Human vs. Machine: Cryptocurrency Microstructure and Market Participants *

Andrew Caosun† Markus Pelger‡ Khizar Qureshi§ Greg Zanotti¶

This draft: August 13, 2026

Abstract

This paper is built on the first dataset, to our knowledge, that separates the two populations trading in a modern electronic market. Coinbase's public data feed exposed, for every order in its Bitcoin–US dollar market, whether the order arrived through the machine interface or through the trading screen. We discovered this identifying field and reconstructed the complete order book from roughly 20 billion messages over eighteen months, with every one of 6.7 billion orders assigned to a human or a machine population. The data reveal two facts about human traders. Human trades are slow, with the two populations living on clocks eighty times apart. Human trades are predictable, since headline tone and recurring order-book states forecast human trading far better than machine trading. That combination is exploitable, and we price its consequence. Simply knowing a trader's population sharply improves out-of-sample forecasts of where liquidity will go next. In fleeting states that resemble spoofing, machines revise and protect their quotes within seconds while humans absorb the signal slowly. Around bursts of aggressive trading, the human share of aggressive volume nearly quadruples just as the cost of trading jumps about 70 percent. Adding up these episodes, the humans caught in them pay an extra 9.1 basis points on their exposed trades, more than triple their normal cost of trading. The same accounting applied to machine traders yields 1.99 basis points, so the human burden is 4.6 times larger, and only humans retain a positive net cost. In bear markets these states occur 58 percent more often per day. Market quality is a distribution across participants, and the slower side pays precisely when trading matters most, a gap that matters for how markets are monitored and designed.

Keywords: cryptocurrency, limit order book, high-frequency trading, algorithmic trading, market manipulation, retail investors

*We thank Ruizhe Jia, David Mazières, Christine Parlour, Geoff Ramseyer and seminar and conference participants at Stanford University and at the Santa Clara University Crypto Conference, the International Conference on AI in Finance and the Stanford IOG Research Hub Conference for helpful comments. We thank the Stanford Future of Digital Currency Initiative (FDCI) and the Stanford IOG Research Hub for generous research funding.

†Stanford University, Department of Management Science & Engineering, Email: acaosun@stanford.edu.

‡Stanford University, Department of Management Science & Engineering, and NBER, Email: mpelger@stanford.edu.

§Stanford University, Department of Management Science & Engineering, Email: kqureshi@stanford.edu.

¶Stanford University, Department of Management Science & Engineering, Email: gzanotti@stanford.edu.

1. Introduction

Modern financial markets are segmented by speed, information, access technology, and decision protocol. Some orders are submitted manually after a person observes a screen, interprets a price move, or reacts to news. Other orders are generated by systems that continuously ingest the full state of the limit order book, update queue positions, and cancel or replace stale quotes within fractions of a second. The central economic question is therefore not only whether prices are efficient or whether liquidity is deep, but also *for whom* the market is efficient and *against whom* the market is liquid. Our answer begins from two facts that this paper documents at the order level. Trades placed by people are slow, and trades placed by people are predictable. Slowness alone would not matter if manual trading were unpredictable, and predictability alone would not matter if people could react in time. Together they create a margin that technology can exploit, and the paper's central exercise is to see that margin directly and to price it. The implication is that every headline measure of market quality, spreads, depth, price efficiency, is an average over all participants, and the average is the wrong statistic. Market quality is a distribution across trader populations, and it is the distribution, not the mean, that determines who actually receives the liquidity the market appears to offer.

Cryptocurrency exchanges provide a sharp setting in which to study this question. They combine continuous centralized limit order books, twenty-four-hour trading, high retail participation, substantial volatility, and a comparatively thin regulatory perimeter. They also concentrate a large share of price discovery in venues whose trading protocols resemble modern equity and futures markets but whose participants differ markedly in sophistication, intermediation structure, and protection from abusive order-book conduct. These features make crypto markets a natural setting for the microstructure of technological inequality where the market is transparent at the book level, but not symmetric in speed. It is accessible to retail traders, but not equal in execution quality. It is liquid in normal times, but potentially fragile precisely when human traders demand immediacy, meaning that they need to trade at once rather than wait.

This paper studies the interaction between human and machine trading in a high-frequency cryptocurrency limit order book. Our dataset is unique in two ways. We observe every order and every trade on a major exchange, and for each order we know how it entered the market. The data are the complete market-by-order feed of Coinbase, the largest US-regulated cryptocurrency exchange and an economically large market in its own right, covering January 2021 through June 2022. Orders reach an exchange through one of two doors. Trading software connects through the machine interface, the API, and attaches its own tracking number to each order so that its code can manage the order later. A person clicking on the trading screen supplies no such number. We detected that Coinbase's public data stream exposed this field, and recognizing that a book-keeping artifact identifies the technology of order entry is a creative measurement contribution of this paper. The field labels every one of 6.7 billion orders, covering $432.78 billion of executed transactions, as API-routed or UI-associated. Coinbase stopped disseminating the field after our eighteen-month window, so the dataset cannot be constructed again. It is not only unique but

unrepeatable. The reason this decomposition has not been done before is simple. No one could see the participants, and studies of equity markets must infer automation from market-level speed proxies. Throughout the paper we refer to the API-routed population as "machine" traders and to the UI-associated population as "human" traders. These labels are population-level proxies for trading technology rather than verified identities, every claim in the paper is a claim about these route-defined populations, and the human cohort is a measured subset of manual trading rather than the venue's entire retail activity. The level of observation is an order-book message, so additions, cancellations, and matches can be assigned to the population that submitted or supplied the relevant order. This permits a decomposition of liquidity provision, liquidity demand, cancellation behavior, taker aggressiveness, and trading costs by trader population.

The empirical focus is the interaction between two forms of asymmetry. The first is the standard information asymmetry of market microstructure, where some order flow is more informative, and liquidity suppliers respond by changing quotes, depth, and execution terms (Kyle, 1985; Glosten and Milgrom, 1985; Hasbrouck, 1991; Back et al., 2018; Easley et al., 2016). The second is technological asymmetry, where even when all participants observe the same public book and the same public news, some participants can update orders and condition on order-flow states far more quickly than others (Hendershott et al., 2011; Brogaard et al., 2014; Chaboud et al., 2014; Hasbrouck and Saar, 2013; Easley et al., 2021). Our main premise is that these two asymmetries are not separable. In a market with fast cancellation, queue priority, and order-resolved depth, technology can transform public information into private execution advantage, and can transform transient displayed liquidity into a signal that slower traders may misinterpret.

The paper asks one question. How does the predictability of human trading get exploited in a modern electronic market, and what does it cost? The answer proceeds in three steps. The prediction step shows that human trading is forecastable from its own past, from the state of the book, and from public news, while machine trading is not. The exploitation step identifies the states where that predictability meets machine speed. We measure order-book configurations that resemble spoofing, where displayed depth is used as a misleading signal, and momentum ignition, where bursts of aggressive trading start price moves that attract chasers, and we benchmark both against public news arrivals that no trader controls. In all three settings machines act first and humans absorb, exactly the conditions under which a predictable trader can be traded against. The pricing step shows what the combination costs. Humans caught in these episodes pay an extra 9.1 basis points on their exposed trades, more than triple their normal cost of trading, and the harm is heavily right-tailed. Machines trading through the same events face only 1.99 basis points on their own exposure, so the burden is 4.6 times larger for humans, and after gains and losses are netted only the human cohort still pays. A market-phase decomposition shows that bear markets hurt through frequency rather than severity, with momentum-ignition states occurring 1.58 times as often per day.

