EROS
QUANTITATIVE INSTITUTE

WORKING PAPER SERIES
Market Microstructure & Behavioral Finance ·
October 2026

# Echo, Not Forecast:

## Trade-Call Bursts on Binance Square and Subsequent Perpetual Returns

**Shiyu Guan** **IRIS**
EROS Quantitative Institute, China

### Abstract

We examine whether bursts of explicit "trade calls" on exchange-embedded social feeds predict subsequent cryptocurrency returns or conversely serve as a reliable contrarian indicator. Continuously archiving the recommendation feed of Binance Square every 10 minutes from July 24 to October 15, 2026 (12 weeks), we capture 181,509 distinct posts across 9.34 million feed impressions. Using a multi-stage large language model (LLM) annotation pipeline with cross-audit, we extract ticker linkages, directional stances, and content categories, aligning feed-visibility timestamps with hourly OKX perpetual futures prices and calibrated execution costs. Under a preregistered research design featuring an interleaved weekly discovery–validation split and an exhaustive trial ledger (10 hypothesis families, 936 total runs including failures and diagnostics), we establish three findings. **(i) Pre-burst momentum echo**: When a token triggers a trade-call burst on the recommendation feed, it has already outperformed the tradable universe by +5.9% to +15.8% over the preceding 24 hours (24/24 configurations positive, date-clustered $t \in [4.4, 7.6]$; +13.7%, $t = 8.4$ in locked validation). Call bursts are a lagging reflection of realized price run-ups. **(ii) Zero gross drift and friction-driven net losses for followers**: Buying immediately after a burst (holding 1 or 4 hours) yields gross excess returns between −1.08% and +0.41%; under zero transaction costs, 10 of 24 configurations are positive and 14 are negative, with all $|t| < 1.9$. Once realistic execution fees and bid–ask spread/impact costs are deducted, 23 of 24 configurations lose money (up to −293.0 USDT on 2,000 USDT capital), and 85%–92% of total friction stems from spread and market impact in thinner altcoins. **(iii) Out-of-sample falsification of contrarian fading**: Fading call bursts over a 72-hour horizon (`S1b`) appears highly profitable in discovery (+336.1 USDT, $t = 2.56$, Sharpe 3.12), yet nearly half of the profit comes from a single token (`SOON`) and the strategy reverses to −191.7 USDT ($t = -0.83$) in locked validation. A matched price-reversal control experiment (`D1`) reveals that called tokens underperform uncalled tokens with identical prior run-ups when shorted in validation (paired net difference −503.0 bp, $t = -2.46$), proving that the apparent discovery reversal is driven purely by mechanical mean reversion after sharp price spikes rather than social media sentiment. Finally, a price-free measurement shows that 7-day net bullishness on Binance Square is virtually uncorrelated with actual account long–short ratios on OKX (−0.08) and Gate (−0.11), whereas the two exchanges' long–short ratios correlate at +0.56, indicating a fundamental disconnect between vocal posters and leveraged position holders.

**Keywords:** Binance Square · social trade calls · retail attention · LLM annotation · market microstructure · preregistration · out-of-sample falsification

## 1. Introduction

### *1.1. Research question*

Exchange-integrated social feeds have become a primary venue where retail cryptocurrency investors consume market narratives and trade ideas. Among these platforms, Binance Square—embedded directly within the world's largest cryptocurrency exchange—hosts a constant stream of commentary, project promotions, and explicit "trade calls" (posts urging readers to buy or short specific tokens). This environment motivates two opposing retail heuristics: momentum followers buy into tokens experiencing a sudden surge of bullish calls on their feed, whereas self-styled contrarians treat a flood of retail calls as a local top signal and enter short positions.

Both heuristics rest on a common, untested premise: **that bursts of trade calls on an algorithmic recommendation feed contain incremental predictive information for subsequent returns beyond the price path that already occurred.** In this paper, we subject both hypotheses to a preregistered, look-ahead-free empirical test with realistic perpetual futures execution costs, addressing three sequential questions:

1. **Pre-event trajectory**: Where does a token's price stand by the time a burst of trade calls actually surfaces on a user's recommendation feed?
2. **Following the burst (Follow)**: If an investor trades in the direction of a call burst or sudden cold-coin hype for 1 to 4 hours, what are the gross and net expected returns, and how much of the outcome is dictated by market microstructure friction?
3. **Fading the burst (Fade)**: Does shorting called tokens over 24 to 72 hours survive out-of-sample validation, and can any observed post-burst pullback be distinguished from mechanical price reversal following an extreme run-up?

### *1.2. Literature context*

Our study connects to the classic literature on retail investor attention and social media sentiment. Barber and Odean[1] demonstrate that individual investors are net buyers of attention-grabbing stocks—those experiencing extreme one-day returns or abnormal volume—creating temporary price pressure that subsequently reverses. Da, Engelberg, and Gao[2] confirm this attention-induced price pressure using search frequency. On internet message boards, Antweiler and Frank[3] show that posting volume predicts volatility and trading volume, but that return predictability is economically negligible relative to transaction costs; Cookson and Niessner[4] further link social network disagreement to trading activity.

Translating these insights to exchange-embedded crypto feeds introduces three distinct institutional and methodological challenges. First, **algorithmic feed latency**: retail users do not observe a chronological firehose of all posts at their creation timestamp (`posted_at`); they observe a curated recommendation feed where posts surface only after algorithmic distribution, introducing a median visibility delay of roughly one hour and a 90th-percentile delay of 16 hours. Second, **severe tail illiquidity**: trade calls disproportionately target small- and mid-cap altcoins whose order books impose substantial bid–ask spreads and price impact. Third, **multiple-testing inflation**[5, 6]: because social signals admit countless parameterizations (window lengths, burst multiples, author deduplication, holding horizons), unpreregistered backtests easily manufacture false positives.

This work also complements our institute's prior investigation of Chinese equity retail forums (6.26 million LLM-annotated posts on NGA's stock board), which established that "the distribution of retail attention carries cross-sectional information, whereas the directional content of retail opinions does not." Here we shift to hourly cryptocurrency perpetuals and focus specifically on explicit trade calls.

### 1.3. Summary of findings and contributions

Across 12 weeks of continuous archiving (181,509 distinct posts) and 936 logged trial runs, our conclusion is unambiguous: **trade-call bursts on Binance Square are an echo of past returns, not a forecast of future returns.**

> **Core Empirical Verdict (Three-Tier Structure):**
>
> **1. Pre-burst momentum echo (robust across all samples):** Call bursts do not precede rallies; they follow them. Over the 24 hours prior to entry, called tokens have already outperformed the tradable universe by +5.9% to +15.8% (24/24 discovery configurations positive, $t \in [4.4, 7.6]$; +13.7%, $t = 8.4$ in locked validation).
>
> **2. Zero gross edge and friction-driven losses when following:** Post-entry 1–4 hour gross excess returns hover near zero (−1.08% to +0.41%; all $|t| < 1.9$ under zero costs). Because called altcoins suffer wide spreads and market impact (averaging $\approx$ 0.9 USDT friction per 200 USDT trade), 23 of 24 follow configurations lose money net of costs, and none is significantly positive.
>
> **3. Out-of-sample falsification of contrarian fading:** Shorting call bursts for 72 hours (`S1b`) yields +336.1 USDT ($t = 2.56$) in discovery but collapses to −191.7 USDT ($t = -0.83$) in locked validation. A matched price-reversal control (`D1`) outperforms `S1b` by 503.0 bp ($t = -2.46$) in validation, showing that the apparent in-sample reversal is purely mechanical pullback after a price spike.

Our contributions are threefold. First, we construct a look-ahead-free recommendation-feed archive anchored strictly on feed-visibility timestamps (`first_seen_at`) rather than authoring timestamps (Section 2). Second, we provide a transparent preregistered falsification workflow—logging all 936 runs, decomposing microstructure costs across retail fee tiers (Sections 5–6), and documenting how an apparent 3.12-Sharpe contrarian strategy fails out of sample and under matched-momentum controls (Section 7). Third, we document a stark cross-source disconnect between social media stance and exchange account long–short ratios (Section 8).

