3905/jpm.1988.409150.

Andrei Shleifer and Robert W. Vishny. The limits of arbitrage. The Journal of Finance, 52(1):35–55, 1997. doi:10.1111/j.1540-6261.1997.tb03807.x. https://doi.org/10.1111/j.1540-6261.1997.tb03807.x.

James E. Smith and Robert L. Winkler. The optimizer’s curse: Skepticism and postdecision surprise in decision analysis. Management Science, 52(3):311–322, 2006. doi:10.1287/mnsc.1050.0451. https://doi.org/10.1287/mnsc.1050.0451.

Halbert White. A reality check for data snooping. Econometrica, 68(5):1097–1126, 2000. doi:10.1111/1468-0262.00152. https://doi.org/10.1111/1468-0262.00152.

A Algebraic derivations

A.1 Basis-minus-crossing identity

Start with the right-hand side of $B - X$ from Equations (7) and (8). Within the common factor Q, the long-entry terms are $-P_L^e + m_L^e$, and the long-exit terms are $P_L^x - m_L^x$. Their sum is $P_L^x - P_L^e$ plus $m_L^e - m_L^x$. The short-entry terms are $P_S^e - m_S^e$, and the short-exit terms are $-P_S^x + m_S^x$. Their sum is $P_S^e - P_S^x$ plus $-m_S^e + m_S^x$. Adding the two legs gives

$$\begin{aligned} B - X = {} & Q(P_L^x - P_L^e) + Q(P_S^e - P_S^x) \\ & + Q(m_L^e - m_L^x) + Q(m_S^x - m_S^e) \\ & + Q(m_S^e - m_S^x) - Q(m_L^e - m_L^x). \end{aligned}$$

Every mid-price term cancels, leaving the four-fill price P&L. The derivation uses no assumption about the sign of the spread, the size of crossing, or convergence to zero.

A.2 Error-budget identity

Let the scenario components be $C_0, B_0, X_0, \Phi_0, O_0$, and use unindexed letters for the outcome. Their nets are

$$\begin{aligned} S_0 &= C_0 + B_0 - X_0 - \Phi_0 - O_0, \\ \Pi &= C + B - X - \Phi - O. \end{aligned}$$

Subtracting yields

$$\begin{aligned} \Pi - S_0 = {} & (C - C_0) + (B - B_0) - (X - X_0) \\ & - (\Phi - \Phi_0) - (O - O_0). \end{aligned}$$

Define signed adverse contributions as $-\Delta C$, $-\Delta B$, ΔX, $\Delta\Phi$, and ΔO. Their sum is $S_0 - \Pi$. Thus a positive adverse sum is an overestimate, while a negative value is an underestimate. No absolute values are taken, so offsetting components remain visible.

A.3 Joint cash variance

For $D = S - L$, linearity of expectation gives $\mathbb{E}[D] = \mathbb{E}[S] - \mathbb{E}[L]$. Variance gives

$$\mathrm{Var}(D) = \mathrm{Var}(S) + \mathrm{Var}(L) - 2\,\mathrm{Cov}(S, L).$$

If both standard deviations equal one and $\mathrm{Corr}(S, L) = \rho$, this becomes $2(1 - \rho)$. The values 0.447 and 1.414 in the main text follow for $\rho = 0.9$ and $\rho = 0$, respectively. This is a moment identity, not a distributional calibration.

B Reference audit algorithms

B.1 Quantity-closed book walk

The following protocol is sufficient for an independent implementation using any exact-decimal arithmetic library:

1. Verify contract type, common base asset, settlement currency, multipliers, steps, and minimums. Return unsupported/NA if any required identity is unknown.
2. Construct the common feasible quantity lattice. Reject an empty positive grid.
3. Walk both entry books at the candidate Q. If either book is insufficient, reduce Q according to the pre-specified rule or report an incomplete fill; never synthesize depth.
4. Record the four actions as long entry, short entry, long exit, and short exit, all at the same Q. Preserve level-level prices and quantities.
5. Compute each VWAP as level price times level quantity divided by Q. Compute each fee from actual fill notional and its evidenced rate.
6. Attribute mids to the same entry and exit information versions. Compute B, X, and price P&L and require their algebraic residual to be zero within exact arithmetic.