The paper makes four contributions. The first is the dataset itself, an order-resolved reconstruction of an entire market with every trader population identified, built by replaying more than

20 billion messages in sequence, a substantial engineering undertaking and, because the identifying field is gone, an unrepeatable one. The construction preserves the distinction between displayed, executed, and canceled liquidity, which is essential in a market where much of the strategic interaction occurs through orders that never execute. The second is a two-population microstructure framework that decomposes signed order flow into its human and machine components, bridges classical informed-trading models and empirical order-book prediction, and shows that knowing who trades improves forecasts of where liquidity will go. The third is the measurement of strategic order-book states. We construct transparent, state-conditioned detectors for spoofing-like and ignition-like activity, validate them against placebos and public news, and document how each population behaves around them. The fourth is an incidence framework that prices technological asymmetry. A market can look liquid on average and still transfer rents from slower to faster traders exactly in the episodes when the slower side most needs immediacy, which matters directly for surveillance and market design.

Together the steps make the chain visible end to end, and the lesson is general. Once the participants are visible, market quality becomes distributional, and the slower, more predictable side pays precisely when trading matters most.

Section 2 describes the market-by-order data, the reconstruction of the limit order book, the route labels and their validation, and the measurement conventions used throughout the paper. Section 3 studies human-machine maker and taker flows and develops the two-population price-impact framework. Section 4 studies spoofing-like and momentum-ignition-like activity, including event-time evidence on participation, spreads, volatility, and liquidity. Section 5 studies news and trader behavior using timestamped CoinDesk and Cointelegraph headlines and cross-exchange order-book shocks from Binance. Section 6 translates spoofing, momentum ignition, and news into realized execution costs, risk-capital costs, participation-capacity losses, and liquidation-risk amplification. Section 7 concludes. The sections form one arc. We validate the labels, model what they change, detect the states where a speed advantage should matter most, benchmark against public information that no trader controls, and price the consequences as rates per dollar of exposure. An Internet Appendix contains the mathematical derivations, additional stylized facts and diagnostics, detector implementation details, supplementary news and cost results, and an institutional fee-schedule analysis.

Related Literature. The paper connects six literatures that have largely evolved separately. The first is the canonical literature on adverse selection, price impact, and limit order books. Theoretical models of informed trading explain why order flow moves prices, why depth is endogenous, and why liquidity is costly when trades are correlated with information (Kyle, 1985; Glosten and Milgrom, 1985; Glosten, 1994). Empirical work shows that trade innovations, order-book states, and order-flow imbalance contain information about subsequent prices and liquidity (Hasbrouck, 1991; Biais et al., 1995; Eisler et al., 2012; Cont et al., 2014; Bouchaud et al., 2018; Cont et al., 2023). We build on this tradition but replace the representative order flow with a labeled order-flow vector,

meaning that flow is carried separately for each trader population. This distinction matters because the same aggregate imbalance can arise from human liquidity demand, machine liquidity supply, machine cancellations, or cross-population trades. Aggregation therefore confounds the mechanism. By estimating population-specific impact and liquidity coefficients, the paper asks whether the adverse-selection object is attached to order flow itself or to the technology and population that generate it.

The second literature studies algorithmic trading, high-frequency trading, and speed. A large empirical literature evaluates whether automated trading improves market quality, contributes to price discovery, anticipates institutional order flow, or changes the incentives to acquire information (Hendershott et al., 2011; Brogaard et al., 2014; Chaboud et al., 2014; Menkveld, 2014; Hirschey, 2021; Weller, 2018; Kang et al., 2022; Garriott and Riordan, 2024). Related work studies low-latency activity, the price process, and the microstructure of the machine age (Hasbrouck and Saar, 2013; Aït-Sahalia and Brunetti, 2020; O'Hara, 2015; Easley et al., 2021, 2024). We extend this literature in two directions. First, the setting is a crypto exchange with substantial direct retail participation and weaker market-wide safeguards than regulated equity and futures markets. Second, the data distinguish human and machine order-flow contributions at the order-book event level, allowing us to observe not only whether algorithms are active, but also who provides liquidity to whom, who removes liquidity from whom, and how those interactions evolve around strategic order-book states.

The third literature studies retail trading and the distribution of trading gains. Evidence from equity markets shows that retail traders often pay large performance and execution costs, that aggressive retail orders can be especially costly, and that retail liquidity provision can be compensated in some states but not others (Barber and Odean, 2000; Kaniel et al., 2008; Barber et al., 2009; Barrot et al., 2016; Eaton et al., 2022; Sağlam et al., 2019). This paper brings the retail question to a market in which retail participation is central rather than peripheral and in which the relevant trading horizon is often seconds rather than days or months. The resulting object is not portfolio performance but microstructure performance: slippage, effective spreads, markouts, adverse selection, and changes in future liquidity conditional on human-machine interaction.

The fourth literature studies cryptocurrency markets. Existing work documents segmentation and arbitrage frictions across exchanges, the role of stablecoin flows, manipulation episodes, pump-and-dump schemes, token valuation, and whether crypto markets differ from traditional markets (Makarov and Schoar, 2020; Griffin and Shams, 2020; Gandal et al., 2018; Li et al., 2021; Cong et al., 2021; Bartolucci and Kirilenko, 2020; Kogan et al., 2024; Aleti and Mizrach, 2021). Much of this literature is cross-exchange, daily, on-chain, or trade-level. Our contribution is inside the exchange and inside the second. We study the full event-level evolution of one limit order book, where the relevant frictions are queue position, fleeting displayed liquidity, cancellation speed, and the decomposition of liquidity supply and demand across human and machine populations.

The fifth literature studies manipulation, spoofing, collusion, and adversarial behavior in automated markets. Strategic traders can profit by exploiting the predictable response of other traders,

by creating temporary price pressure, or by altering displayed depth without intending to trade (Brunnermeier and Pedersen, 2005; Tao et al., 2022; Cartea et al., 2022, 2023, 2025a; Tremacoldi-Rossi, 2024). Recent methodological work proposes structural, learning-based, and statistical tools for identifying spoofing-like behavior from order-book data (Byrd, 2022; Gu et al., 2024; Aggarwal and Khan, 2024). Our empirical design combines these ideas with cohort labels. This allows us to ask not only whether a suspicious order-book state is present, but whether it predicts human liquidity demand, human liquidity withdrawal, machine liquidity taking, or future net liquidity.

The sixth literature develops high-frequency prediction and text-based measures of information arrival. Machine learning models extract predictive structure from order books, order-flow imbalance, and high-frequency event sequences (Ntakaris et al., 2018; Zhang et al., 2019; Kolm et al., 2021; Briola et al., 2024; Ait-Sahalia et al., 2022; Cartea et al., 2025b; Goyenko et al., 2024). Separately, the textual finance literature studies how public information, news tone, and news categories move prices and trading activity (Tetlock, 2007; Tetlock et al., 2008; Boudoukh et al., 2013; Foucault et al., 2016; Aleti and Bollerslev, 2024). Our setting merges these two literatures. News is not only a public signal about fundamentals; it is also a synchronization device that changes the strategic state of the order book. We therefore study both direct reactions to news and reactions to the order-flow responses induced by news.