## 2. Data and Recommendation-Feed Architecture

### 2.1. Continuous feed sampling and visibility latency

We collect data from the web recommendation feed endpoint of Binance Square (`feed-recommend/list`). From July 24 to October 16, 2026 (84 complete calendar days, 12 weekly blocks), our collector queried the first 40 pages of the recommendation feed every 10 minutes, supplemented by post-detail and sitemap queries. This process recorded 9,337,057 raw feed impressions, yielding 181,509 deduplicated posts (Table 1).

Crucially, **this sample represents the platform's algorithmic recommendation stream rather than the universe of all published posts**. For studying retail trade-call chasing, the recommendation feed is the exact object of interest: isolated posts that never receive algorithmic distribution cannot trigger coordinated retail order flow.

Table 1: Overview of the Binance Square archive, LLM annotation, and perpetual market mapping.

| | |
|---|---|
| Archive window | 2026-07-24 – 2026-10-16 (84 days, 12 weekly blocks) |
| Raw recommendation-feed impressions | 9,337,057 (every 10 min, ≈ 144 rounds/day) |
| Deduplicated distinct posts | 181,509 |
| Visibility delay (`first_seen_at` − `posted_at`) | Median ≈ 1 h, P90 ≈ 16 h (0 negative) |
| Distinct post bodies re-read by LLMs | 176,912 |
| Valid stance coverage across blocks | 89.1% – 93.9% |
| Posts mentioning at least one token | 115,798 (63.8% of all posts) |
| Distinct raw ticker strings extracted | 1,699 |
| Unambiguously mapped OKX perpetuals | 408 (24.0% of strings; 78.8% of token posts) |
| Posts excluded for homonym ambiguity | 1,707 (0.94% of all posts, e.g., `CASH`) |
| Primary execution venue | OKX USDT perpetuals, certified 1-hour bars |
| Tradable universe definition | Rolling 30d OKX top 150 by USD turnover |

Notes: Source and integrity checks are documented in the Session A Data Audit. An hourly candle is certified usable only if all 60 constituent 1-minute candles are present and verified.

In high-frequency social media studies, timestamp alignment is a primary source of look-ahead bias. A post is authored at `posted_at`, but only surfaces in a user's recommendation feed at `first_seen_at` after moderation and algorithmic ranking. Across our archive, the median delay between `posted_at` and `first_seen_at` is approximately 1 hour (ranging from 0.45 to 2.72 hours across weekly blocks), while the 90th percentile reaches 10.3 to 21.0 hours (≈ 16 hours overall). Aligning signals at `posted_at` implicitly assumes traders can act on posts hours before the recommendation algorithm distributes them—a flaw that conflates pre-distribution price surges (which itself causes the algorithm to promote the post) with post-signal returns. **We therefore anchor all signal timestamps strictly at `first_seen_at`**, executing trades at the open of the first complete hourly candle at least 5 minutes after `first_seen_at`.

### *2.2. Multi-model LLM annotation and cross-audit*

To convert 181,509 noisy social posts into structured records, we deploy a three-stage LLM annotation pipeline extracting: (1) mentioned token symbols, (2) overall directional stance (bullish, bearish, mixed, neutral), and (3) content category (explicit trade call, market analysis, news, promotion/airdrop giveaway, personal update, etc.).

In Stage 1, `doubao-seed-2.1-lite` generated initial labels for 143,754 posts (with 30,506 posts carrying baseline labels from `qwen3.8-max`). In Stages 2 and 3, `GLM-5.3 Flash` and `GPT-6.1 Sol` independently re-read all 176,912 unique post texts. Production signals rely on `GPT-6.1 Sol` while retaining inter-model disagreement flags; a subsequent cross-audit of 527 sampled posts corrected unambiguous edge-case errors affecting 27 posts.

To guard against LLM look-ahead leakage (models using pretraining knowledge of late-2026 crypto prices to infer stance), we verified model knowledge cutoffs: `GPT-6.1 Sol` has a documented knowledge cutoff of April 30, 2026, strictly predating our July 24, 2026 sample start. Because official documentation for `GLM-5.3 Flash` does not disclose a verifiable cutoff date, we record this as a formal limitation in Section 8. Furthermore, directional stance attribution enforces a **strict single-coin rule**: a post's bullish or bearish stance is assigned to a token only if the post references exactly one valid token; multi-coin roundups contribute to attention volume but not to directional calls.

### *2.3. Instrument mapping, market data, and the mid-September feed break*

Of the 1,699 distinct token strings extracted, 408 map unambiguously to historical OKX USDT perpetual contracts after removing homonym collisions (such as `CASH`, affecting 1,707 posts). While 24.0% of raw ticker strings map to OKX perpetuals (reflecting thousands of micro-cap on-chain tokens), these 408 contracts cover 91,205 of the 115,798 token-mentioning posts (78.8%). Hourly prices, 30-day rolling turnover universes (top 150), daily contract specifications (`lotSz`, `minSz`, `ctVal`), and realized funding rates are drawn from pinned snapshots of our laboratory's data lake.

Prior to computing any return series, a price-free structural audit of the corpus uncovered two foundational properties:

First, **a mid-September platform recommendation regime shift**. With our collector code, proxy exit, and polling frequency held completely constant, the composition of the Binance Square recommendation feed shifted abruptly in mid-September 2026 (Table 2): daily distinct authors appearing in the feed nearly doubled from $\approx$ 1,050 to 1,800–2,200, the share of first-time weekly authors rose from 14% to 30%, and even among pre-existing authors, the share of explicit trade calls declined from 26.7% to 18.5%. This platform-driven shift toward lighter, campaign-driven posts implies that raw post-count thresholds would suffer severe non-stationarity under a naive chronological train/test split.

Table 2: Structural shift in the Binance Square recommendation feed in mid-September 2026.

| Feed metric | Pre-break (Aug 20–Sep 01) | Post-break (Sep 14+) |
|---|---|---|
| Daily distinct posts in recommendation feed | $\approx$ 2,850 | 3,300 – 4,000 |
| Daily distinct authors in recommendation feed | $\approx$ 1,050 | 1,800 – 2,200 |
| Share of new authors among weekly posts | 14% | 30% |
| Share of explicit calls among pre-break authors | 26.7% | 18.5% |

Notes: Derived from the price-free measurement study (collector held constant). Among post-break new authors, 64% appear only once and write shorter posts (median 127 characters vs. 227 for established authors).

Second, **social stance is orthogonal to exchange positioning** (detailed in Section 8). Together, these two facts motivated our share-normalized signal definitions and interleaved weekly split.

## 3. Research Design and Anti-Overfitting Protocol

### *3.1. Event definitions and twenty-four follow configurations*

To prevent single-account spamming from triggering signals, we deduplicate posts by `(author_id, symbol)` within rolling 6-hour windows, keeping only the earliest visible post. We then formalize retail momentum-chasing through two preregistered follow families totaling 24 configurations:

1. **Follow explicit call bursts (`S1a`, 16 configurations)**: Restricted to single-token posts classified as explicit trade calls. A token triggers a burst when, over the past $W \in \{6, 24\}$ hours, it accumulates at least $M \in \{3, 5\}$ deduplicated trade calls and its share of all token-mentioning posts in that window reaches at least $K \in \{3, 5\}$ times its own trailing 14-day mean share (normalizing by share neutralizes the Table 2 feed expansion). Trade direction follows the net majority of bullish versus bearish calls in the window (overwhelmingly long in practice), with repeat triggers for the same token within 24 hours merged into a single event. Positions enter at the next hourly open and hold for $H \in \{1, 4\}$ hours, yielding $2 \times 2 \times 2 \times 2 = 16$ configurations (`S1a-W{6,24}-K{3,5}-M{3,5}-H{1,4}`).
2. **Buy sudden hype in neglected coins (`S2b`, 8 configurations)**: Targets "cold" tokens outside the top 20 by trailing 28-day attention share. A long entry triggers when the token's 24-hour discussion share exceeds the $K \in \{95, 99\}$th percentile of its own trailing 28-day distribution with at least $M \in \{3, 5\}$ distinct authors, held for $H \in \{1, 4\}$ hours ($2 \times 2 \times 2 = 8$ configurations, `S2b-W24-K{95,99}-M{3,5}-H{1,4}`).

