

INFORMATIONAL INERTIA IN A DECENTRALIZED PREDICTION MARKET: EVIDENCE FROM THE MAY 2024 CPI LEAK*

EMREHAN AKTUĞ[†]

Sabancı University

ORHAN TORUL[‡]

Boğaziçi University

June 2, 2026

Abstract

On May 15, 2024, the U.S. Bureau of Labor Statistics inadvertently uploaded the CPI report to its public server at 08:00 ET, thirty minutes before the scheduled release. We use this verified timing shock as a single-episode descriptive case study of how quickly a decentralized retail prediction market incorporates a high-signal macroeconomic surprise. Polymarket's Fed rate-cut contracts exhibit a near-zero top-of-book response throughout the 35-minute interval between the inadvertent upload and the first detectable price adjustment at 08:35 ET, while CME 30-Day Federal Funds futures begin adjusting within seconds of the upload and complete the bulk of their move before the official release. The 35-minute interval has two parts. The pre-release thirty minutes are consistent with rational non-adjustment inside a 1–2 percentage-point no-arbitrage band given the modest realized shock. The remaining five minutes, by contrast, fall after the official release: by 08:30 ET the information is universally public and CME has completed the bulk of its move, so the bid–ask logic that rationalizes the pre-release window cannot account for the continued absence of repricing. Change-point detection, randomization inference, synthetic control, and quote-update hazard analysis, reported as bounding exercises rather than sharp hypothesis tests, are jointly consistent with limited updating rather than noisy immediate repricing. The episode highlights a sizable monitoring-infrastructure difference between a thin decentralized retail venue and the corresponding institutional benchmark, and provides a pre-growth baseline for future measurement of real-time efficiency on Polymarket.

Keywords: Prediction markets, Price discovery, Informational efficiency, Macroeconomic news, Market microstructure, Event study, Polymarket

JEL Codes: G14, E44, D82

*We thank two anonymous reviewers and our colleagues at Boğaziçi and Sabancı University for helpful comments and suggestions. Torul gratefully acknowledges financial assistance by the Faculty of Economics and Administrative Sciences at Boğaziçi University for the submission of this manuscript. All errors are our own.

[†]Address: Sabancı University Department of Economics, 34956, Tuzla, Istanbul, Türkiye.

E-Mail: emrehan.aktug@sabanciuniv.edu

[‡]Corresponding author. Address: Boğaziçi University, Department of Economics, 34342, Bebek, Istanbul, Türkiye.

E-Mail: orhan.torul@bogazici.edu.tr

1 Introduction

How quickly does a decentralized prediction market incorporate major macroeconomic news? The answer matters because prediction markets are increasingly used as real-time forecasting devices, and their credibility depends in part on how rapidly they absorb high-signal public information. In institutional bond and foreign-exchange markets, scheduled macroeconomic announcements are incorporated within seconds.¹ Comparable evidence for retail-accessible prediction markets, however, remains limited, especially at minute frequency and in settings with a verified information-arrival timestamp.² We exploit a rare natural experiment—the inadvertent early release of U.S. CPI data on May 15, 2024—to provide such evidence.

On May 15, 2024, the Bureau of Labor Statistics (BLS) inadvertently uploaded CPI data to its public server at 08:00 ET (Eastern Time), thirty minutes before the scheduled release. The report indicated cooling inflation, a surprise favorable to rate-cut expectations. We compare the minute-level response of Polymarket’s “How many Fed rate cuts in 2024?” contracts with the response of CME (Chicago Mercantile Exchange) 30-Day Federal Funds futures, our primary institutional benchmark, and supplement the comparison with Bitcoin and Ethereum spot prices on Binance as contextual evidence on what continuously monitored, algorithmically traded markets did at the same moment. The BLS-confirmed upload time provides a precisely dated $T = 0$, offering an unusually clean lower bound on the start of the information-asymmetry window (the window cannot have opened earlier than the verified upload). Unlike a scheduled release, however, the early upload did not generate common-knowledge information arrival: the timestamp identifies when the information became publicly accessible, not when market participants actually observed it.

Our contribution is to describe the discovery gap that this single verified episode reveals. Two canonical frictions make the measurement interesting *ex ante* rather than mechanical. First, information that is publicly *accessible* need not be publicly *attended to*: in a [Grossman and Stiglitz \(1980\)](#) equilibrium, costly information acquisition sustains a non-zero delay between availability and incorporation whenever the expected private gain is small relative to the cost of monitoring, and rational-inattention models ([Sims, 2003](#)) imply that agents optimally ignore signals whose realized magnitudes lie below an endogenous attention threshold. Second, a decentralized retail venue without designated market makers, programmatic feeds, or low-latency APIs lies far out on the cost-of-monitoring dimension, so the predicted incorporation lag for an unannounced early upload is longer than for a scheduled release. What the episode shows is that the lag has two conceptually distinct components: a 30-minute pre-release non-response over the asymmetric-information window (08:00–08:30 ET) and a further 5-minute post-release latency (08:30–08:35 ET) before quoted probabilities adjust in response to the now-universally-public release. The pre-release component is consistent with rational non-adjustment inside a 1–2 percentage-point no-arbitrage band given the modest realized shock. The post-release residual is the substantive puzzle, since once CME has already moved on the same information the bid–ask logic that explains the pre-release window no longer applies. *Informational inertia*, in our usage, refers to this post-release residual. The episode provides a pre-growth baseline for real-time informational efficiency in a thin decentralized retail prediction market.

¹See [Andersen and Bollerslev \(1998\)](#); [Balduzzi et al. \(2001\)](#); [Evans and Lyons \(2002\)](#); [Fleming and Remolona \(1999\)](#).

²The analogous speed in retail-accessible prediction markets ([Arrow et al., 2008](#); [Berg et al., 2008](#)), where contracts isolate single macro outcomes ([Manski, 2006](#); [Wolfers and Zitzewitz, 2004](#)), has not been measured at minute frequency against a verified timing shock.

2 Institutional Setting and Data

The report released inadvertently at 08:00 ET indicated cooling inflation (0.3% MoM vs. 0.4% expected), a signal that typically raises the probability of Federal Reserve rate cuts ([Bureau of Labor Statistics, 2024](#)). Our primary outcome is the minute-level price path of Polymarket’s “How many Fed rate cuts in 2024?” contract series, traded on its Central Limit Order Book (CLOB). The series consists of mutually exclusive outcome tokens (“0 cuts,” “1 cut,” ..., “6+ cuts”) that settle to \$1 if the outcome occurs and \$0 otherwise. The CLOB has no designated market maker; aggregate cumulative lifetime trading volume across tokens was of order \$1M–\$20M, and event-window depth on May 15 was thin.