The closest papers establish the objects that organize our design, but none observes the human-machine margin at the order level. Kyle (1985), Glosten and Milgrom (1985), and Hasbrouck (1991) explain why order flow moves prices and why liquidity suppliers charge for adverse selection; we let the information content of order flow differ by trading technology. Hendershott et al. (2011), Brogaard et al. (2014), and Easley et al. (2021) measure automation at the market level; we move to an order-resolved split between manual and automated access. In crypto, Makarov and Schoar (2020), Griffin and Shams (2020), and Kogan et al. (2024) work across exchanges, at daily horizons, or from trades; we go inside one venue at sub-minute horizons, where we observe the book, the side of each trade, the cohort of the trader, and the cost paid by slower participants. Brunnermeier and Pedersen (2005) formalizes profiting from another trader's need to trade, and recent spoofing and surveillance designs show how transient displayed depth distorts inference; we use those mechanisms as interpretable state variables and study their distributional consequences rather than legal intent. Finally, Tetlock (2007), Boudoukh et al. (2013), and Foucault et al. (2016) show that public news, language, and speed matter for trading; we add cohort-level order-book responses and separate headlines that follow recent price trends from headlines that do not, which matters because crypto journalism often reacts to price action.

2. High-Frequency Crypto Market Microstructure

This section describes the data and the measurement framework. The analysis is not based on trades alone, nor on time-sampled quote snapshots. It is based on the event-by-event evolution of a centralized limit order book. In this electronic market, liquidity is not a static stock but

rather a stochastic inventory of executable interest. Liquidity appears when traders submit passive orders, disappears when aggressive orders consume queues, and disappears just as consequentially for trading costs when fast traders cancel before execution. A market can therefore look liquid in a low-frequency summary while being locally fragile in the states in which slower traders most need immediacy.

The cryptocurrency setting sharpens this because trading is continuous, with no open or close. Market participation follows global time zones, macro-information cycles, and platform-specific frictions rather than a single national trading calendar. At the same time, the matching protocol is the familiar price-time-priority limit order book. The market is therefore sufficiently close to canonical electronic markets to discipline the analysis, but sufficiently different to test which textbook limit-order-book facts survive in a continuous, retail-intensive, technologically heterogeneous setting.

2.1. Market-by-Order Data and Sample Construction

We use the event-level market-by-order stream for Coinbase BTC–USD from January 1, 2021 through June 30, 2022. Coinbase is the largest US-regulated cryptocurrency exchange. As of December 2021 it held $278 billion of assets on platform and accounted for 75.5 percent of trading on US-regulated exchanges and 11.5 percent worldwide. The paper studies this centralized limit order book throughout; decentralized venues, on-chain settlement, and custody are outside its scope. The feed preserves the identifier and lifecycle of each observed order rather than reporting only aggregate depth at a price level. That distinction is central to the paper. It permits exact reconstruction of the displayed limit order book, measurement of how long orders live before they trade or are canceled, and propagation of an order-entry-route indicator from submission through all later activity, so that on every trade we know the route of both sides: the resting order that supplied the quote (the maker) and the incoming order that traded against it (the taker). The raw archive contains 20.21 billion messages and 6.71 billion unique order submissions. Processing an archive of this scale into an analysis-ready book is a substantial engineering task: every message must be applied in sequence to a live representation of all resting orders, so that each execution and cancellation is attributed to the order, price level, and cohort it affects. Since many messages arise from the same order and the same local market state, message rows are not independent economic observations. Our order-level estimands therefore use unique orders, while statistical uncertainty is assessed by resampling or clustering in calendar time.

Venue and event stream. Coinbase operates a continuous double-auction limit order book under price–time priority. Trading proceeds twenty-four hours per day and seven days per week, without opening or closing auctions. The full-order-book channel contains five economically distinct message types that together trace every order from arrival to exit.[1] Every message carries a venue

[1] A `received` message records the arrival of a new order and assigns its identifier; `open` places the nonmarketable residual in the book; `match` records an execution between an incoming order and a resting order; `change` records an

sequence number. Starting from the most recent valid full-book snapshot and replaying sequence-valid messages in order recovers both the displayed book and the observed lifecycle of each order over intervals with complete capture.

Reconstruction, audit flags, and analysis samples. The pipeline audits every row and maintains two samples. The analysis-event sample removes rows flagged as incomplete, pre-snapshot, or redundant, and the lifecycle samples track each unique order from its arrival record to its exit record.[2] The audit finds no duplicate order identifiers and no processed day with nonmonotone timestamps. One deliberate restriction matters for later results. Cancellation records and execution records count different objects, since one filled order can generate several matches, so we never form cancellation-to-execution ratios from them. Figure IA.7 in the Internet Appendix reports the raw daily message ratio by route as a descriptive statistic, and Internet Appendix Section IA.B details the treatment of terminal messages.

Coverage and scale. The requested window contains 546 calendar days, of which 543 have a processed daily file.[3] Sequence placeholders account for 42.91 million rows, or 0.21% of the raw stream. Redundant and pre-bootstrap rows account for 320.03 million rows, or 1.58%. Because these categories overlap, the filters jointly remove 1.80% of raw observations and leave 19.85 billion analysis-event records. The final event stream contains 6.71 billion `received`, 6.51 billion `open`, 6.39 billion `done`, 248.90 million `match`, and 14,894 `change` messages.

Table I reports the principal sample statistics. The 248.90 million match messages represent 10.28 million BTC and $432.78 billion of executed notional (dollar value traded), corresponding to daily averages of 18,931 BTC and $797 million on processed days. Quote statistics are sampled in clock time rather than message time so that high-activity periods do not receive mechanically greater weight. Specifically, we retain the last valid quote in each minute. The median processed day contains all 1,440 quote minutes, and at least 95% of processed days have full minute coverage. The median daily time-sampled spread is one cent, which is a single tick.[4] Five-minute realized volatility, defined as the square root of the sum of squared five-minute log returns within the day, has a median of 3.48%.

Every order-book reconstruction, cohort assignment, execution measure, and trader-behavior outcome in the paper is computed from Coinbase data. The cross-exchange analysis of Section 5.4 additionally uses Binance prices as an external reference, and the Binance data are described there.

amendment, principally from self-trade prevention; and `done` removes the remaining order.

[2]The flags are `missing`, `skip`, and `ignore`, marking sequence placeholders for uncaptured events, observations preceding the valid replay snapshot, and messages redundant for event counting though needed for lifecycle reconstruction. Lifecycle samples begin from sequence-valid `received` records and canceled `done` records and deduplicate by `order_id`.

[3]The missing daily files for January 26–27, 2021 and June 9, 2021 are due to a disruption of the Coinbase data feed on our side, unrelated to market activity.

[4]Relative to the prevailing midpoint, one cent corresponds to approximately 0.00286 basis points, so the quoted spread is economically negligible in normal times, which is exactly why state-dependent costs rather than average spreads are the relevant object in this market.