Alongside `S1a` and `S2b`, our protocol preregistered eight additional hypothesis families (Table 3): contrarian fading of call bursts over 24/72 hours (`S1b`), shorting cold-coin hype over 24/72/120 hours (`S2a`), shorting promotion/giveaway bursts (`S3`), cross-sectional sentiment contrarian and momentum portfolios (`S4` / `S4r`), disagreement predicting volatility and returns (`S5v` / `S5r`), and large-author leadership (`S6`, which was frozen as "insufficient data" prior

to any return calculation because point-in-time follower counts at first observation were null across all 181,509 posts).

**Non-independence of the 24 configurations**: The 24 follow configurations analyzed in Sections 4–6 **do not constitute 24 independent hypothesis tests**. Configurations with $H = 1$ and $H = 4$ share the exact same underlying trigger events and entry timestamps; adjacent threshold grids (e.g., $K = 3$ vs. $K = 5$) also share substantial event overlap. In total, the 24 configurations represent 8 base event definitions for `S1a` and 4 for `S2b`, providing a sensitivity map across parameter choices rather than independent trials.

### *3.2. Interleaved weekly split, preregistration, and the 936-run ledger*

To eliminate data dredging and selective reporting, we enforce three structural safeguards:

Table 3: Complete trial ledger across all 10 hypothesis families: 936 runs, failures included.

| Family | Meaning | Attempts | Completed | Failed |
|---|---|---|---|---|
| S1a | Follow call bursts (1/4 h) | 113 | 112 | 1 |
| S1b | Fade call bursts (24/72 h) | 140 | 140 | 0 |
| S2a | Short cold-coin hype | 85 | 85 | 0 |
| S2b | Buy cold-coin hype | 61 | 61 | 0 |
| S4 | Fade one-sided sentiment | 37 | 36 | 1 |
| S4r | Follow one-sided sentiment | 37 | 36 | 1 |
| S5v | Disagreement to volatility | 49 | 48 | 1 |
| S3 | Promotion posts, short | 155 | 154 | 1 |
| S5r | Disagreement to returns | 259 | 258 | 1 |
| S6 | Large authors (no data) | 0 | 0 | 0 |
| Total | | 936 | 930 | 6 |

Notes: Source: `TRIAL-COUNTS.json`. Comprises 916 discovery runs plus 10 original and 10 engineering-corrected validation runs (Appendix D). Each run bundles main and unrounded accounts, capital variants, and diagnostics (`D1`–`D3`); these are not 936 independent hypotheses. No family passed all out-of-sample validation gates.

First, **preregistration with hashed amendments**. All signal rules, directional priors, discovery qualification gates, and validation criteria were committed before accessing return data. Twelve subsequent engineering or specification amendments (`A1`–`A12`) and two paper-stage summary preregistrations (`P1` follower table, `P2` retail cost tiers) were individually documented and SHA256-hashed.

Second, **interleaved weekly discovery–validation split**. To ensure balanced representation before and after the mid-September feed break, we partition the 84-day sample into twelve 7-day blocks (B0, . . . , B11). Within every consecutive triad of weeks, the first two blocks are assigned to **Discovery** (8 blocks, 56 days) and the third block is locked in **Validation** (B2, B5, B8, B11; 4 blocks, 28 days, with two blocks pre-break and two post-break). Any discovery trade whose holding window would cross into a validation block is excluded in its entirety. Code-level assertions blocked access to validation prices until discovery freeze files were pushed to the remote repository.

Third, **universal trial logging**. Every execution across the project—936 runs in total, including 6 early engineering failures—is recorded in an append-only ledger (Table 3). Discovery qualification required an event family to simultaneously achieve date-clustered net excess $t \geq 2.0$, at least 25 distinct event dates, realistic account Sharpe $\geq 1.0$, positive net P&L in at least 60% of active discovery blocks, and positive net P&L after removing the top three contributing tokens. **Nine of the ten families failed already in discovery; only `S1b` qualified for validation, where it too was falsified.**

*3.3. Account simulation and microstructure cost calibration*

All strategies are evaluated via lot-rounded account simulations on 2,000 USDT initial capital (up to 10 concurrent positions, 10% target notional or ≈ 200 USDT per trade). Excess returns are measured relative to the equal-weighted tradable top-150 universe, and all event-level $t$-statistics are clustered by calendar date (UTC+8) to account for cross-sectional correlation on busy market days.

Each trade incurs three cost components: (1) **exchange fees**: 2.5 bp per side in our baseline research tier (R), reflecting 50% rebate on OKX's 5 bp taker schedule; (2) **bid–ask spread and market impact (price cost)**: calibrated from our institute's live execution engine as a function of each token's trailing 30-day median daily turnover and daily volatility, multiplied by a 1.5× safety factor; and (3) **funding payments**: realized OKX funding rates accrued at settlement timestamps during the holding window. Section 6 further evaluates zero-cost (G), unrebated retail perpetual (U1, 5.0 bp/side), and retail spot approximation (U2, 10.0 bp/side, zero funding) tiers.

## 4. Finding 1: By the Time Calls Surface, the Move Has Happened

Figure 1 answers our first question by plotting the average cumulative excess price path from 24 hours before entry ($t = -24$) to 120 hours after entry ($t = +120$) for the primary burst definition (`W24-K5-M3`).

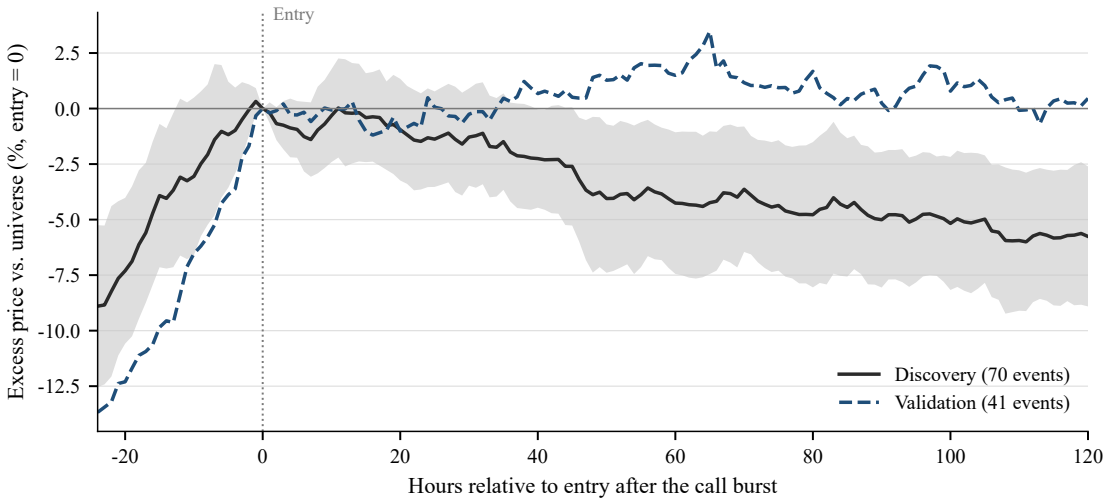



Figure 1: Event-time excess price path around a trade-call burst relative to the equal-weighted tradable universe (entry open normalized to 0). The solid line shows the discovery period (70 complete-path events across 8 weekly blocks) with its date-clustered 95% confidence band in grey; the dashed line shows the locked validation period (41 complete-path events across 4 weekly blocks) under the identical `W24-K5-M3` definition.

The trajectory splits sharply at entry ($t = 0$). To the left of $t = 0$, both the discovery and validation curves climb steeply: over the 24 hours preceding entry, tokens triggering a `W24-K5-M3` burst have already outperformed the market by +8.9% ($t = 5.1$) in discovery and +13.7% ($t = 8.4$, 41 events across 14 dates) in locked validation.