For each outcome token k associated with C_k cuts, we extract its mid-quote implied probability p_k and compute the expected number of cuts $E = \sum_k p_k C_k$. We then apply the generalized logit $\ell \equiv \ln(E/(6 - E))$, which we refer to as “log-odds” throughout. This transformation maps the bounded interval $(0, 6)$ to the real line, mitigating boundary compression in arithmetic averages; in our sample $E \in (0.8, 1.3)$, so no truncation is invoked. We use raw mid-quote p_k rather than probabilities normalized by $\sum_j p_j$; because the deviation of $\sum p_k$ from 1 is small and stable across windows (0.982 on May 15, 0.980 on placebos; Table 2), the normalization would alter E by at most 2% and the *abnormal* log-odds by a negligible amount. As a technical caveat, the unnormalized probabilities imply a structural ceiling on E slightly below 6, so the log-odds transform does not reach the upper pole; this does not bind in our sample, where $E \in (0.8, 1.3)$. The “6+ cuts” token is right-censored (we set $C_k = 6$), which imparts a slight downward bias on E whenever that token carries positive mass and therefore dampens rather than amplifies any post-leak upward move; the upward repricing we observe is, in that sense, a conservative *lower bound* on the magnitude the cooling surprise warranted.

Our primary benchmark for the institutional response is the CME 30-Day Federal Funds future (ZQU4), the standard market-based instrument for U.S. rate-cut expectations and the venue most directly comparable in payoff structure to the Polymarket Fed contracts. To keep the institutional-versus-retail, centralized-versus-decentralized, and algorithmic-versus-manual-monitoring distinctions separate, we map each market we use along all three (Table 1). The contrast that drives our reading of the May 15 episode is the monitoring-apparatus dimension, not decentralization per se: CME and Binance both feature programmatic participation by news-aware systems with low-latency infrastructure, while Polymarket in May 2024 lacked designated market makers, algorithmic news feeds, and low-latency APIs.

Table 1: Three-Dimension Taxonomy of Venues Considered

Venue	Participant base	Matching infrastructure	Monitoring apparatus
CME ZQU4	Institutional	Centralized	Automated, algorithmic
Binance BTC/ETH	Mixed (retail + automated)	Centralized	Automated, algorithmic
Polymarket Fed cuts	Retail	Decentralized (Polygon CLOB)	Manual / event-driven

Notes: The three dimensions are conceptually distinct. CME is the primary institutional benchmark; Binance BTC/ETH is reported as continuous-monitoring contextual evidence; Polymarket is the venue under study. CME ZQU4 denotes the September 2024 CME 30-Day Federal Funds futures contract, the standard market-based instrument for U.S. rate-cut expectations; BTC/ETH are Bitcoin and Ethereum spot pairs on Binance. The contrast highlighted by the May 15 episode is between automated/algorithmic and manual/event-driven monitoring apparatuses rather than between centralized and decentralized matching per se.

Bitcoin (BTC/USDT) and Ethereum (ETH/USDT) prices from Binance are included as contextual evidence on what continuously monitored markets did during the same window, not as a benchmark for direct CPI-to-rate-cut transmission. Binance is a centralized exchange with conventional matching infrastructure; its role here is to provide a liquid, round-the-clock, algorithmically traded reference point on

the monitoring-infrastructure dimension, not a stand-in for decentralization. Both assets have emerged as macro-sensitive proxies through the dollar and real-yield channels, but the transmission is indirect and any leak-window co-movement is observational rather than evidence of price discovery. We report descriptive return statistics in Section 3.1 for completeness and read the resulting comparisons as a continuous-monitoring reference rather than as an evidentiary benchmark.³

We restrict the placebo group to the four CPI release days in 2024 immediately following May 15 (June 12, July 11, August 14, and September 11) to ensure that Polymarket contract structure, participant base, and market depth are comparable across event and placebo windows; this restriction limits statistical power, as discussed in Section 4.

Table 2: Average Abnormal Log-Odds and Coherence across Windows

Event Class	Early Window [0, 30m]	Main Window [0, 60m]	Avg. Sum(P) [0, 30m]
Leak Event (May 15, 2024)	0.0053	0.0354	0.982
Normal Placebo Days	0.0007	-0.0024	0.980

Notes: Mean abnormal log-odds relative to a pre-leak baseline computed as the mean log-odds over the 10-minute window 07:45–07:55 ET (centered on $T = -10$ minutes relative to the 08:00 ET leak). The early window corresponds to 08:00–08:30 ET. The placebo group includes four normal CPI release days in 2024 (June 12, July 11, August 14, September 11); its Avg. Sum(P) column is computed over the equivalent 08:00–08:30 ET window on those placebo days. A positive abnormal log-odds value implies an upward shift in the expected number of rate cuts, consistent with the cooling inflation surprise observed on May 15. For scale, using $dE/d\ell = E(6-E)/6 \approx 0.86$ at the pre-leak $E \approx 1.04$, a leak-window log-odds shift of 0.0053 translates to a change in the expected number of rate cuts of approximately 0.005 cuts; the full-window shift of 0.0354 translates to approximately 0.030 cuts—small in absolute terms, consistent with the minor magnitude of the inflation surprise (0.1 percentage point vs. consensus). The early-window figure (0.0053) is the mean abnormal log-odds *level* averaged across all minutes of the 08:00–08:30 ET window (explicitly, 08:00–08:30 ET for the early window and 08:00–09:00 ET for the main window, both relative to the 08:00 ET leak); this same statistic is used in the randomization inference test in Appendix A.

3 The 35-Minute Interval: Pre-Release Window and Post-Release Latency

The 35-minute interval between the inadvertent 08:00 ET upload and the first detectable Polymarket price adjustment at 08:35 ET is best read as the sum of two conceptually distinct components: a 30-minute pre-release non-response over the 08:00–08:30 ET asymmetric-information window, and a 5-minute post-release latency over 08:30–08:35 ET after the universally public official release. We treat the two separately throughout this section, and the formal tests reported in the Appendix as bounding exercises rather than sharp hypothesis tests.