Table I Sample and cohort composition

Measure	Estimate or UI share	95% block interval	Unknown share
Panel A: Sample construction and market activity			
Processed calendar days	543 of 546		
Raw message rows	20.21 billion		
Analysis-event rows	19.85 billion		
Unique received orders	6.71 billion		
Unique cancellation records	6.39 billion		
Match events	248.90 million		
Executed notional	$432.78 billion		
Median time-sampled spread	0.00286 bps (one tick)		
Median five-minute realized volatility	3.48%		
Panel B: UI-associated share of known-route activity			
New orders	1.31%	[1.11%, 1.52%]	0.00%
Cancellation records	1.10%	[0.92%, 1.32%]	0.06%
Maker match count	7.12%	[6.02%, 8.32%]	0.87%
Maker executed notional	4.29%	[3.75%, 4.81%]	1.65%
Taker executed notional	2.39%	[2.00%, 2.80%]	0.00%

Notes. The sample spans January 1, 2021 through June 30, 2022. Analysis-event rows exclude sequence placeholders, pre-bootstrap rows, and messages marked redundant. Lifecycle records are deduplicated by order identifier. Panel B reports the UI-associated share relative to API-routed plus UI-associated observations. Unknown-route observations are excluded from that denominator and reported separately. Confidence intervals use 999 circular moving-block bootstrap replications with seven-day blocks. Spread statistics use the last valid quote in each minute rather than weighting observations by message intensity.

2.2. Order-Entry Route and Cohort Validation

Cohort construction and interpretation. Orders reach Coinbase through two doors. Trading software talks to the exchange through its machine interface, the API, which is how bots, market-making systems, and the backends of third-party trading apps submit orders. A person logs into the website or the app and clicks. When software submits an order through the API, it can attach an optional field called `client_oid`, an order identifier that the client's own system generates so that its code can track the order, avoid duplicates, and match fills back to its records. The field exists purely for bookkeeping by programs. A person clicking on the trading screen has no such identifier to supply. Coinbase does not publish a direct indicator of whether a decision was made manually or by an algorithm. Our creative measurement contribution is to recognize that this bookkeeping field, which Coinbase's public stream exposed on every arriving order, identifies the technology of order entry. Coinbase ceased disseminating the field after June 2022, so the sample can be neither extended nor reconstructed. A nonempty identifier provides affirmative evidence that the order was submitted through the API. An empty identifier defines the residual UI-associated cohort, recorded as `napi=true` in the processed data. The machine label is positive evidence while the human label

is the absence of it, which is why we treat the human cohort as a residual category and validate it below rather than assume it. The route label is carried through the order lifecycle by joining on `order_id`. On an execution, `napi` identifies the route of the resting maker and `a_napi` identifies the route of the incoming taker.

The mapping measures infrastructure rather than legal identity. An API user can omit the optional client identifier, a manually chosen instruction can reach the venue through a third-party API, and software can in principle operate a browser interface. We therefore treat API-routed and UI-associated as the measured cohorts. The terms machine and human are economic shorthand, used after documenting that the two route-defined populations operate on sharply different clocks. The route is also selected rather than randomly assigned. Later comparisons characterize participants who sort into different trading infrastructures and should not be read as causal effects of routing an otherwise identical order through a different interface.

Composition. UI-associated activity is small in message count but economically visible in executions. It represents 1.31% of new orders, with a seven-day moving-block 95% interval of [1.11%, 1.52%]. Its monthly share ranges from 0.27% to 3.91%, which makes date-local comparisons important because pooled contrasts can otherwise load on changing market conditions and cohort composition. UI-associated orders account for 7.12% of maker match messages, 4.29% of maker executed notional, and 2.39% of taker executed notional. Orders already resting at the beginning of a reconstruction window can have an unknown route when their `received` message predates the bootstrap snapshot. Unknown labels are absent among observed new orders, equal 0.06% of cancellation records, and account for 1.65% of maker notional. We remove unknown labels from route comparisons and report their shares separately.

Validation estimands and inference. Every result in the paper rests on the quality of these labels, so we validate them before using them. The logic is simple. If the labels capture the technology of order entry, the two populations must act on different clocks. The sharpest fingerprint of automation is how fast a trader cancels a resting order, because canceling within a fraction of a second physically requires software while a person reading a screen cannot do it. We therefore measure cancellation latency, the time between placing an order and canceling it, and ask how many times longer human cancellations take than machine cancellations. We compare the whole distribution rather than a single average, so the separation cannot come from picking a favorable statistic. Formally, let $Q_r(q)$ denote the pooled qth latency quantile for route $r \in \{A, U\}$, and let $R(q) = Q_U(q)/Q_A(q)$ for $q \in [0.05, 0.95]$, so that values above one indicate that UI-associated orders remain in the book longer before cancellation. For inference we resample whole days, because orders within a day share market conditions and are not independent, and we report one confidence band that covers the entire curve, so the conclusion cannot rest on a single favorable quantile.[5]

[5]Inference uses 999 circular moving-block bootstrap replications with a seven-day block length, which preserves within-day dependence and short-run serial dependence across adjacent days. The band is a simultaneous 95% max-deviation band.

Table II Cancellation-latency validation of the route cohorts

Latency quantile	API-routed (seconds)	UI-associated (seconds)	Pooled UI/API ratio [95% CI]	Equal-day UI/API ratio [95% CI]
Median	0.16	13.36	81.42 [33.32, 273.91]	113.20 [79.16, 161.90]
90th percentile	3.94	227.98	57.86 [33.33, 86.96]	87.00 [66.70, 113.50]
95th percentile	9.20	874.39	95.06 [70.46, 127.68]	107.88 [83.45, 139.50]

Notes. Latency is elapsed time from `open` to a canceled `done` for unique orders. Pooled ratios compare quantiles of the order-weighted route distributions. Confidence intervals resample days in 999 circular moving-block replications with seven-day blocks. Equal-day estimates are exponentiated means of daily log quantile ratios, with confidence intervals based on Newey–West standard errors using seven lags. For independent pooled draws, $\Pr(T_A < T_U) = 0.88$, with a block-bootstrap interval of [0.86, 0.90].

We complement the order-weighted estimand with an equal-day comparison that averages the daily log quantile ratios and is algebraically equivalent to a two-cohort comparison with date fixed effects.[6] The two estimands answer distinct economic questions: the pooled ratio describes a randomly selected cancellation, while the equal-day ratio describes the typical trading day and prevents the most message-intensive dates from dominating the result.

Cancellation horizons. The route separation is large throughout the latency distribution. The API-routed median cancellation latency is 0.16 seconds, compared with 13.36 seconds for UI-associated orders: a trader who cancels within a fifth of a second is running software, while a trader who takes thirteen seconds is plausibly reading a screen. At the 90th percentile, the corresponding values are 3.94 and 227.98 seconds. At the 95th percentile, they are 9.20 and 874.39 seconds. The pooled UI/API ratios are therefore 81.4 at the median, 57.9 at the 90th percentile, and 95.1 at the 95th percentile. The associated seven-day block intervals are [33.3, 273.9], [33.3, 87.0], and [70.5, 127.7]. The same distinction appears at operationally meaningful thresholds. Among API-routed cancellations, 41.9% occur within 100 milliseconds and 77.4% occur within one second. The corresponding UI-associated shares are 0.29% and 24.3%. Figure IA.6 in the Internet Appendix plots the pooled latency distributions on a logarithmic clock.