Table 4: Twenty-four follow configurations in discovery: pre-burst run-up, post-entry gross excess returns, and net account results.

| Config | Events | Pre 24h (%) | $t$ | 1h (%) | 4h (%) | 72h (%) | Trades | Net P&L | Net-excess $t$ |
|---|---|---|---|---|---|---|---|---|---|
| S1a-W6-K3-M3-H1 | 102 | +7.1 | +4.9 | +0.40 | +0.19 | −1.1 | 186 | −95.1 | −1.20 |
| S1a-W6-K3-M3-H4 | 102 | +7.1 | +4.9 | +0.40 | +0.19 | −1.1 | 181 | −131.5 | −1.89 |
| S1a-W6-K3-M5-H1 | 35 | +12.1 | +4.7 | +0.25 | −0.62 | −4.1 | 67 | +21.4 | +0.16 |
| S1a-W6-K3-M5-H4 | 35 | +12.1 | +4.7 | +0.25 | −0.62 | −4.1 | 65 | −24.2 | −0.64 |
| S1a-W6-K5-M3-H1 | 74 | +9.4 | +5.3 | +0.41 | +0.19 | −1.3 | 130 | −62.8 | −1.16 |
| S1a-W6-K5-M3-H4 | 74 | +9.4 | +5.3 | +0.41 | +0.19 | −1.3 | 127 | −130.7 | −1.98 |
| S1a-W6-K5-M5-H1 | 25 | +15.8 | +5.0 | −0.02 | −0.67 | −6.7 | 48 | −9.0 | −0.77 |
| S1a-W6-K5-M5-H4 | 25 | +15.8 | +5.0 | −0.02 | −0.67 | −6.7 | 46 | −30.1 | −0.86 |
| S1a-W24-K3-M3-H1 | 118 | +7.3 | +4.4 | +0.33 | +0.19 | −3.2 | 196 | −105.6 | −1.73 |
| S1a-W24-K3-M3-H4 | 118 | +7.3 | +4.4 | +0.33 | +0.19 | −3.2 | 190 | −189.0 | −1.94 |
| S1a-W24-K3-M5-H1 | 66 | +11.5 | +4.5 | −0.09 | +0.03 | −3.9 | 112 | −108.9 | −2.14 |
| S1a-W24-K3-M5-H4 | 66 | +11.5 | +4.5 | −0.09 | +0.03 | −3.9 | 111 | −182.2 | −1.94 |
| S1a-W24-K5-M3-H1 | 70 | +8.9 | +5.1 | −0.18 | −0.87 | −4.2 | 109 | −136.5 | −2.35 |
| S1a-W24-K5-M3-H4 | 70 | +8.9 | +5.1 | −0.18 | −0.87 | −4.2 | 107 | −131.1 | −2.10 |
| S1a-W24-K5-M5-H1 | 39 | +14.7 | +7.2 | −0.15 | −1.03 | −5.0 | 61 | −37.0 | −0.93 |
| S1a-W24-K5-M5-H4 | 39 | +14.7 | +7.2 | −0.15 | −1.03 | −5.0 | 60 | −61.6 | −1.43 |
| S2b-W24-K95-M3-H1 | 172 | +5.9 | +6.2 | −0.03 | −0.19 | −0.3 | 303 | −293.0 | −3.11 |
| S2b-W24-K95-M3-H4 | 172 | +5.9 | +6.2 | −0.03 | −0.19 | −0.3 | 301 | −263.3 | −1.90 |
| S2b-W24-K95-M5-H1 | 160 | +5.9 | +6.8 | +0.00 | −0.26 | −0.7 | 264 | −288.3 | −2.96 |
| S2b-W24-K95-M5-H4 | 160 | +5.9 | +6.8 | +0.00 | −0.26 | −0.7 | 264 | −271.8 | −1.67 |
| S2b-W24-K99-M3-H1 | 104 | +9.1 | +7.2 | −0.07 | −1.07 | −0.5 | 180 | −183.0 | −2.76 |
| S2b-W24-K99-M3-H4 | 104 | +9.1 | +7.2 | −0.07 | −1.07 | −0.5 | 177 | −203.2 | −1.58 |
| S2b-W24-K99-M5-H1 | 102 | +9.1 | +7.6 | −0.14 | −1.08 | −0.8 | 166 | −216.9 | −3.21 |
| S2b-W24-K99-M5-H4 | 102 | +9.1 | +7.6 | −0.14 | −1.08 | −0.8 | 164 | −212.9 | −1.58 |

Notes: Source: `results/p1_follower_table.csv`. "Events" counts triggers whose full $[-24\text{h}, +120\text{h}]$ price path lies within discovery blocks; "Trades" counts executed account trades. Pre 24h is the excess run-up over the 24 hours before entry with its date-clustered $t$-statistic; 1h/4h/72h are post-entry gross excess returns before costs. Net P&L is in USDT on 2,000 capital at baseline research costs (R tier). H1 and H4 share identical events.

Table 4 confirms that this pre-burst run-up is an invariant feature across all 24 configurations (24/24 positive):

- Across the 16 `S1a` configurations, pre-entry 24-hour excess run-ups range from +7.1% to +15.8%, with date-clustered $t$-statistics between 4.4 and 7.2. Stricter burst thresholds capture even larger prior rallies: requiring at least 5 calls and a 5× attention multiple (`W6-K5-M5` and `W24-K5-M5`) corresponds to prior 24-hour excess gains of +15.8% ($t = 5.0$) and +14.7% ($t = 7.2$).
- Across the 8 `S2b` cold-coin hype configurations, pre-entry 24-hour excess run-ups range from +5.9% to +9.1% ($t \in [6.2, 7.6]$).

This pattern reflects the causal mechanics of social content creation and algorithmic curation: **price rallies generate trade calls, not vice versa.** When an altcoin experiences a sharp intraday breakout, creators rush to publish bullish calls to capture engagement, and the recommendation algorithm promotes those trending posts to retail screens with a one-to-sixteen-hour lag. By the time a user perceives a "chorus of calls" on Binance Square, +6% to +16% of excess appreciation is already in the rearview mirror.

## 5. Finding 2: Buying After the Burst Yields Zero Gross Drift and Net Losses

Looking to the right of $t = 0$ in Figure 1, the steep upward slope vanishes immediately upon entry. Over the subsequent 1 to 4 hours, the price path flattens completely, before drifting mildly lower in discovery or moving sideways in validation over the next 72 to 120 hours.

Across all 24 configurations in Table 4:

- **1-hour post-entry gross excess return**: ranges narrowly from −0.18% to +0.41%;

- **4-hour post-entry gross excess return**: ranges from −1.08% to +0.19% (negative in 15 of 24 configurations);
- **72-hour post-entry gross excess return**: negative across all 24 configurations in discovery (−0.3% to −6.7%), but near zero in validation ($\approx -1.0\%$, $t = -0.3$), confirming the absence of positive follow-through at any horizon.

When executed in the 2,000 USDT account simulation under baseline research costs (R tier), net P&L is overwhelmingly negative (Figure 2 and the final columns of Table 4):

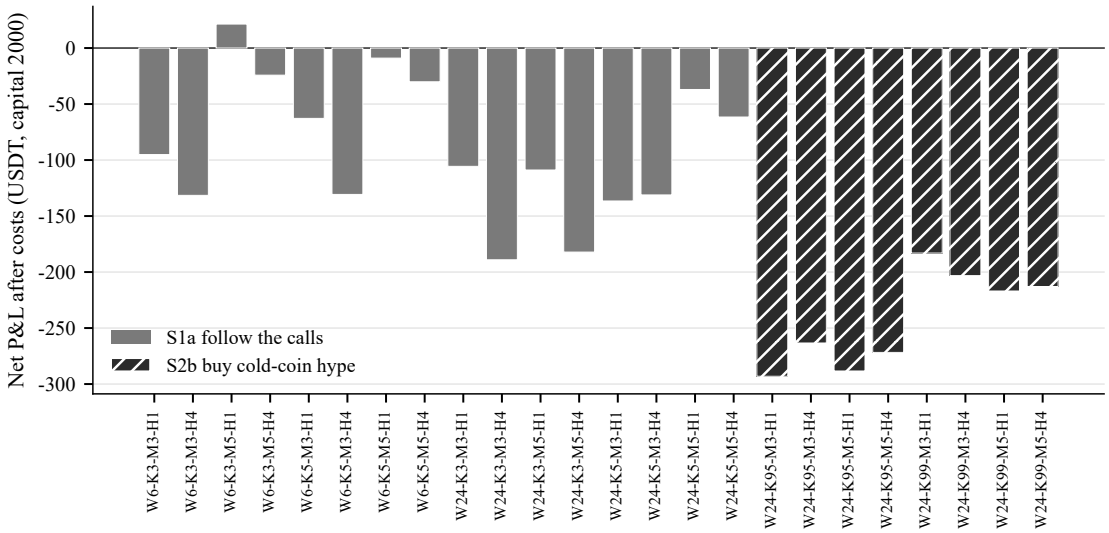



Figure 2: Net account P&L after baseline transaction costs across the 24 follow configurations in discovery (8 weeks, 2,000 USDT initial capital). Solid grey bars denote `S1a` (follow explicit calls, 16 variants); hatched dark bars denote `S2b` (buy cold-coin hype, 8 variants). Twenty-three of 24 variants lose money.