3.1 Pre-release leak window: near-zero response in a no-arbitrage band

Timing. The Bureau of Labor Statistics confirmed in a post-incident statement that CPI files were inadvertently loaded to its public server at 08:00:00 ET (Bureau of Labor Statistics, 2024), providing a clean $T = 0$ and, unlike rumor-based leaks, a verified rather than inferred upload moment.⁴

Polymarket non-response. Between 08:00 and 08:30 ET, Polymarket Fed contracts exhibited negligible abnormal movement. The mean abnormal log-odds level was 0.0053 (Table 2). Randomization inference across matched placebo mornings places this non-response indistinguishably within the routine

³Qualitative inspection of E-mini S&P 500 futures (ES) price data suggests no visible abnormal movement at 08:00 ET on May 15, 2024, consistent with equity algorithmic traders not monitoring the BLS server pre-release; formal return statistics and placebo comparisons for ES are not reported and constitute a direction for further investigation.

⁴The 08:00:00 ET timestamp refers to public accessibility on the BLS server, as confirmed by the post-incident BLS statement. Any earlier internal staging of the file on non-public directories is irrelevant to market participants and does not affect the onset of the information-asymmetry window.

pre-release distribution: the rolling-window design (120 overlapping shifts) yields $p = 0.157$, and the coarser four-block non-overlapping permutation places the May 15 statistic at the median of the four-day reference distribution (see Appendix A) (Fisher, 1935). We treat these as bounding exercises rather than sharp tests, for reasons developed below.

Economic magnitude. The leak-window magnitude is small by design: the CPI surprise itself was 0.1 percentage point against consensus, and the corresponding abnormal log-odds shifts translate via the local slope $dE/d\ell = E(6 - E)/6 \approx 0.86$ at the pre-leak $E \approx 1.04$ to roughly 0.005 expected cuts in the 30-minute pre-release window and 0.030 cuts over the full 08:00–09:00 ET window. The contribution of this paper is therefore not the magnitude of any putative mispricing but the *speed of incorporation conditional on a verified, modest shock*: even a small directionally correct response should have been observable in real time, and the same instrument that registered 0.030 cuts of cumulative repricing by 09:00 registered only 0.005 over the 30-minute window in which CME completed nearly 60% of its move.

Crypto context. Bitcoin and Ethereum were positive over the leak window (BTC: +0.31%; ETH: +0.12%, computed from the 07:59 ET close to the 08:29 ET close using Binance 1-minute OHLCV data), though only Bitcoin’s move exceeds typical pre-release drift. The four matched placebo CPI mornings exhibit positive and somewhat dispersed pre-release drift (BTC: $\bar{r} = +0.23\%$, $s = 0.33\%$; ETH: $\bar{r} = +0.38\%$, $s = 0.30\%$). A one-sample test that accounts for the variance of a single new observation, $t = (r_{\text{event}} - \bar{r}_{\text{placebo}}) / (s_{\text{placebo}} \sqrt{1 + 1/n})$ ($n = 4$, $\text{df} = 3$), yields $t = 0.22$ ($p = 0.84$) for BTC and $t = -0.78$ ($p = 0.49$) for ETH; with only four placebo days, the formal t -tests have negligible power and are reported for completeness rather than as inference, consistent with the contextual—not evidentiary—role assigned to these series in Section 2. What the numbers do support is a *gradient* along the monitoring-apparatus axis (Table 1): seconds at the CME institutional benchmark, noisy minute-level drift in 24/7 algorithmic crypto spot markets, and essentially zero movement on Polymarket through 08:30 ET—consistent with differences in macro-monitoring infrastructure rather than with decentralization per se.

No-arbitrage interpretation. The 30-minute pre-release non-response is consistent with rational non-adjustment inside a standard no-arbitrage band rather than with a behavioral failure. Polymarket operates on Polygon, where a round-trip trade in May 2024 cost approximately \$0.01–0.05 in gas (transaction fees) at prevailing rates; the platform charged no taker fee at the time. The binding friction is the bid–ask spread, which on thin Fed rate-cut tokens was typically 1–2 cents wide. Because these binary tokens settle to \$1 or \$0, one cent of price corresponds mechanically to one percentage point of implied probability, so the spread defines an approximately 1–2 percentage-point no-arbitrage band around the mid-quote, with a half-spread of 0.5–1 percentage point. The leak-window abnormal log-odds of 0.0053 corresponds via the local slope to a ≈ 0.005 -cut shift in E , and the per-token probability shift is at most approximately 0.005 if concentrated on a single marginal token, and smaller if distributed across tokens—in either case comparable to, and most plausibly below, the half-spread. We therefore make no behavioral-failure claim for the 08:00–08:30 ET window; on this venue, with this spread, this pre-release pattern is what the no-arbitrage logic predicts. The residual that the paper calls *informational inertia* is the post-release window, treated separately in Section 3.2.

What the design can detect. A natural question is whether non-response merely reflects insufficient power. Under a location-shift model, the rolling-window randomization design at the 5% level has a minimum detectable effect (MDE) of approximately 0.0106, 0.0122, and 0.0156 log-odds at 50%, 80%, and 90% power respectively—equivalently, roughly 0.009, 0.010, and 0.013 expected cuts using the local slope (see Appendix A for the full derivation). The observed early-window statistic (0.0053 log-odds, ≈ 0.005 cuts) is less than half the 80%-power MDE, so any reaction approaching even a fraction of the CME’s

pre-release adjustment would have been detectable. Non-rejection therefore bounds the leak-window response tightly near zero rather than merely reflecting low power; the full-window statistic (0.0354 log-odds) lies well above the MDE and corresponds to the PELT (Pruned Exact Linear Time) change-point identified at $T = 35$ minutes.

3.2 Post-release latency: 5-minute delay after universal public release

Institutional benchmark. The CME 30-Day Federal Funds Future (ZQU4) adjusted within seconds of the 08:00 ET leak, with professional trading desks completing nearly 60% of the total 08:00–09:00 ET price adjustment before the official release (Figure 1). By 08:30 ET the cooling-CPI signal was therefore not only universally public but already substantially reflected in the closest institutional analogue to the Polymarket Fed contracts.

Polymarket adjustment delay. Material price adjustment on Polymarket did not commence until 08:35 ET. This 5-minute post-release delay is endogenously identified through change-point detection⁵ (Figure 2) and is the residual that the bid–ask–band rationalization in Section 3.1 cannot accommodate: by 08:30 ET the information was universally public and CME had completed the bulk of its move—a materially larger shock than the 1–2 percentage-point no-arbitrage band on Polymarket Fed tokens—so the 5-minute window between 08:30 ET and 08:35 ET is not explained by spread frictions of the leak-window kind. This 5-minute component is the *informational inertia* of the paper’s title; the term is not applied to the pre-release window.