Table II shows that this conclusion is not an artifact of order weighting. The equal-day ratios are 113.2 at the median, 87.0 at the 90th percentile, and 107.9 at the 95th percentile, and every HAC interval lies far above one. A distribution-free comparison yields the same ordering. For independent draws from the two pooled distributions, $\Pr(T_A < T_U) = 0.88$, with a moving-block interval of [0.86, 0.90].

Figure 1 reports $R(q)$ over the central 90% of both distributions. The ratio ranges from 27.1 to 219.4. More importantly, the lower edge of the simultaneous 95% band never falls below 8.73. The technological distinction is therefore not confined to a particular percentile, a chosen latency threshold, or a handful of extreme observations. It is a distribution-wide separation in the speed with which the two populations revise displayed liquidity. These measures are cancellation-conditioned

[6] For quantile q, $\delta_{dq} = \log Q_{U,d}(q) - \log Q_{A,d}(q)$ and $R^{\text{day}}(q) = \exp\left(D^{-1}\sum_{d=1}^{D}\delta_{dq}\right)$, with the long-run variance of the daily mean estimated by a seven-lag Newey–West estimator.

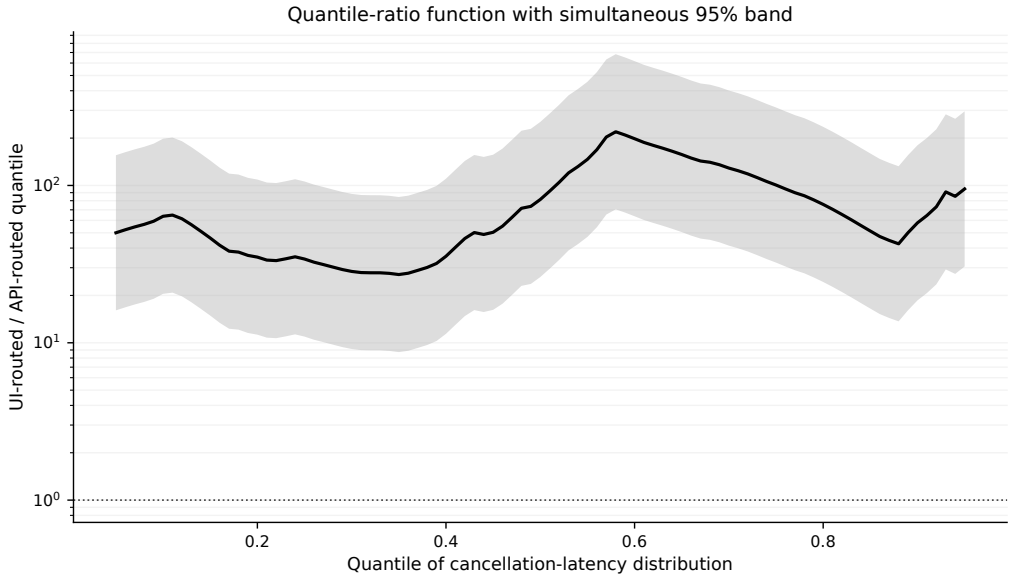


Figure 1. Cancellation-latency quantile ratios. The solid line plots $Q_U(q)/Q_A(q)$ for $q \in [0.05, 0.95]$. The shaded region is a simultaneous 95% max-deviation band constructed from 999 circular moving-block bootstrap replications with seven-day blocks. The horizontal reference line equals one.

revealed-action clocks: they record how quickly each population actually acts on its resting orders, not latent decision times, network latencies, or causal effects of technology.

Stability and label noise. The magnitude of the latency gap varies over time, but its sign and economic scale are stable. The pooled monthly median ratio exceeds 5.5 in every one of the 18 sample months. Across eleven subsamples formed by half-year, realized-volatility quartile, and trading-volume quartile, the median ratio always exceeds 23 and every block interval excludes one. The inference is also insensitive to the persistence imposed by the bootstrap. Increasing the block length from one to 28 days leaves the median ratio at 81.4 and keeps the lower confidence limit above 24.3.

Two contamination experiments (Figure IA.10 in the Internet Appendix) examine classification error. The first experiment reassigns the fastest UI-associated cancellations to the API cohort. This is the contamination pattern most directly implied by API users who omit `client_oid`. Removing these API-like observations from the UI distribution increases the estimated separation. The second experiment symmetrically mixes each route distribution with the other. Even when 20% of each distribution is replaced by observations from the other route, the median latency ratio remains 10.6 and the 90th-percentile ratio remains 3.82. These exercises cannot transform a route field into verified trader identity, but they show that the main validation result does not depend on a small set of perfectly classified orders.

Secondary diagnostics. Three secondary diagnostics corroborate the routing interpretation, although none is used as an identifying assumption. First, after normalizing for the large difference in mean intensity, UI-associated submissions are more clustered in time than API-routed submissions: a scale-invariant excess-dispersion statistic averages 1.74 for UI-associated flow and 0.43 for API-routed flow, and the difference of 1.31 has a seven-day block interval of [1.06, 1.57]. Second, the two routes have systematically different intraday profiles, with UI-associated activity relatively concentrated late in the UTC day and API-routed activity during 08:00–16:00 UTC; these differences are stable but modest relative to the latency evidence. Third, UI-associated orders bunch more at round dollar notionals, with standardized UI/API ratios of 1.50, 1.79, and 2.55 around \$100, \$500, and \$1,000 multiples, although the block intervals include one and the monthly estimates are unstable, so we treat bunching as suggestive corroboration. Internet Appendix Section IA.B.1 reports the formal definitions, the underlying distributions, and the intraday composition.

The route field thus separates populations that operate on economically different clocks and exhibit different submission patterns. The cancellation-latency result survives order weighting, equal-day weighting, simultaneous distributional inference, market-state splits, alternative assumptions about serial dependence, and substantial contamination of the route labels. In the remainder of the paper, human and machine refer to the UI-associated and API-routed populations established here. These labels should be understood as population-level proxies for trading infrastructure and automation, not as verified identities or claims about the intent behind any individual order.

The next subsection formalizes the market-by-order state variables used in the empirical analysis.

2.3. Market-by-Order Data and the Limit-Order-Book State

This subsection sets up the small set of objects used throughout the paper. They describe what the book looks like at any moment, how trading activity is measured, and how much a trade costs. All of them are standard objects in the empirical microstructure literature (Hasbrouck, 1991; Easley et al., 2016). The novelty of the paper is not in the measures themselves but in observing each of them separately for each trader population. Readers familiar with limit-order-book notation can skim to the maker/taker example at the end; the formal queue-level representation is spelled out in the footnote.[7] Replaying every message in sequence maintains a live copy of the full book, so the representation preserves queue position, order lifetime, maker cohort, taker cohort, and the difference between displayed liquidity that trades and displayed liquidity that is withdrawn. This is what a market-by-price feed, which reports only total depth at each price, cannot deliver, and it is the property on which the entire analysis rests: the labels ride on individual orders, not on

[7] Let event time be indexed by $n = 1, 2, \ldots$, with timestamp T_n. Immediately after event n, the bid and ask books are price-indexed FIFO queues $\mathcal{B}_n = \{\mathcal{Q}_n^b(p) : p \in \mathcal{P}\}$ and $\mathcal{A}_n = \{\mathcal{Q}_n^a(p) : p \in \mathcal{P}\}$, where $\mathcal{P}$ is the exchange tick grid and each queue element carries an order identifier, residual size, entry time, and route tag. Each message is a tuple $e_n = (T_n, \text{type}_n, \text{id}_n, \text{side}_n, p_n, x_n, g_n)$ with $g_n \in \{H, M\}$ the cohort proxy, and the book evolves by a deterministic transition map $(\mathcal{B}_n, \mathcal{A}_n) = \Phi(e_n; \mathcal{B}_{n-1}, \mathcal{A}_{n-1})$: an add inserts an order into its queue, a cancellation removes residual quantity, and a match removes quantity from the resting order at the execution price, walking through several queues if the marketable order exceeds depth at the first price.

aggregates.