First, **23 of the 24 configurations record net losses**, with a median net loss of −131.3 USDT (6.6% of initial capital over 8 weeks, despite unlevered 10% position sizing). Within `S1a`, 15 of 16 variants are negative (losing up to −189.0 USDT, net-excess $t = -1.94$). Within `S2b`, **all 8 variants lose money**, with net losses between −183.0 and −293.0 USDT (9.2% to 14.7% of capital, $t \in [-3.21, -1.58]$).

Second, **the lone positive variant is statistically indistinguishable from zero and fragile**. Only `S1a-W6-K3-M5-H1` (67 trades) posts a positive net result of +21.4 USDT (+6.0 bp mean net excess), with a date-clustered $t$-statistic of just +0.16. Concentration diagnostics (`D3`) show that this small gain relies entirely on three tickers (`USELESS` +14.8 USDT, `NEAR` +13.2 USDT, `LAB` +12.8 USDT); removing the top three coins flips net P&L to −19.3 USDT. Separating long and short legs (`D2`), its 58 long trades earn +54.0 bp gross ($t = 1.85$) but only +11.0 bp net ($t = 0.25$), while its 9 short trades lose −33.4 bp net ($t = -0.53$).

## 6. Finding 3: Microstructure Cost Decomposition and Retail Fee Sensitivity

### *6.1. Why accounts lose when gross returns are flat*

Do followers lose money because buying after a burst actively predicts an immediate crash, or because transaction costs tax a zero-expectation coin toss? In supplementary experiments `P1` and `P2`, we decompose every trade's net P&L into four exact components:

$$\text{Net P\&L} = \text{Gross P\&L} - \text{Fees} - \text{Price Cost} - \text{Funding Paid}.$$

Table 5 summarizes account performance across four preregistered cost tiers, and Figure 3 plots the vertical shift in net P&L for each configuration (full configuration-level breakdowns appear in Tables A.8 and A.9).

Table 5: Summary of net P&L across 24 follow configurations under four cost tiers (discovery, 8 weeks, 2,000 USDT capital).

| Tier | Cost specification | Lose | Win | Net P&L range | Median |
|---|---|---|---|---|---|
| G | Zero cost (no fees, no spread/impact, no funding) | 14 | 10 | −96.8 to +73.6 | −5.4 |
| R | Baseline research (2.5 bp/side + spread/impact + funding) | 23 | 1 | −293.0 to +21.4 | −131.3 |
| U1 | Retail perpetual (5.0 bp/side + spread/impact + funding) | 23 | 1 | −321.3 to +14.8 | −145.1 |
| U2 | Retail spot approx. (10.0 bp/side + spread/impact, 0 funding) | 23 | 1 | −379.0 to +1.0 | −171.6 |

Notes: Source: `results/p2_net_pnl_by_tier.csv` (preregistration hash `4333bbd3…`). Under the zero-cost G tier, date-clustered $t$-statistics across all 24 configurations lie within $[-1.09, +1.85]$. Tier U2 reuses perpetual spread/impact estimates while setting funding to zero as a first-order spot approximation.

Row G of Table 5 provides a crucial clarification: **without transaction costs, buying after a burst neither wins nor loses**. At zero cost, 10 configurations are positive and 14 are negative (6 of 8 `S1a` 1-hour variants have positive gross P&L between +29.4 and +73.6 USDT), the overall median is −5.4 USDT, and every single configuration has $|t| \leq 1.85$. This is why we refrain from claiming that following calls "always loses" at the gross price level: post-burst 1–4 hour expected excess return is essentially zero.

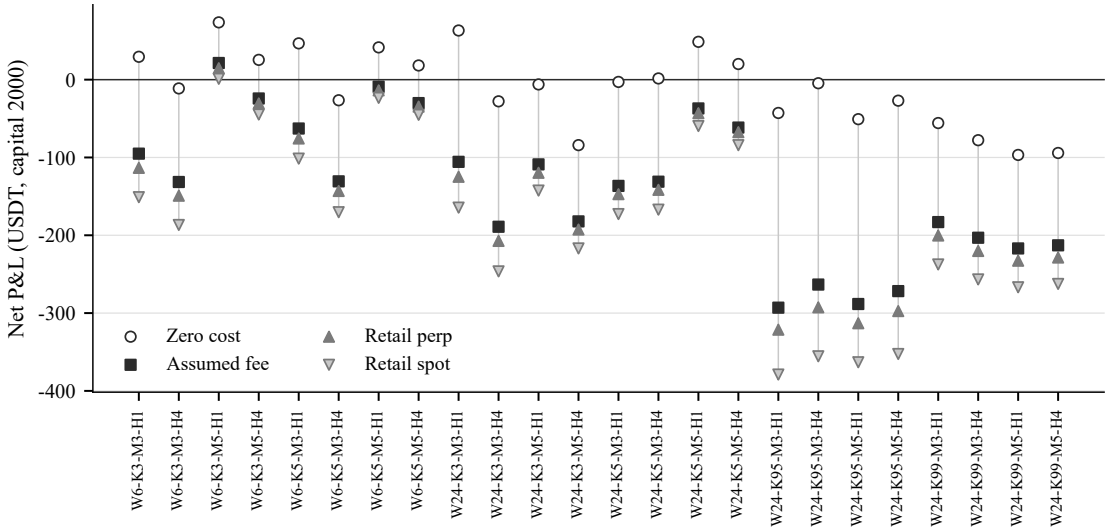



Figure 3: Net P&L across the 24 follow configurations under four cost tiers. Open circles (Zero cost, G) scatter around zero; imposing any realistic transaction friction—Baseline research (R, solid squares), Retail perpetual (U1, dark upward triangles), or Retail spot (U2, light downward triangles)—shifts virtually the entire distribution below zero.

On a per-trade basis (≈ 200 USDT notional per trade, `results/p1_per_trade.csv`), the median gross P&L across the 24 configurations is −0.04 USDT (range −0.76 to +1.10 USDT), whereas median round-trip execution friction (fees plus spread/impact) is 0.87 USDT (range 0.68 to 1.48 USDT), pulling median net P&L to −0.98 USDT per trade.

Crucially, Table A.8 reveals that **bid–ask spread and market impact (45.0 to 233.8 USDT per configuration) dwarf explicit exchange fees (4.5 to 29.3 USDT), accounting for 85% to 92% of total execution costs**. While median round-trip cost is 12.4 to 22.3 bp, mean round-trip cost reaches 34 to 74 bp due to right-skewed illiquidity in smaller altcoins (funding payments are minor, netting −0.3 to −5.1 USDT in received funding). Because social call bursts cluster heavily in thinner altcoins rather than BTC or ETH, market-order followers pay a steep liquidity toll to enter a zero-drift trade.

### *6.2. Deterioration under retail fee schedules*

Replacing our 50%-rebated research fee (2.5 bp/side) with standard unrebated retail taker fees (Table 5, rows U1 and U2) leaves 23 of 24 configurations in the red while deepening

median losses from −131.3 USDT (R) to −145.1 USDT in retail perpetuals (U1) and −171.6 USDT in retail spot (U2, maximum loss −379.0 USDT). The sole positive variant under R (`S1a-W6-K3-M5-H1`, +21.4 USDT) shrinks to +14.8 USDT under U1 and just +1.0 USDT under U2.

## 7. Finding 4: Why Fading the Calls Does Not Work Either

Given that called tokens have already surged +6% to +16% before entry and exhibit negative 72-hour discovery drift in Table 4, a natural contrarian hypothesis arises: **can traders profit by shorting trade-call bursts and holding for several days?**

We tested this exact hypothesis in family **`S1b` (fade call bursts for 24 or 72 hours)**. Its trajectory from discovery triumph to validation failure provides a textbook illustration of in-sample illusion versus out-of-sample falsification (Figure 4).