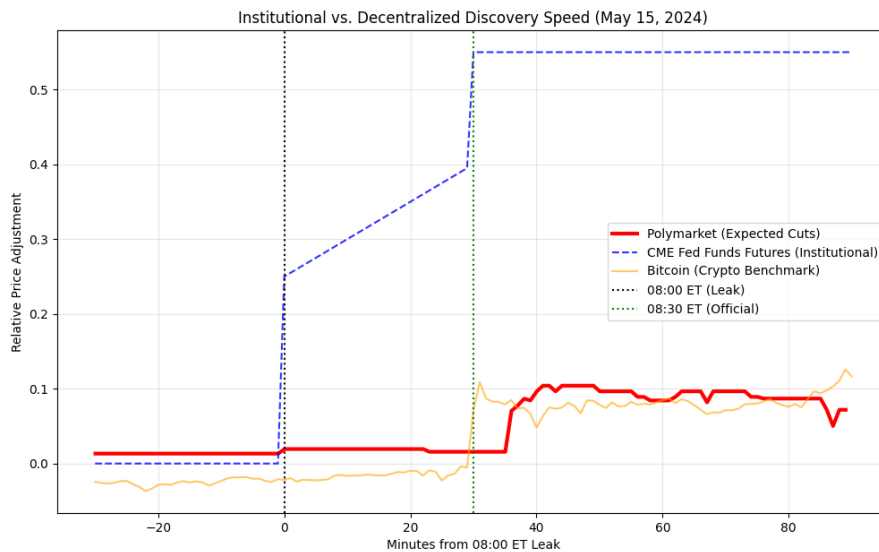


Figure 1: Institutional vs. Decentralized Discovery Speed. The figure plots the relative price adjustment of CME Fed Funds Futures (primary institutional benchmark), Bitcoin (continuous-monitoring contextual reference), and Polymarket (decentralized retail prediction market) during the May 15 CPI leak. “Relative Price Adjustment” normalizes each series to the fraction of its total 08:00–09:00 ET move realized by each minute. CME completed nearly 60% of its eventual adjustment before the 08:30 ET official release; Polymarket remained flat until 08:35 ET.

⁵We use the PELT change-point algorithm (Killick et al., 2012) with a change-in-mean cost function. The penalty parameter is set to the BIC criterion; the minimum segment length is set to 5 minutes to preclude spurious micro-breaks. Results are robust to MBIC penalization and minimum segment lengths of 3 and 10 minutes. Full replication code is available in the data repository.

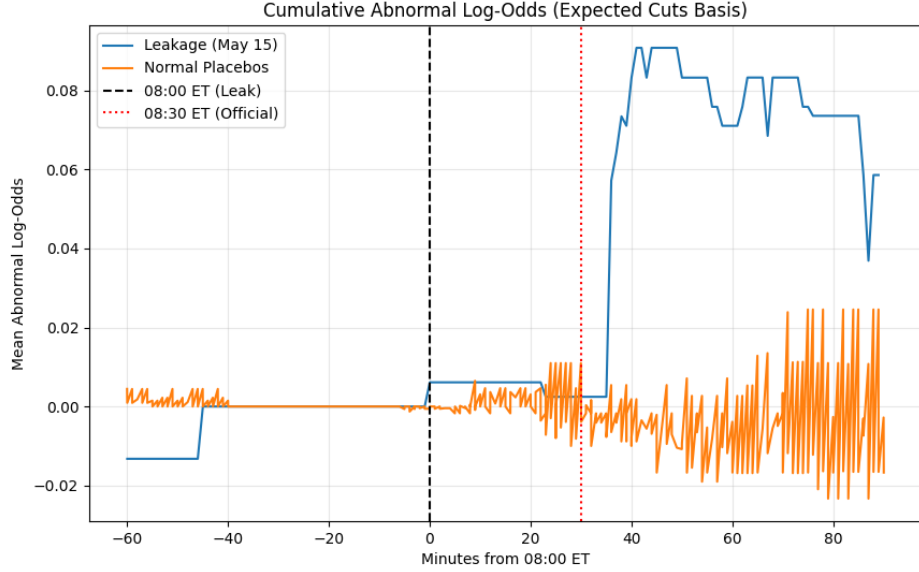


Figure 2: Cumulative Abnormal Log-Odds around Information Arrival. Abnormal log-odds is the generalized log-odds of the expected number of rate cuts, $\ell = \ln(E/(6 - E))$, measured relative to the mean over the 07:45–07:55 ET pre-leak baseline. Vertical lines mark the 08:00 ET leak and the 08:30 ET official release. The first significant break in Polymarket prices is identified at $T = 35$ minutes relative to the leak by the PELT change-point algorithm (Killick et al., 2012).

Microstructure pointer. Inter-arrival times between successive quote updates during the leak window were exceptionally long: only 3 best-bid or best-ask updates were recorded across all outcome tokens during the 08:00–08:29 ET window on May 15, compared with 11, 16, 20, and 50 updates on the four placebo mornings (mean ≈ 24). Appendix C reports the Kaplan–Meier survival function and a formal log-rank test ($p = 0.016$ against the pooled placebo, with the inferential caveats noted there); the pattern is consistent with stale quoting and limited updating rather than active transacting at pre-leak prices. Distinguishing pure absence of monitoring from strategic deeper-book withdrawal requires depth data that are not available for the May 15 event window; we therefore avoid mechanism claims at that level (see Section 4).

Coherence pointer. The summed contract probabilities remained close to 1 during the 08:00–08:30 ET window—absolute deviation ≈ 0.018 , qualitatively similar to the pre-leak deviation of 0.020 (Table E.1)—and rose to ≈ 0.088 during 08:30–09:00 ET as asynchronous repricing across tokens began. This rules out a collapse-of-coherence reading for the leak window and is consistent with normal post-release dislocation rather than leak-specific confusion.

4 Limitations

We summarize the main limitations of the design and the consequent scope of the claim.

1. *Single verified event* ($N = 1$). The analysis centers on a single verified early-release incident and is descriptive of this episode rather than a universal estimate of prediction-market efficiency; no comparable verified incident has been documented on Kalshi, PredictIt, or other platforms.
2. *Routine versus unannounced reaction lag.* An unannounced leak finds participants less prepared than a scheduled release, so the routine reaction lag on Polymarket is plausibly no greater than

what we observe here: the 35-minute interval is an *upper bound* on the routine discovery lag—equivalently, a *lower bound* on the discovery speed the contract series could plausibly achieve under standard, anticipated releases. The 5-minute post-release component (Section 3.2) may itself reflect a baseline release-processing latency on the platform; quantifying this on ordinary CPI mornings is a direct test we flag for future work once a larger sample of scheduled releases is available.