At event n, let B_n and A_n denote the best bid and ask prices and Q_n^b and Q_n^a the depth displayed at them. Three top-of-book summaries appear throughout. The midprice $m_n = (A_n + B_n)/2$ is the market's reference price at that instant. The quoted spread $s_n = A_n - B_n$ is the cost of buying and immediately selling one unit. The top-of-book imbalance $I_n = (Q_n^b - Q_n^a)/(Q_n^b + Q_n^a)$ measures which side of the market is heavier at the best prices. It lies between minus one and plus one and is positive when more depth is displayed on the bid side. Dividing by total depth makes the measure comparable across quiet and busy periods. The intuition is a race between the two best queues. The midprice moves when one of them is exhausted, so when the bid queue is deep and the ask queue is thin, incoming selling has a thick cushion to chew through while incoming buying only has to eat a sliver, and the price is likely to tick up. Imbalance is in effect a measure of relative time to depletion, which is why it is one of the strongest short-horizon price predictors in electronic markets. It also represents revealed demand, since resting orders are commitments to trade at the current price. Both readings matter later in the paper, because a signal that everyone trades on is also a signal worth forging.

We use both calendar time and event time. Calendar-time intervals are left-closed, right-open windows $[t, t + \Delta)$, with Δ ranging from one second to one hour depending on the exercise. Event-time objects are indexed by the sequence of book updates, and are used for sign-memory, imbalance-response, and post-order event studies. Calendar time measures the intensity of market activity, while event time measures the strategic response of the book to the next k opportunities to update.

For an execution or parent marketable order j, let T_j be arrival time, P_j the volume-weighted execution price, X_j notional size, and $\epsilon_j = +1$ for buyer-initiated trades and $\epsilon_j = -1$ for seller-initiated trades. The feed reports directly which side initiated each trade, so trade signs are observed rather than inferred from price movements. We use signed slippage, effective spread, and markouts,

$$\mathrm{Slip}_j = \epsilon_j \frac{P_j - m_{T_j^-}}{m_{T_j^-}}, \tag{1}$$

$$\mathrm{EffSpr}_j = 2\epsilon_j \frac{P_j - m_{T_j^-}}{m_{T_j^-}}, \tag{2}$$

$$\mathrm{MO}_j(\tau) = \epsilon_j \frac{m_{T_j+\tau} - P_j}{m_{T_j^-}}. \tag{3}$$

In words, slippage is the gap between the price a trade actually pays and the midquote at the moment it arrives, the effective spread is twice that gap, and the markout asks what happened next, namely how far the price moved for or against the trade within τ seconds. Slippage and effective spread measure the immediate price of demanding liquidity, while the markout measures the ex post information content of the trade. A negative markout for a taker is an immediacy cost not followed by favorable price movement; a positive markout is a trade in the direction of

subsequent price discovery.

A *maker* order is a limit order that rests in the book and supplies displayed liquidity. A *taker* order is an order that executes immediately against resting liquidity. For example, if the best bid is 100 with two BTC displayed and the best ask is 101 with one BTC displayed, a one-BTC limit sell at 101 is a maker order if it rests at the ask, while a one-BTC market buy is a taker order because it removes the ask. Slippage is the signed distance between the execution price and the midquote just before the trade. In the same example, the midquote is 100.50, so a market buy at 101 pays slippage of 0.50 per BTC before fees. Slippage matters because it is the first dollar measure of the cost of speed. It is the price paid for immediacy before one asks whether the trade was informed or whether the price later moved in the same direction.

2.4. Stylized Facts of the BTC-USD Book

Before using the reconstructed book, we verify that it reproduces the stylized facts that decades of research have established for electronic limit-order markets. This serves two purposes. It validates the pipeline, since a reconstruction error would show up as a distorted version of well-known regularities, and it establishes the baseline against which the human-machine results are interpreted. Throughout, clock-time figures use Coordinated Universal Time, large marketable orders are cohort-specific notional sizes above the empirical 95th percentile, and extreme imbalance means I_n above the 95th or below the 5th percentile of its rolling-day distribution.[8] Four facts carry the later analysis directly and are presented where the analysis uses them. Displayed depth is small relative to the trading it must absorb (Section 4), imbalance forecasts both price pressure and liquidity depletion (Section 4.2), large orders reprice immediacy more than they move the midpoint (Section 4.3.3), and signed flow has lead-lag structure across the two populations (Section 3.3).

Comparison with canonical electronic markets. The canonical-facts summary in Internet Appendix Section IA.A.2 records the comparison with the canonical limit-order-book facts point by point. The broad conclusion is that BTC-USD strongly resembles textbook electronic markets in event space. Order signs have memory, liquidity is concentrated near the spread, order sizes and queue sizes are heavy-tailed, and displayed depth is small relative to traded volume. The key difference is in calendar time and participant heterogeneity. The crypto market does not inherit a national-market U-shape but has a global participation cycle. These differences aside, the resemblance is itself a finding. A market with no opening bell, no designated market makers, and no consolidated tape converges to the same event-time regularities as regulated equity markets, which suggests that these regularities are produced by the mechanics of price-time priority itself rather than by regulation or institutional structure. What crypto adds is the ability to see the participants. The aggregate facts conceal an important cohort dimension, because manual and automated traders populate different regions of the state space.

[8] Profile figures report, for each clock bin, the cross-day mean with interquartile bands across days, and event-time figures report means of the transformed outcome at fixed offsets after the conditioning events. Internet Appendix Section IA.A.1 records the estimator notation and varies these definitions in robustness checks.

Three measurement principles follow for the remainder of the paper. Liquidity must be measured conditionally, because the relevant state is the joint configuration of spread, imbalance, recent flow, and replenishment rather than an average. Event-time and clock-time effects must be separated, because activity is periodic on the calendar clock while order-book reactions occur over the next few events. And the human-machine labels are not cosmetic controls. If manual and automated traders arrive in different states, use different order sizes, and react on different clocks, then aggregate order flow is a mixture of financially distinct objects. Section 3 measures these interactions directly.

3. Interactions of Humans and Machines

The previous section treats the book as a single stochastic object, but the evidence in this section shows that this aggregation hides the most important economics of the market. Human and automated traders do not merely submit different amounts of the same order flow. They arrive in different states, use different order-size distributions, update passive liquidity at different frequencies, and leave different footprints in future depth and prices. The relevant distinction is therefore not a loose comparison between slow and fast traders. It is a decomposition of the liquidity-production function into the agents who keep queues alive, the agents who consume those queues, and the agents whose flow is informative for subsequent book conditions.