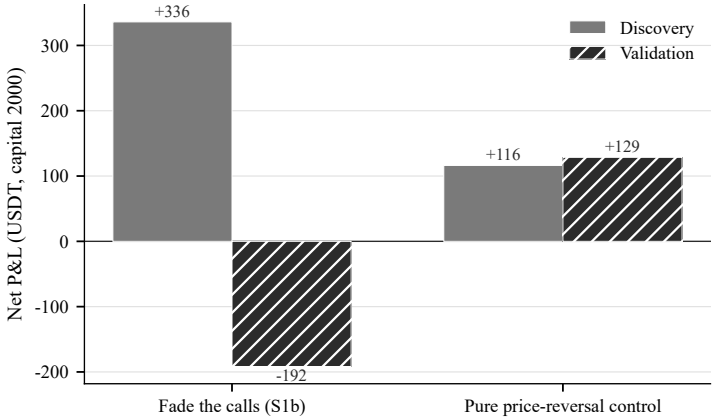



Figure 4: Net P&L (USDT on 2,000 capital) for the contrarian fade strategy (`S1b-W24-K5-M3-H72`) versus its matched pure price-reversal control account (`D1`) in discovery (8 weeks) and locked validation (4 weeks). Fading the calls collapses out of sample, whereas shorting matched prior run-ups without calls earns modest gains in both periods.

### *7.1. In-sample illusion versus locked validation reversal*

In the 8-week discovery period, the selected candidate `S1b-W24-K5-M3-H72` (fading bursts of $\geq 3$ calls at a 5× share multiple over 24 hours, held for 72 hours) appeared remarkably strong: across 70 executed trades (61 shorts, 9 longs; 26 event dates), it generated +336.1 USDT in net profit (+16.8% return on capital), +346.6 bp mean net excess return, a date-clustered $t$-statistic of +2.56, and an annualized Sharpe ratio of 3.12.

Had our study stopped at in-sample backtesting, `S1b` would have been hailed as compelling evidence that Binance Square is a profitable contrarian indicator. Even within discovery, however, concentration diagnostics (`D3`) raised a warning flag: a single volatile token (`SOON`) contributed +154.4 USDT (45.9% of total net profit), and the top three tokens (`SOON`, `TRUMP`, `GPS`) accounted for 87.8%; removing those three reduced net profit across the remaining 67 trades to +41.0 USDT.

Once rules were frozen and the 4-week locked validation set (B2, B5, B8, B11) was opened, the strategy failed outright (Table 6): across 52 validation trades, `S1b` lost −191.7 USDT (−9.6% of capital), with mean net excess flipping to −353.5 bp ($t = -0.83$). Losses occurred both before the mid-September feed break (B2 + B5: 26 trades, −113.7 USDT) and after it (B8 + B11: 26 trades, −78.0 USDT), as well as on Gate perpetuals (−398.1 USDT) and among pre-break established authors (−309.9 USDT).

Table 6: Contrarian fade strategy (`S1b-W24-K5-M3-H72`) in discovery vs. locked validation and matched price-reversal control (`D1`).

| Metric / specification | Disc. (8w) | Val. (4w) | Diagnostic verdict |
|---|---|---|---|
| `S1b` main trades (short / long) | 70 (61/9) | 52 (46/6) | Mostly shorting bullish call bursts |
| `S1b` main net P&L (USDT) | +336.1 | −191.7 | Fails validation net-profit gate |
| `S1b` mean net excess (bp, $t$) | +346.6 (+2.56) | −353.5 (−0.83) | Sign reversal out of sample |
| `S1b` net P&L ex top 3 coins | +41.0 | Negative | Discovery driven by `SOON` (45.9%) |
| Val. split: pre- / post-break | — | −113.7 / −78.0 | Negative in both regimes (26 trades each) |
| Val. Gate / established authors | — | −398.1 / −309.9 | Fails across venue and fixed author set |
| `D1` **control**: pure reversal P&L | +116.2 | +128.6 | Shorting quiet run-ups profits in both |
| `D1` **paired diff.** (called − ctrl, bp) | +176.7 (+0.66) | −503.0 (−2.46) | Called coins revert *less* in validation |

Notes: Capital 2,000 USDT; date-clustered $t$-statistics in parentheses. At each `S1b` trigger, the `D1` control selects an uncalled tradable token with the closest trailing 24-hour run-up and executes an identical 72-hour short (engineering correction details in Appendix D).

### *7.2. Mechanism attribution: the matched price-reversal control (`D1`)*

To separate mechanical mean reversion after large price spikes from any incremental reversal caused by social calls, we preregistered diagnostic **`D1` (Matched Price-Reversal Control)**: whenever a token triggers an `S1b` entry, we simultaneously match it to a "quiet" token from the same tradable universe that experienced a nearly identical prior 24-hour return but had *no* trade-call burst on Binance Square, shorting the control token under identical account rules.

The bottom rows of Table 6 and the right bars of Figure 4 expose the underlying mechanism:

1. **Pure price reversal earns modest gains in both periods**: The control account—which ignores Binance Square entirely and simply shorts tokens after sharp 24-hour run-ups—earns +116.2 USDT in discovery and +128.6 USDT in validation.
2. **Social calls add zero contrarian value in discovery and significantly hurt short sellers in validation**: The paired net excess difference between called tokens and matched quiet tokens is statistically indistinguishable from zero in discovery (+176.7 bp, $t = +0.66$, driven by `SOON`) and turns **significantly negative in validation (−503.0 bp, $t = -2.46$)**. Conditional on the same prior price run-up, tokens experiencing a call burst on Binance Square subsequently pull back *less* than quiet tokens, costing short sellers roughly 5 percentage points per trade relative to a pure momentum-reversal baseline.

Thus, the apparent in-sample profitability of fading Square calls was an artifact of mechanical return reversal following extreme price spikes, amplified by single-coin luck (`SOON`). **Binance Square is neither a valid momentum signal nor a reliable contrarian indicator.**

## 8. Discussion: Words vs. Positions, Implications, and Limitations

### *8.1. Why do trade calls lack predictive power? The disconnect between posters and traders*

Why do intense bursts of bullish calls neither propel prices higher over the next 1–4 hours nor mark crowded retail buying that subsequently unwinds beyond normal price mean reversion?

Our price-free measurement study (Appendix C) offers a direct microstructural explanation. For each token in the top-150 universe, we compute rolling 7-day net bullish stance on Binance Square (net7 $= (N_{\text{bull}} - N_{\text{bear}})/N_{\text{total}}$, requiring $\geq 5$ single-coin call/analysis/news posts) and compare it against actual **account long–short ratios** on OKX and Gate perpetuals across 65 days (July 25 to September 27, 2026; Table 7).

Table 7: Cross-sectional and within-coin correlations between Binance Square 7-day stance and exchange account long–short ratios (Jul 25 – Sep 27, 2026).

| Pairwise relationship | Full (65d) | Pre-break | Post-break |
|---|---|---|---|
| Square 7d net bullish (`net7`) vs. OKX account LSR (XS mean) | −0.08 | −0.12 | +0.06 |
| Square 7d net bullish (`net7`) vs. Gate account LSR (XS mean) | −0.11 | −0.16 | +0.07 |
| Square 7d attention (`att7`) vs. combined exchange LSR rank | −0.03 | −0.04 | 0.00 |
| Square 7d net bullish (`net7`) vs. OKX account LSR (within-coin) | −0.19 | — | — |
| **Benchmark**: OKX account LSR vs. Gate account LSR (XS mean) | **+0.56** | — | — |

Notes: Cross-sectional (XS) values are daily Spearman rank correlations averaged over time (≈ 60 eligible tokens/day). Within-coin correlation pools 3,475 token-day observations across 76 tokens. No price or return data were accessed.

While account long–short ratios on OKX and Gate correlate strongly with each other in the cross section (+0.56), confirming that exchange positioning metrics capture a coherent retail footprint, **Binance Square net bullishness is virtually orthogonal to exchange account positioning**: cross-sectional Spearman correlations average just −0.08 (OKX) and −0.11 (Gate), flipping sign across the mid-September feed break. Within individual tokens over time, Square bullishness and OKX long–short ratios correlate negatively at −0.19.