3. *Small placebo set.* The placebo group is restricted to four CPI mornings to preserve comparability of contract structure and depth. Statistical power is correspondingly limited, and we recharacterize the randomization, synthetic control, and survival exercises as bounding rather than as sharp hypothesis tests (Appendices A–C). The MDE calculation in Section 3.1 indicates that the leak-window response is bounded tightly near zero by the design, but a larger placebo sample would naturally tighten that bound further.
4. *Unobserved depth and strategic withdrawal.* The Polymarket quote-update feed exposes top-of-book but not full depth. With only 3 updates across all tokens during the leak window, the top-of-book is nearly static; what we cannot observe is whether resting size behind the inside quotes was withdrawn. Distinguishing pure absence of monitoring from deeper-book strategic withdrawal requires depth data not available for the May 15 event window.
5. *Scope of the informational-inertia claim.* We apply the label *informational inertia* only to the 5-minute post-release component (Section 3.2). The 30-minute pre-release window is bounded by a no-arbitrage zone (Section 3.1) and is not, in our reading, evidence of behavioral failure on the venue.
6. *Cross-venue arbitrage frictions.* The implied policy reading is correspondingly narrow. This episode does not condemn prediction markets as a class; it documents the efficiency cost of a thin, CLOB-only venue lacking designated market makers, liquidity rebates, automated news-aware strategies, and the institutional monitoring apparatus that concentrates around scheduled releases. Cross-venue arbitrage against CME faces material frictions—USDC on/off-ramping, settlement delays, and binary-contract hedging basis—that make instantaneous price transmission from institutional futures into Polymarket tokens uneconomical at the minute frequency even when both venues are observed in real time. Whether closing the gap requires incentive-based liquidity provision, algorithmic-friendly APIs, lower cross-venue frictions, or simply deeper natural participation is an open design question that future events, on Polymarket or on platforms such as Kalshi and PredictIt, will help adjudicate.
7. *Geographic and timezone friction.* A simple timezone friction is one potential alternative explanation: Polymarket restricts U.S.-based participation, so 08:00 ET might coincide with sleeping traders. A geographic check rules this out: the window corresponds to 13:00 BST in London, 14:00 CEST in Europe, and 20:00 SGT in Asia—prime trading hours for Polymarket’s primary non-U.S. user base. Being awake is necessary but not sufficient; the binding constraint is the absence on the platform in May 2024 of automated macro-monitoring infrastructure (data-release feeds, programmatic trading, low-latency APIs) of the kind that renders institutional venues timezone-irrelevant. The two explanations are complementary, not competing.

5 Conclusion

We use the verified May 15, 2024 CPI leak as a single-episode descriptive case study of how quickly a decentralized retail prediction market incorporates a high-signal macroeconomic surprise. Over the 35-minute interval between the inadvertent 08:00 ET upload and the first detectable Polymarket price adjustment at 08:35 ET, Polymarket Fed contracts exhibited a near-zero top-of-book response while CME 30-Day Federal Funds futures completed the bulk of their adjustment well before the official release. We separate this interval into two conceptually distinct components: a 30-minute pre-release non-response (08:00–08:30 ET) that is bounded by a no-arbitrage zone given the modest realized shock and the 1–2 percentage-point bid–ask band, and a 5-minute post-release latency (08:30–08:35 ET) that the same bid–ask logic cannot rationalize once the information is universally public and CME has already moved. This 5-minute residual is what we mean by *informational inertia* in this paper. Polymarket liquidity and algorithmic participation have grown substantially since mid-2024, and this episode provides a pre-growth baseline against which future measurements of real-time efficiency on the venue can be calibrated.

Data Availability Statement

The full replication package—a sequential Python analysis pipeline, processed minute-level Polymarket, Binance BTC/USDT, and ETH/USDT series, a derivative minute-level summary of the CME 30-Day Federal Funds future (ZQU4), and all figures and tables in the manuscript and online appendix—is archived at the Harvard Dataverse (Aktuğ and Torul, 2026), <https://doi.org/10.7910/DVN/7FWIQ0>. The record is currently in draft status during peer review; editors and referees may request access at any time, and the record will be released publicly upon acceptance. Binance OHLCV data are publicly accessible via the Binance API. CME ZQU4 tick data are licensed from CME DataMine (<https://datamine.cmegroup.com/>) and cannot be redistributed; the derivative minute-level series included in the replication package is sufficient to reproduce the institutional-benchmark comparison.

References

- Aktuğ, E. and Torul, O. (2026). Replication data for: Informational inertia in a decentralized prediction market: Evidence from the May 2024 CPI leak. Harvard Dataverse, <https://doi.org/10.7910/DVN/7FWIQ0>.
- Andersen, T. G. and Bollerslev, T. (1998). Deutsche mark–dollar volatility: Intraday activity patterns, macroeconomic announcements, and longer run dependencies. *The Journal of Finance*, 53(1):219–265.
- Arrow, K. J., Forsythe, R., Gorham, M., Hahn, R., Hanson, R., Ledyard, J. O., Levmore, S., Litan, R., Milgrom, P., Nelson, F. D., et al. (2008). The promise of prediction markets. *Science*, 320(5878):877–878.
- Balduzzi, P., Elton, E. J., and Green, T. C. (2001). Economic news and bond prices: Evidence from the U.S. Treasury market. *Journal of Financial and Quantitative Analysis*, 36(4):523–543.
- Berg, J. E., Nelson, F. D., and Rietz, T. A. (2008). Results from a dozen years of election futures markets research. *Handbook of experimental economics results*, 1:742–751.