3.1. Cohort-Specific Order-Book Primitives

The building blocks are the Section 2.3 objects computed separately for each population. For each interval t and cohort $g \in \{H, M\}$, with H the manual/retail cohort and M the automated/API cohort, the signed marketable notional $q_{g,t} = \sum_{j:T_j \in [t,t+\Delta),\ g_j=g} \epsilon_j X_j$ measures how many dollars cohort g buys minus sells through immediately executing orders, and $L_{g,t}(\eta)$ and $C_{g,t}(\eta)$ are the cohort's passive additions and cancellations within the price band η. The cohort-specific contribution to displayed liquidity is

$$\ell_{g,t}(\eta) = L_{g,t}(\eta) - C_{g,t}(\eta) - M_{g,t}, \tag{4}$$

where $M_{g,t}$ is marketable notional submitted by cohort g. The sign convention is from the perspective of the book: passive additions add tradeable supply, while cancellations and marketable executions remove supply from the queues that later traders face. In short, $\ell_{g,t}$ answers whether cohort g made the book deeper or thinner during bar t.

When an execution occurs, both sides of the trade are observed, and the resulting cross-cohort execution matrix is used in the cross-exchange analysis of Section 5.4. Let $X^{\text{take,make}}_{g,h,t}$ denote notional in interval t in which the taker belongs to cohort g and the resting maker belongs to cohort h. The maker-taker matrix is

$$\mathbf{X}_t = \begin{pmatrix} X^{\text{take,make}}_{H,H,t} & X^{\text{take,make}}_{H,M,t} \\ X^{\text{take,make}}_{M,H,t} & X^{\text{take,make}}_{M,M,t} \end{pmatrix}. \tag{5}$$

The off-diagonal cells are direct measures of human-machine interaction. The cell $X^{\text{take,make}}_{H,M,t}$ records

human demand for immediacy supplied by machine quotes, while $X^{\text{take,make}}_{M,H,t}$ records machine demand for immediacy against human displayed interest.

3.2. Order Placement, Size, and the Geography of Human-Machine Liquidity

Passive placement gives the first view of the interaction. Both cohorts place large mass close to the best quotes, yet the two distributions are not the same (Figure IA.3 in the Internet Appendix). Automated liquidity is spread more continuously, both across nearby price levels and across small to intermediate order sizes. Human passive liquidity is more concentrated at round sizes. The pattern is consistent with different control problems. Automated systems continuously trade off queue priority, adverse selection, and inventory, while manual traders use coarser price and size choices that reflect displayed screens, round-number preferences, and portfolio constraints.

The second margin is aggressive trading. We condition on recent signed marketable flow, $R^{(a)}_t = \sum_{j:T_j\in(t-a,t]} \epsilon_j X_j$, which measures net buying pressure over the preceding a seconds. Figure 2 plots cohort-normalized marketable order size against $R^{(a)}_t$ for ten-, thirty-, and sixty-second lookbacks. The plotted curves are conditional bin means of order size relative to the cohort's own average, so values are centered near one and the slopes show how strongly size responds to the recent flow state, $\widehat{m}_{g,a}(b) = |\mathcal{I}_{g,a,b}|^{-1} \sum_{j\in\mathcal{I}_{g,a,b}} X_j/\bar{X}_g$, where $\mathcal{I}_{g,a,b}$ contains marketable orders from cohort g whose recent flow lies in bin b. Human order size falls sharply as recent flow moves from selling pressure toward balance and then rises only modestly with buying pressure. Automated order size is larger throughout and varies more symmetrically around zero. The economically relevant fact is therefore not an isolated tail observation but a persistent difference in state dependence. Human liquidity demand arrives in discrete, directionally conditioned bursts, whereas automated trading remains closer to continuous queue and inventory management. This distinction matters for execution costs because the human disadvantage reflects not only the spread paid conditional on trading, but also the states in which human traders choose to demand immediacy.

3.3. Order-Flow Memory and Cross-Population Lead-Lag

A durable fact in limit-order-book markets is that marketable-order signs are persistent. Let ϵ_i be the sign of the ith marketable order and let $Z_i = \epsilon_i X_i$ be signed notional. The event-time autocorrelations are

$$\rho_\epsilon(k) = \frac{\text{Cov}(\epsilon_i, \epsilon_{i+k})}{\text{Var}(\epsilon_i)}, \tag{6}$$

$$\rho_Z(k) = \frac{\text{Cov}(Z_i, Z_{i+k})}{\text{Var}(Z_i)}. \tag{7}$$

Figure 3 shows that signs remain positively autocorrelated over long event lags and that sign persistence is stronger than signed-notional persistence. This distinction matters because persistent direction can be generated by order splitting, herding, or common reaction to information, while signed-notional persistence is dampened by size variation and depth constraints. The right panel

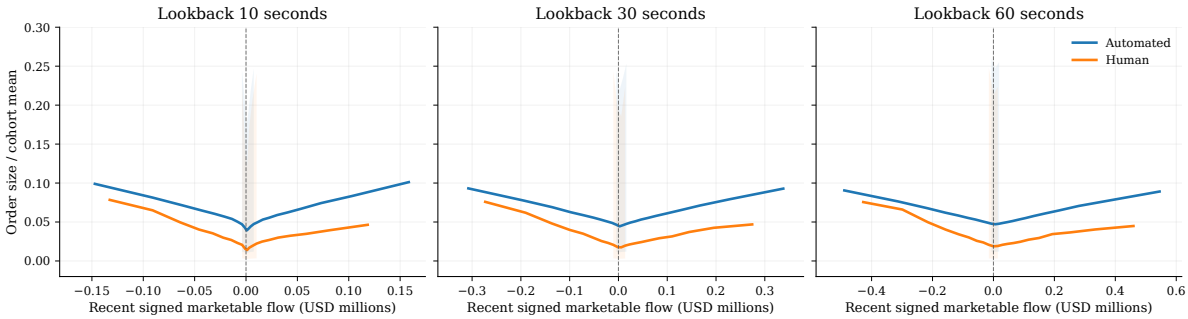


Figure 2. Cohort-specific marketable order size and recent signed flow. Unknown-route orders are excluded, the axis covers the central 90 percent of the recent-flow distribution, and confidence bands are shown where they are narrow enough to read. Each panel reports conditional means after removing unknown cohort labels and restricting recent flow to the central 90 percent of its pooled distribution. Order size is normalized by the corresponding cohort mean. Confidence bands are retained only in central bins where they remain visually informative. Human order size is more strongly conditioned on recent directional flow, whereas automated order size is larger and more symmetric around balanced flow.

studies cross-population lead-lag behavior. For cohort-specific sign or signed-notional series X_t^H and X_t^M on a one-second grid, define

$$\Gamma_{HM}(\ell) = \mathrm{Corr}(X_t^H, X_{t+\ell}^M). \tag{8}$$

Positive ℓ means that automated flow follows human flow. The cross-correlation is concentrated around the event time with a sharp peak at contemporaneous and near-contemporaneous lags. Human and automated trading are therefore synchronized around short-lived states rather than independent streams passing through the book. Even modest correlations are financially meaningful at this frequency because a small increase in the probability of same-direction order flow can materially raise adverse-selection risk when top-of-book depth is thin.

The important fact for the paper is that human and machine order flow is not independent noise. It is synchronized around short-lived book states, which raises adverse-selection costs precisely when top-of-book depth is thin. The direction of the lead is itself striking. On average, automated flow follows human flow at short lags, so the population most able to react is reacting to the population least able to. This ordering is not an artifact of humans somehow anticipating events, as the externally timed evidence of Section 5 shows machines responding first when information arrives from outside the venue. The figure connects the canonical long-memory fact to the human-machine mechanism that this section measures.