This divergence indicates that **vocal content creators on Binance Square and leveraged traders putting capital at risk on perpetual exchanges are largely distinct populations responding to price with opposite incentives**. Following a +10% price jump, perpetual account holders tend to take profit on longs or open contrarian shorts (lowering the account long–short ratio), whereas Square creators—motivated by platform creator rewards, referral commissions, and follower growth—publish bullish posts chasing the green candle. Treating engagement-seeking social commentary as a proxy for marginal buying power or crowded positioning is the root error behind both follow and fade strategies.

### *8.2. Practical and methodological implications*

1. **For market participants**: (i) *Check the prior 24-hour chart*—when a coin floods the recommendation feed, +6% to +16% of excess appreciation has typically already occurred; (ii) *Respect microstructure drag in zero-edge trades*—in small-cap perpetuals, spread and impact costs (34–74 bp mean round trip) turn flat gross returns into steady account bleed; (iii) *Do not treat Square as a contrarian short signal*—shorting surging altcoins carries severe squeeze risk, and called coins actually revert less in validation than quiet coins with identical run-ups.
2. **For quantitative researchers**: Aligning social signals at authoring time (`posted_at`) rather than feed-visibility time (`first_seen_at`) injects severe look-ahead bias; omitting locked validation splits, top-coin concentration gates, and matched-momentum controls (`D1`) allows single-outlier artifacts like `SOON` ($t = 2.56$ in discovery) to masquerade as alpha.

### *8.3. Limitations*

Six boundary conditions qualify our conclusions:

1. **Sample duration**: The archive spans 12 weeks (84 days) within a single market regime. Although the interleaved split balances the mid-September feed break, adjacent discovery and validation weeks are less independent than multi-year out-of-sample tests.
2. **Recommendation feed vs. full firehose**: Results apply to the web recommendation feed (`feed-recommend/list`), not private following timelines or unrecommended posts.
3. **LLM labels and knowledge cutoffs**: Annotations rely on LLMs rather than human gold labels. While primary model `GPT-6.1 Sol` has a verified April 30, 2026 cutoff prior to our sample, `GLM-5.3 Flash` lacks an officially disclosed cutoff date.

4. **Venue and universe coverage**: Simulations cover the 408 tokens mapped to OKX USDT perpetuals within the top-150 turnover universe (78.8% of token posts), excluding on-chain-only micro-caps.
5. **Unobserved KOL effect**: Because point-in-time follower counts at first observation were null in the feed API and backfilling current follower counts would induce severe look-ahead bias, hypothesis family `S6` (large-author leadership) was frozen as "insufficient data." Our findings address aggregate feed bursts, not whether tracking specific key opinion leaders is profitable.
6. **Shared event paths across follow and fade**: Families `S1a` and `S2b` did not run as separate validation accounts after failing discovery, but because `S1a` and validated `S1b` share the exact same underlying call-burst events, the validation price path is directly observed in the dashed curve of Figure 1.

## 9. Conclusion

Archiving 181,509 recommended posts on Binance Square over 12 weeks and logging 936 preregistered trial runs, we find that trade-call bursts are a lagging echo of +5.9% to +15.8% prior 24-hour price rallies. Buying immediately after a burst yields zero gross drift and consistent net losses (23/24 configurations negative) due to altcoin spread and impact costs; fading the burst fails in locked validation (−191.7 USDT, $t = -0.83$) and underperforms a pure price-reversal control by 503.0 bp ($t = -2.46$). Paired with near-zero correlation between Square sentiment and exchange account long–short ratios, the evidence shows that neither chasing nor fading recommendation-feed call bursts offers a positive trading edge.

## Appendix A. Supplementary Summary Tables

Tables A.8 and A.9 report the configuration-level cost decomposition and four-tier fee sensitivity from supplementary experiments `P1` and `P2`, generated directly by `code/make_tables.py`.

Table A.8: Where the money goes: gross P&L, fees, spread and impact, and funding across 24 follow configurations (discovery, USDT on 2,000 capital).

| Config | Gross | Fees | Spread+impact | Funding paid | Net | Median round trip (bp) |
|---|---|---|---|---|---|---|
| S1a-W6-K3-M3-H1 | +29.4 | 18.2 | 107.6 | −1.3 | −95.1 | 13.7 |
| S1a-W6-K3-M3-H4 | −11.3 | 17.6 | 105.1 | −2.5 | −131.5 | 13.7 |
| S1a-W6-K3-M5-H1 | +73.6 | 6.6 | 46.1 | −0.5 | +21.4 | 12.4 |
| S1a-W6-K3-M5-H4 | +25.4 | 6.3 | 45.0 | −1.7 | −24.2 | 12.4 |
| S1a-W6-K5-M3-H1 | +46.6 | 12.8 | 97.0 | −0.3 | −62.8 | 15.6 |
| S1a-W6-K5-M3-H4 | −26.5 | 12.4 | 94.5 | −2.6 | −130.7 | 15.5 |
| S1a-W6-K5-M5-H1 | +41.5 | 4.7 | 46.0 | −0.3 | −9.0 | 16.7 |
| S1a-W6-K5-M5-H4 | +18.2 | 4.5 | 45.5 | −1.6 | −30.1 | 16.7 |
| S1a-W24-K3-M3-H1 | +63.2 | 19.2 | 150.9 | −1.3 | −105.6 | 19.5 |
| S1a-W24-K3-M3-H4 | −27.9 | 18.2 | 145.9 | −2.8 | −189.0 | 19.3 |
| S1a-W24-K3-M5-H1 | −6.1 | 10.8 | 93.3 | −1.4 | −108.9 | 17.3 |
| S1a-W24-K3-M5-H4 | −84.1 | 10.5 | 90.9 | −3.4 | −182.2 | 17.4 |
| S1a-W24-K5-M3-H1 | −2.9 | 10.6 | 127.5 | −4.5 | −136.5 | 22.2 |
| S1a-W24-K5-M3-H4 | +1.6 | 10.4 | 127.0 | −4.7 | −131.1 | 22.2 |
| S1a-W24-K5-M5-H1 | +48.6 | 6.0 | 84.1 | −4.6 | −37.0 | 22.3 |
| S1a-W24-K5-M5-H4 | +20.1 | 5.8 | 81.0 | −5.1 | −61.6 | 22.1 |
| S2b-W24-K95-M3-H1 | −42.8 | 28.3 | 223.2 | −1.2 | −293.0 | 18.8 |
| S2b-W24-K95-M3-H4 | −4.6 | 29.3 | 233.8 | −4.4 | −263.3 | 19.0 |
| S2b-W24-K95-M5-H1 | −50.8 | 24.7 | 213.5 | −0.7 | −288.3 | 18.3 |
| S2b-W24-K95-M5-H4 | −27.0 | 25.6 | 223.0 | −3.9 | −271.8 | 18.3 |
| S2b-W24-K99-M3-H1 | −55.8 | 17.2 | 112.8 | −2.8 | −183.0 | 19.6 |
| S2b-W24-K99-M3-H4 | −77.8 | 17.0 | 111.0 | −2.6 | −203.2 | 19.5 |
| S2b-W24-K99-M5-H1 | −96.8 | 15.7 | 107.2 | −2.7 | −216.9 | 19.4 |
| S2b-W24-K99-M5-H4 | −94.2 | 15.7 | 105.6 | −2.6 | −212.9 | 19.4 |

Notes: Source: `results/p1_follower_table.csv`. Net = Gross − Fees − Spread+impact − Funding paid. Negative funding paid indicates net funding received. Median round trip = (Fees + Spread+impact)/one-way notional; mean round-trip costs range from 34 to 74 bp due to illiquid tail tickers.