- Bureau of Labor Statistics (2024). Statement from BLS commissioner on the early release of CPI data. U.S. Department of Labor. Available at: <https://www.bls.gov/cpi/>. Accessed: April 12, 2026.
- Evans, M. D. and Lyons, R. K. (2002). Order flow and exchange rate dynamics. *Journal of Political Economy*, 110(1):170–180.
- Fisher, R. A. (1935). *The design of experiments*. Oliver and Boyd.
- Fleming, M. J. and Remolona, E. M. (1999). Price formation and liquidity in the U.S. Treasury market: The response to public information. *The Journal of Finance*, 54(5):1901–1915.
- Grossman, S. J. and Stiglitz, J. E. (1980). On the impossibility of informationally efficient markets. *American Economic Review*, 70(3):393–408.
- Killick, R., Fearnhead, P., and Eckley, I. A. (2012). Optimal detection of changepoints with a linear computational cost. *Journal of the American Statistical Association*, 107(500):1590–1598.
- Manski, C. F. (2006). Interpreting the predictions of prediction markets. *Economics Letters*, 91(3):425–429.
- Sims, C. A. (2003). Implications of rational inattention. *Journal of Monetary Economics*, 50(3):665–690.
- Wolfers, J. and Zitzewitz, E. (2004). Prediction markets. *Journal of Economic Perspectives*, 18(2):107–126.

Appendix: Robustness and Microstructure Analysis (A–E)

Editorial note: Appendices B–E are intended for online publication only.

This appendix reports supplementary evidence designed to clarify the interpretation of the main result. Given the small placebo set ($N = 4$ matched CPI mornings) and the single verified event, we treat the exercises in this appendix as *bounding exercises* rather than as sharp hypothesis tests; the main-text summary of what the design can and cannot detect appears in Section 3.1. The common message across these exercises is that the May 15 leak-window non-response in Polymarket is best understood as limited updating and stale quoting rather than noisy immediate repricing.

A. Randomization Inference

To assess whether the absence of repricing during the May 15 leak window is unusual, we construct an empirical placebo distribution using temporal shifts across matched CPI mornings. Specifically, for each of the four matched CPI mornings, we apply 30 one-minute temporal shifts to the [08:00–08:30 ET] window, stepping the start time from 07:30 to 07:59 ET in one-minute increments; across four mornings this yields 120 placebo windows. All placebo windows fall strictly pre-release: the design compares the May 15 leak window against thirty-minute windows of routine pre-release pricing on ordinary CPI mornings, which is the relevant counterfactual given that the question is whether the leak window was unusual relative to what a quiet pre-release interval normally looks like. For each window we compute the mean abnormal log-odds. Figure A.1 shows that the May 15 statistic (0.0053) lies within the main mass of the placebo distribution, yielding a two-sided p -value of 0.157. Because adjacent rolling windows overlap by 29 of 30 minutes, the effective number of independent placebo draws is substantially smaller than 120 and the serial dependence could distort the p -value in either direction; we therefore also report the coarser

four-block permutation test using the four non-overlapping 08:00–08:30 ET placebo windows as the sole reference distribution. In that test, the May 15 statistic ranks third of five observations (event plus four placebos); with only five independent draws, no rank-based two-sided permutation test can reject at conventional levels, and we report the rank fraction (3/5) purely as a descriptive summary—the event sits at the median of the four-day reference distribution—rather than as an inferential p -value. A potential concern with the rolling design is that the latest placebo windows (starting at 07:55–07:59 ET) extend into the minutes just before the scheduled 08:30 ET release on placebo days, where mild anticipatory positioning could inflate placebo variance and bias toward non-rejection; the block test, whose windows terminate exactly at 08:30 ET, is in fact *more* exposed to such positioning than the rolling design, since the entire block window sits within the minutes immediately preceding the scheduled release. That the block test still places the May 15 statistic inside the reference mass therefore corroborates, rather than undermines, the rolling-design conclusion. The qualitative conclusion is the same under either construction: the May 15 leak-window path is indistinguishable from routine pre-release inactivity.

The implication is narrow but important. The leak-window path on May 15 is statistically indistinguishable from routine pre-release inactivity in these contracts. This result strengthens the interpretation that the market did not incorporate the leaked information before 08:30 ET; it does not imply that May 15 exhibited an unusually large abnormal null relative to ordinary mornings.

To address the natural concern that non-rejection merely reflects low power, we compute the minimum detectable effect (MDE) of the rolling-window design. Under a location-shift model, a two-sided permutation test at the 5% level using this placebo distribution (mean ≈ 0.0005 , standard deviation ≈ 0.0062 , $N = 120$) has a critical absolute deviation of 0.0099 log-odds. Shifting the placebo distribution and computing empirical rejection probabilities, the MDE is approximately 0.0106, 0.0122, and 0.0156 log-odds at 50%, 80%, and 90% power, respectively—equivalently, roughly 0.009, 0.010, and 0.013 expected cuts using the local slope $dE/d\ell \approx 0.86$. The observed early-window statistic (0.0053 log-odds, ≈ 0.005 cuts) is less than half of the 80%-power MDE, so any reaction approaching even a fraction of the CME’s pre-release adjustment would have been detectable with high probability; the non-rejection therefore bounds the leak-window response tightly near zero rather than merely reflecting insufficient power. For comparison, the full-window statistic (0.0354 log-odds) lies well above the MDE, consistent with the PELT change-point at $T = 35$ minutes.

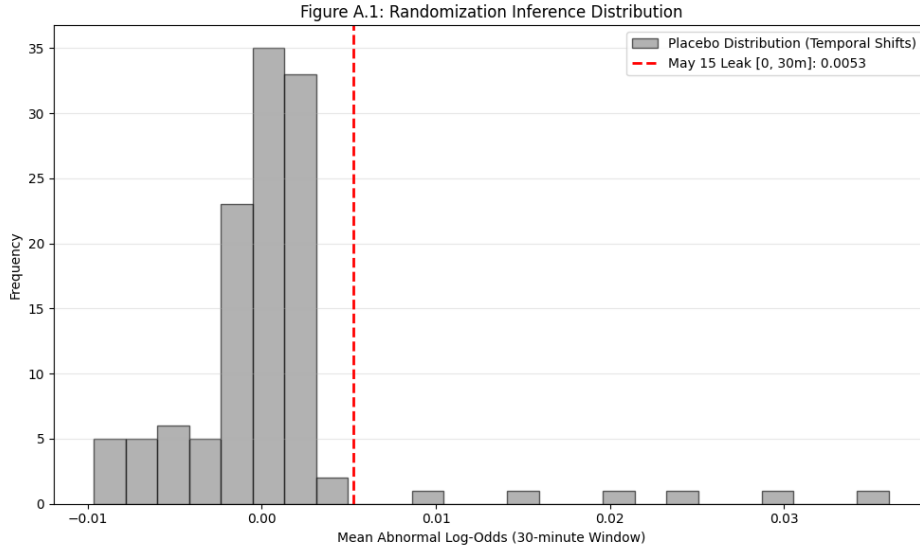


Figure A.1: Randomization Inference Distribution. The histogram reports placebo values of mean abnormal log-odds (the generalized log-odds of expected cuts relative to the 07:45–07:55 ET baseline) across 120 matched pre-release windows—30 one-minute temporal shifts of the 08:00–08:30 ET window on each of the four matched CPI mornings. The observed May 15 leak-window statistic (red solid line; positioned centrally) marks the non-response; the two-sided permutation p -value is 0.157.