B.2 Event reconciliation

1. Construct the expected eligible-event schedule for each leg from the contract version effective in the declared half-open window.
2. Store expected event time, venue-recorded settlement time, and local receipt time as separate fields.
3. Match expected events to records one-to-one using a documented venue tolerance. Reject a second use of the same record.
4. Do not merge same-minute records until venue semantics establish whether they are separate cash flows, corrections, or updates.
5. For each matched event, archive quantity immediately before the event, settled rate, and official valuation mark. If the mark is unavailable, label any substituted value as a proxy.
6. Classify each expected event as matched cash, verified no event, missing record, ambiguous semantics, or ineligible. Only the first enters exact cash; a verified empty set yields zero; unresolved states block exact totals.
7. Reconcile computed cash to account-level cash when available. Differences remain reconciliation items rather than being forced to zero.

B.3 Scenario-key validation

Before adding or comparing components, normalize the scenario key and require exact agreement on route, contract version, quantity, quote currency, fee basis,

holding view, target time, and information version. Evaluate the final objective for every candidate horizon before selecting one. The key-validation rule rejects mixed-horizon sums rather than attempting to repair them with an assumed conversion.

C Evidence-state dictionary

Table 7: Minimum evidence states for event-time accounting

State	Required evidence	Prohibited interpretation
Verified empty event set	Effective schedule and eligibility rule establish no eligible event	Missing history or unverified coverage
Matched official cash	One-to-one event match, eligible quantity, official rate and mark, or authoritative account cash	A proxy mark or fixed-notional estimate
Snapshot cost proxy	Current public book with freshness, contract, and proxy labels	Future fill, four-book outcome, or official account fee
Four-book paper	Entry and exit books at common Q, four VWAPs, public fee basis	Submitted orders, actual fills, or guaranteed execution
Actual fills	Order/fill records and quantities, reconciled to fees and funding	Complete cost evidence without the remaining records
Coverage unverified	Schedule or record semantics are incomplete or ambiguous	Zero cash, complete coverage, or exact net
NA / unsupported	Required evidence is absent or contract mechanics are outside scope	Zero, loss, conservative value, or denominator removal

D Reproduction data dictionary

The reproduction package contains `paper_route_arm_summary.csv`, `error_attribution_summary.csv`, `toy_examples.json`, and `data_dictionary.csv`. The route-arm file records route, asset, long and short venues, base quantity, arm, price P&L, fees, net before funding, funding amount and evidence, scenario net including reported funding, and exact account net. Amounts are USDT and quantities are base-asset units. Exact account net is NA for every row. Blank fields mean not reported, not zero.

The error file records baseline and realized ex-funding values, signed D_{exec}, and the disclosed basis, crossing, fee, and funding-scenario components. Blank component cells for routes B and C are not decomposed in the supplied summary. The toy file stores decimal strings to avoid binary floating-point changes in worked examples. Input hashes and output checks are written to `outputs/verification.json`.

Core reproduction is standard-library Python. Figure generation additionally uses Matplotlib. The scripts perform no network calls and require no credentials. Running the reproduction script, six unit tests, and figure script takes only a few seconds in the delivered environment. The generated figures are vector SVG and PDF files with source notes and units.

E Attempted configurations and reporting inventory

The source protocol's first version was aborted after a venue rejected an excessive depth request with HTTP 400. It produced no qualifying paired books and no entry or exit outcomes. Version 2 changed only the depth request. Three routes and two holding arms were then retained; the route list was not changed after exit results were observed. All six resulting net-before-funding values are reported, as are the funding proxies, evidence restrictions, and current-state error summaries. The fixed-dollar crossing control and the uniform fee stress are also reported. No alternative profitable route subset is withheld to create a more favorable narrative.

The analysis did not estimate a predictive model, vary model hyperparameters, or construct a portfolio. Therefore there is no model-search count to report. A future selection exercise must archive the full route, parameter, and horizon candidate set and disclose unsuccessful configurations. It must also preserve the aborted-request log and any post-freeze corrections so that protocol development can be distinguished from outcome-driven selection.

F External pre-registration checklist for a confirmation run

A confirmation run should complete the following sequence before outcomes are available: (i) freeze route discovery, quantity, fee basis, freshness rules, arms, and analysis code; (ii) publish or independently timestamp the frozen protocol and its hash; (iii) archive the complete candidate set, including aborted requests; (iv) collect raw public responses and contract versions without orders; (v) compute all arms without post-outcome edits; (vi) release raw responses, hashes, and an independent re-parser; and (vii) report every attempted configuration and every evidence limitation. If predictive claims are added, the protocol must also freeze the training/validation split, loss function, candidate models, multiple-testing procedure, and dependence-aware inference plan before final evaluation.