Table A.9: Account net P&L across 24 follow configurations under four cost tiers (discovery, USDT on 2,000 capital).

| Config | Zero cost | Assumed fee | Retail perp | Retail spot |
|---|---|---|---|---|
| S1a-W6-K3-M3-H1 | +29.4 | −95.1 | −113.3 | −151.1 |
| S1a-W6-K3-M3-H4 | −11.3 | −131.5 | −149.1 | −186.8 |
| S1a-W6-K3-M5-H1 | +73.6 | +21.4 | +14.8 | +1.0 |
| S1a-W6-K3-M5-H4 | +25.4 | −24.2 | −30.5 | −44.8 |
| S1a-W6-K5-M3-H1 | +46.6 | −62.8 | −75.6 | −101.5 |
| S1a-W6-K5-M3-H4 | −26.5 | −130.7 | −143.0 | −170.4 |
| S1a-W6-K5-M5-H1 | +41.5 | −9.0 | −13.7 | −23.5 |
| S1a-W6-K5-M5-H4 | +18.2 | −30.1 | −34.6 | −45.3 |
| S1a-W24-K3-M3-H1 | +63.2 | −105.6 | −124.8 | −164.5 |
| S1a-W24-K3-M3-H4 | −27.9 | −189.0 | −207.2 | −246.4 |
| S1a-W24-K3-M5-H1 | −6.1 | −108.9 | −119.7 | −142.6 |
| S1a-W24-K3-M5-H4 | −84.1 | −182.2 | −192.6 | −216.9 |
| S1a-W24-K5-M3-H1 | −2.9 | −136.5 | −147.1 | −172.8 |
| S1a-W24-K5-M3-H4 | +1.6 | −131.1 | −141.6 | −167.1 |
| S1a-W24-K5-M5-H1 | +48.6 | −37.0 | −43.0 | −59.7 |
| S1a-W24-K5-M5-H4 | +20.1 | −61.6 | −67.4 | −84.3 |
| S2b-W24-K95-M3-H1 | −42.8 | −293.0 | −321.3 | −379.0 |
| S2b-W24-K95-M3-H4 | −4.6 | −263.3 | −292.6 | −355.5 |
| S2b-W24-K95-M5-H1 | −50.8 | −288.3 | −313.1 | −363.2 |
| S2b-W24-K95-M5-H4 | −27.0 | −271.8 | −297.4 | −352.6 |
| S2b-W24-K99-M3-H1 | −55.8 | −183.0 | −200.3 | −237.5 |
| S2b-W24-K99-M3-H4 | −77.8 | −203.2 | −220.2 | −256.7 |
| S2b-W24-K99-M5-H1 | −96.8 | −216.9 | −232.6 | −266.9 |
| S2b-W24-K99-M5-H4 | −94.2 | −212.9 | −228.6 | −262.6 |
| Negative / positive | 14 / 10 | 23 / 1 | 23 / 1 | 23 / 1 |

Notes: Source: `results/p2_net_pnl_by_tier.csv`. Zero cost (G): no friction; Assumed fee (R, baseline research): 2.5 bp/side fee after 50% rebate plus spread/impact and funding; Retail perp (U1): 5.0 bp/side; Retail spot (U2): 10.0 bp/side and zero funding, reusing perpetual spread/impact estimates. Trade quantities are held fixed across tiers without equity resizing.

## Appendix B. Protocol and Econometric Conventions

**Calendar and interleaved split.** Twelve 7-day blocks run from 00:00 UTC+8 on July 24, 2026 to 24:00 UTC+8 on October 15, 2026 (B0–B11; October 16 prices settle open positions from B11). Discovery comprises B0, B1, B3, B4, B6, B7, B9, B10 (56 days); locked validation comprises B2 (Aug 07–13), B5 (Aug 28–Sep 03), B8 (Sep 18–24), and B11 (Oct 09–15) (28 days). Trailing feature windows (`S1`: 14d; `S2`: 28d) may look back across block boundaries, but any discovery trade whose exit would fall inside a validation block is dropped.

**Deduplication and execution.** Posts are deduplicated by `(author_id, symbol)` over rolling 6-hour windows. Only single-token posts contribute to directional stance. Repeat event triggers for the same token within 24 hours are merged. Orders execute at the open of the first complete hourly bar at least 5 minutes after `first_seen_at`, rounded down to OKX lot sizes (`lotSz`, `minSz`, `ctVal`). Standard errors and $t$-statistics are clustered by event date (UTC+8).

## Appendix C. Price-Free Measurement: Feed Break and Positioning Alignment

The measurement study accessed only post records, collector logs, and exchange long–short ratios. Single-token call, analysis, and news posts (58% of token posts) were aggregated into rolling 7-day net bullishness $\text{net7} = (N_{\text{bull}} - N_{\text{bear}})/N_{\text{total}}$ (requiring $N_{\text{total}} \geq 5$) and attention $\text{att7} = \ln(1 + N_{\text{total}})$. Daily cross-sectional Spearman correlations were computed against log account long–short ratios on OKX (`account_long_short_ratio`) and Gate (`lsr_account`) across the top-150 universe from July 25 to September 27, 2026 ($\approx 60$ tokens/day). Restricting posts to pre-break established authors changes correlations by at most 0.04.

## Appendix D. Disclosure of Engineering Corrections

We disclose two mechanical corrections that involved no parameter selection:

1. **Validation funding-window right-endpoint correction**: After inspecting the initial `S1b` validation run (49 trades, net −191.38 USDT, $t = -0.87$), a code audit found an exclusive right-boundary check in the funding export that prematurely truncated terminal funding records for three legitimate 72-hour exits (`APT`, `LTC`, `BCH`) and one rate on `ROBO`. With user authorization, we fixed the boundary check while holding all protocol rules and parameters fixed, yielding 52 trades, net −191.67 USDT ($t = -0.83$, difference −0.295 USDT). Both batches are preserved in the archive; because the fix followed initial inspection, we report it as an engineering-corrected confirmation rather than a blind re-test.
2. **Funding sign convention in `PREREGISTRATION_P1`**: Section 5 of `PREREGISTRATION_P1` wrote the per-trade identity as `gross - fees - price_cost + funding`. Because `funding` in `events.json` records funding *paid* (positive for outflow, negative for inflow), the correct accounting identity subtracts `funding`. The automated check in `code/p1_follower_table.py` caught this on the first run; once corrected, all 24 configurations matched `summary.json` to four decimal places ($< 0.01$ USDT).

## Data and Code Availability

A self-contained replication package (`square_callecho_attachments.zip`) accompanies this paper, allowing full reproduction of all tables, figures, and headline statistics without re-scraping or re-running LLMs:

- **Frozen datasets (`data/`)**: Trade-level logs for all 24 follow configurations (`data/trades/*.csv`), event-time price curves (`data/curves/*.csv`), validation curves, `S1b` and `D1` control summaries, and the 936-run trial counts.
- **Strategy and replication scripts (`code/`)**: Original event construction (`strategy/events.py`) and account simulation (`strategy/account.py`) modules, plus a standalone script `reproduce.py` verified by `SHA256SUMS.json`.
- **Public endpoints**: Binance Square recommendations were queried from `/bapi/composite/v9/friendly/pgc/feed/feed-recommend/list` (with detail endpoint `/bapi/composite/v3/friendly/pgc/special/content/detail/` and sitemap, unauthenticated); OKX perpetual candles, funding rates, and contract specs were retrieved from `/api/v5/market/history-candles`, `/api/v5/public/funding-rate-history`, and `/api/v5/public/instruments`.

To protect user privacy and comply with platform terms, post texts, author identifiers, post IDs, and post-level labels are excluded from the archive.

## References


[1] Barber, B.M., Odean, T., 2008. All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors. *Review of Financial Studies* 21, 785–818.

[2] Da, Z., Engelberg, J., Gao, P., 2011. In search of attention. *Journal of Finance* 66, 1461–1499.

[3] Antweiler, W., Frank, M.Z., 2004. Is all that talk just noise? The information content of internet stock message boards. *Journal of Finance* 59, 1259–1294.

[4] Cookson, J.A., Niessner, M., 2020. Why don't we agree? Evidence from a social network of investors. *Journal of Finance* 75, 173–228.

[5] Harvey, C.R., Liu, Y., Zhu, H., 2016. …and the cross-section of expected returns. *Review of Financial Studies* 29, 5–68.

[6] Bailey, D.H., López de Prado, M., 2014. The deflated Sharpe ratio: Correcting for selection bias, backtest overfitting, and non-normality. *Journal of Portfolio Management* 40, 94–107.