B. Synthetic Control Minute Path

We construct a synthetic counterfactual for the May 15 Polymarket path using normal CPI release mornings as donor days. The donor weights are estimated over the pre-leak interval $T = -60$ to $T = -30$ to minimize the pre-treatment Root Mean Square Prediction Error (RMSPE), subject to non-negativity constraints and a convexity restriction that the weights sum to one. The resulting optimal weights are 1.000 for E004 (June 12, the first post-May-15 CPI release) and 0.000 for the other three donor days, yielding a pre-treatment RMSPE of 0.0055. Figure B.1 shows that the realized May 15 path and the synthetic path are nearly indistinguishable throughout the 08:00–08:30 ET window, consistent with the absence of an abnormal response. The concentration of weight on a single donor is a common feature of synthetic control estimators when the pre-treatment fit is tightly optimized over a small donor pool; consequently, this estimation functions effectively as a nearest-neighbor matched comparison and does not deliver independent inferential content beyond the randomization inference in Appendix A. The pre-treatment fitting window (07:00–07:30 ET) precedes the baseline window used in the main specification (07:45–07:55 ET); we use the earlier window here specifically to match the donor-day path before any pre-release positioning dynamics could begin, and Polymarket CLOB quotes are available continuously (the platform trades 24/7) through this interval.

The two series begin to diverge at 08:30 ET, when the synthetic control (representing normal placebo-day behavior) adjusts promptly to the official release while the realized May 15 path remains flat for an additional five minutes. This reinforces the interpretation that the absence of movement before 08:30 ET is not driven by an unusual drift on May 15; rather, through 08:30 ET the market is indistinguishable from a typical inactive pre-release morning, and the five-minute post-release gap between the realized and synthetic paths is itself the anomaly being measured.

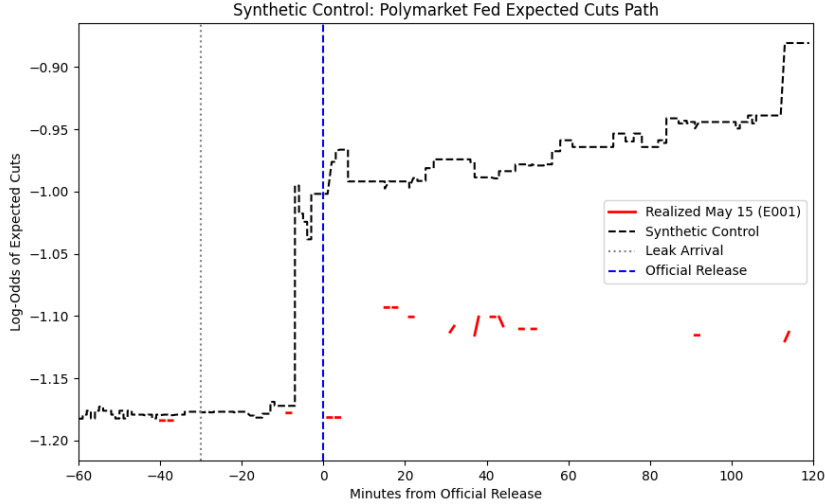


Figure B.1: Synthetic Control Path. The realized May 15 path (red solid line; higher series in the post-release window) closely tracks the synthetic counterfactual (black dashed line) during the leak window and diverges only after the official release.

C. Microstructure: Update Hazard and Inactivity

A natural concern is whether the flat leak-window path reflects active trading at stale prices or simply a lack of updates. We examine the inter-arrival times between price revisions during the 08:00–08:30 ET window—where a revision is defined as a change in the best bid or best ask price for *any* outcome token in the contract series. The analysis pools across all tokens, resetting the clock each time any token’s quote is updated. Figure C.1 plots the resulting survival probability of the no-update state.

On May 15, the inter-arrival durations between successive quote updates were exceptionally long: the Kaplan–Meier curve is shifted far to the right of the placebo baseline, indicating that the market spent the vast majority of the leak window in a no-update state. Raw update counts reinforce this pattern: only 3 best-bid or best-ask updates were recorded across all outcome tokens during the 08:00–08:29 ET window on May 15, compared with 11, 16, 20, and 50 updates on the four placebo mornings (mean ≈ 24). A formal log-rank test comparing the May 15 survival curve to the pooled placebo curve yields a p -value of 0.016, with May 15 exhibiting even fewer updates than the placebo baseline. The most direct interpretation is that the market was exceptionally inactive during this interval, pointing toward limited updating and stale quoting rather than active transacting at pre-leak prices. A competing mechanism is adverse-selection-driven withdrawal: liquidity providers who did observe institutional markets moving may have rationally cancelled limit orders and widened spreads to avoid being picked off, producing quote disappearance through strategic rather than informational channels. The top-of-book (best bid, best ask) is in fact observable from the quote-update feed and is nearly static throughout the window given only three updates; what is *not* observable without full depth snapshots is whether resting size behind the inside quotes was withdrawn. Distinguishing pure absence of monitoring from deeper-book withdrawal therefore requires depth data that are not available for the May 15 event window. The near-total lack of quote updates (3 vs. a placebo mean of 24) is consistent with either pure absence or strategic withdrawal; we note, however, that strategic withdrawal presupposes awareness of the leak, which itself implies a degree of information incorporation—just not one reflected in prices. A further observation partially constrains the strategic-withdrawal alternative: the best bid and best ask are reconstructed mechanically from the

quote-update feed, and with only 3 updates across all tokens the top-of-book prices were nearly static throughout the window. Active top-of-book spread widening would itself have generated further logged updates, so strategic withdrawal, if present, must have operated at deeper levels of the book rather than at the best quotes. Two caveats on the log-rank inference: outcome tokens within the series are mechanically linked (their probabilities must sum close to one), so pooled inter-arrival times are not independent renewal events; and the pooled placebo distribution spans a wide range of activity (11–50 updates across days), so significance could be partly driven by the single most active placebo morning. The raw count contrast (3 vs. 11–50) does not rely on the independence assumption and carries the same direction.

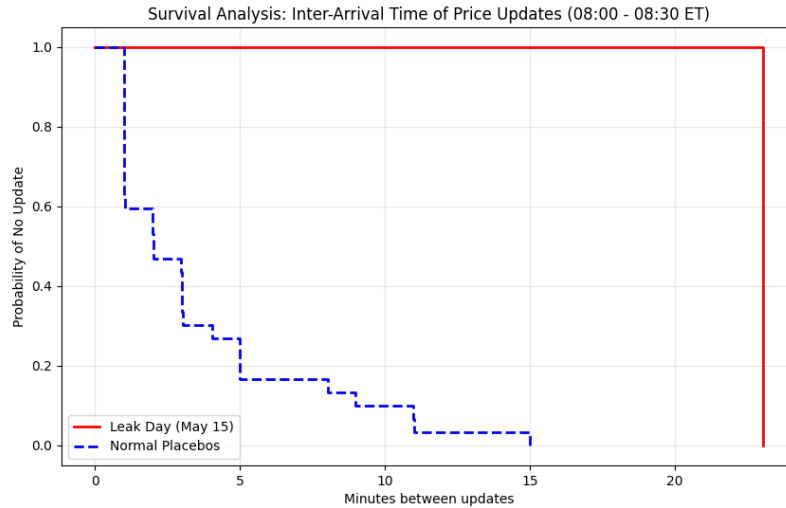


Figure C.1: Survival Analysis of Quote Inter-Arrival Times. The figure plots the Kaplan–Meier survival function of durations between successive best-bid or best-ask updates across all outcome tokens, with the clock reset at each update. The leak day (red solid line) is shifted far to the right, indicating much longer durations between updates compared with the pooled placebo baseline (blue dashed line) (log-rank $p = 0.016$).

D. Measurement Sensitivity

We assess whether the leak-window null result depends on specific measurement choices. Specifically, we re-estimate the early-window abnormal repricing statistic on a 3×2 grid: three baseline anchors (the minute at $T = -60$, $T = -30$, and $T = -10$ relative to the 08:00 ET leak) crossed with two outcome transformations (the generalized log-odds used in the main text and the raw implied probability on the expected-cuts scale). The grid uses single-minute baselines rather than the 10-minute average (07:45–07:55 ET) used in the main specification. Single-minute anchors are noisier than the 10-minute average, so this robustness check varies the anchor’s susceptibility to minute-level idiosyncrasy rather than supplying an independent power test; we read it as a consistency check, not as confirmation of the null. The six resulting statistics are reported in Figure D.1; all lie within a narrow band close to zero, indicating that the near-zero leak-window response is not an artifact of any one normalization choice.

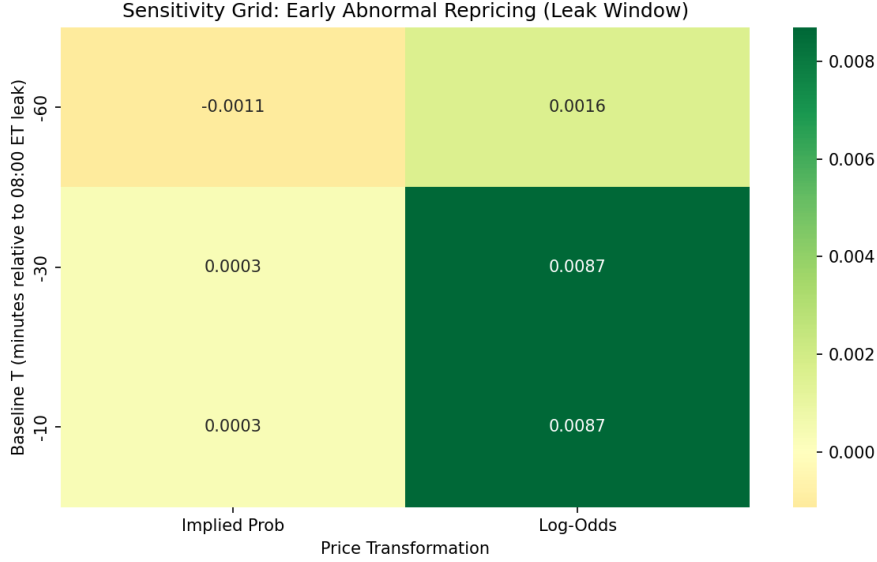


Figure D.1: Sensitivity Grid of Early Abnormal Repricing. The heatmap summarizes alternative constructions of the leak-window repricing statistic across a 3×2 grid that crosses three single-minute baseline anchors ($T = -60, -30$, and -10 minutes relative to the 08:00 ET leak) with two outcome transforms (the generalized log-odds used in the main text and the raw implied probability on the expected-cuts scale), with darker shades indicating values closer to zero.

E. Advanced Coherence Diagnostics

Logical coherence is measured as the absolute deviation of the sum of implied probabilities from the theoretical bound of 1.0. Table E.1 shows that this deviation remained stable at approximately 0.02 during the leak window, consistent with prevailing bid-ask spreads. The post-release deviation of 0.088 reflects asynchronous repricing across outcome tokens following the official release, and closing toward baseline levels over subsequent minutes as cross-contract arbitrage equilibrates the market. To benchmark this magnitude, we compute the same post-release coherence statistic on the four placebo CPI mornings: the placebo mean is approximately 0.070 with a wide range of 0.019–0.188. With only four placebo observations this is a weak reference distribution and we avoid statistical claims; in point-estimate terms the May 15 value of 0.088 sits within the placebo range, consistent with normal post-release dislocation rather than leak-specific confusion. The stable pre-release coherence followed by elevated but transient post-release deviation is consistent with the limited-updating interpretation in Section 3.1 and with the 5-minute post-release latency we label informational inertia in Section 3.2: the market is essentially inactive throughout the 08:00–08:35 ET window and only begins the asynchronous process of price discovery from 08:35 ET onward, after which the cross-token sum temporarily departs from unity while tokens reprice at different rates.

Table E.1: Formal Coherence: Absolute Deviation from $\text{Sum}(P)=1$

Period	Average Absolute Deviation
Pre-Leak (07:00–08:00)	0.020
Leak Window (08:00–08:30)	0.018
Post-Release (08:30–09:00)	0.088

Notes: Coherence is measured as the absolute deviation of summed contract probabilities from one. Leak-window deviations are consistent with bid–ask frictions and rounding. The post-release increase reflects asynchronous adjustment across outcome contracts following the official release.