The stylized facts imply that the price of immediacy is state dependent in a way that loads on the human cohort: human marketable orders arrive disproportionately after directional flow and in volatile states, exactly when liquidity suppliers reprice the option embedded in their quotes, as the event studies of Section 4.3.3 show. A transient book state is therefore costly not only when it moves the midpoint, but also when it changes which traders demand immediacy, which traders

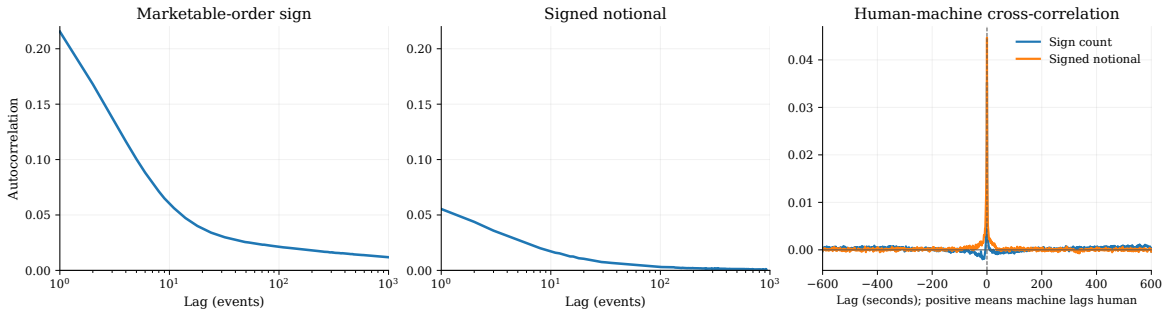


Figure 3. Order-flow memory and human-automated lead-lag. The first two panels plot the event-time autocorrelations of marketable-order signs and signed notional defined in Equations (6) and (7), on a vertical range ending at 0.20, which makes their slow decay over hundreds of events visible. The third panel plots the one-second cross-correlation between human and automated flow defined in Equation (8). Positive lags mean that automated flow follows human flow.

supply it, and at what price, the channel that Section 6 prices in dollars.

3.4. A Two-Population Kyle Model

Why do trades move prices at all? In the canonical models of Kyle (1985) and Glosten and Milgrom (1985), liquidity suppliers cannot tell which orders carry information, so they protect themselves by moving the price against every order; this protection charge is adverse selection. We extend the logic by letting the two populations differ. The model in this subsection is deliberately simple. It is a projection argument in the spirit of Kyle rather than a new equilibrium theory, and its value is that it disciplines measurement and generates the out-of-sample test of Section 3.5. Let the fundamental value after the trading interval be v, with prior mean p_0 and variance σ_v^2, and decompose aggregate signed flow into the vector $\boldsymbol{q} = (q_H, q_M)^\top$. A competitive risk-neutral liquidity supplier sets the price equal to the conditional expectation of value, $p_1 = \mathbb{E}[v \mid \boldsymbol{q}]$. Under joint Gaussianity, the conditional expectation is linear,

$$p_1 = p_0 + \boldsymbol{\lambda}^\top \boldsymbol{q}, \qquad \boldsymbol{\lambda} = \Sigma_{qq}^{-1} \Sigma_{qv}, \tag{9}$$

where $\Sigma_{qq} = \mathbb{V}(\boldsymbol{q})$ and $\Sigma_{qv} = \mathbb{C}\mathrm{ov}(\boldsymbol{q}, v)$. Equation (9) is the two-population analogue of the Kyle impact coefficient: in words, $\boldsymbol{\lambda}$ measures how much prices move per dollar of net buying from each population, given what the other population is doing. The impact assigned to machine flow is the covariance of machine flow with value after projecting out human flow, and the human coefficient is defined symmetrically. A one-population model is therefore valid only when the two flows have the same informational content and the same state dependence, which the preceding evidence rejects.

A stark benchmark makes the economic content transparent. Suppose human flow is pure liquidity trading, $q_H = u_H$, while machine flow tracks value, $q_M = \beta(v - p_0) + u_M$, with u_H and u_M zero-mean components independent of v and of each other. Then $\lambda_H = 0$ and $\lambda_M =$

$\beta\sigma_v^2/(\beta^2\sigma_v^2+\sigma_M^2)$: all price impact attaches to the informed population. The value of the model is not the literal prediction $\lambda_H = 0$, since human flow can be informative around news or herding episodes and automated flow can be inventory driven. The value is that λ_H and λ_M are disciplined measures of residual price impact after conditioning on the other cohort.

The same decomposition explains why execution costs are distributional. Suppose a human taker trades size x when the book has spread s and labeled impact vector $\boldsymbol{\lambda}$. The expected signed cost can be written as

$$\mathbb{E}[\text{Cost}_H(x) \mid \mathcal{F}_{t^-}] \approx \frac{s_t}{2m_t}x + x\,\mathbb{E}\Big[\boldsymbol{\lambda}_t^\top \boldsymbol{q}_t \mid \mathcal{F}_{t^-}, g = H\Big] + \text{DepthCost}_t(x). \tag{10}$$

In words, a human trade pays three costs: half the quoted spread, an impact charge that is largest when informative flow is active, and the extra cost of eating through multiple price levels when the order is large relative to displayed depth. Human-machine heterogeneity matters because the conditional distribution of $\boldsymbol{q}_t$, s_t, and $\text{DepthCost}_t(x)$ differs by cohort and by arrival state: the states in which human orders concentrate are exactly the states in which all three components are elevated. This is the channel that Sections 4 through 6 trace from mechanism to dollars.

Why the decomposition matters and what we do not estimate. The economic point of the decomposition is an omitted-variable argument. A price-impact coefficient estimated on aggregate flow blends two different objects, because the same dollar of net buying means something different when a machine submits it than when a person does. When the two flows co-move, an aggregate coefficient loads on the covariance of both flows with returns, so collapsing q_H and q_M into one scalar discards precisely the variation needed to distinguish liquidity demand from informed or defensive quote revision. We deliberately do not report estimated impact coefficients in this paper. Reading a fitted slope as price impact requires signed flow to be predetermined relative to the return innovation, a strong condition in a high-frequency book where quotes and marketable orders are chosen almost simultaneously, and a defensible estimate therefore requires lagged and vector-autoregressive designs on a common construction that are part of the ongoing analysis rather than of this draft.[9] The framework's testable content is instead delivered in two ways that do not require the coefficients. The forecasting exercise of Section 3.5 uses the labeled flows themselves and asks whether they carry incremental information about the future book, and the event-time evidence of Sections 4 and 5 asks which population moves first when information arrives, the across-venue counterpart of the within-venue impact question.

3.5. Forecasting Future Liquidity with Labeled Flow

The labeled flow variables have direct predictive content for future liquidity. If the human-machine distinction captured nothing real, it would not help forecast the future state of the book.

[9] Where an impact slope is needed as a per-event diagnostic in Section 6, it is estimated locally as described in Internet Appendix Section IA.G.1, and it is used there only to rank events, never as a headline estimate.

The forecast target answers a question every trader cares about. Over the next H seconds, is displayed liquidity near the best prices about to be replenished or about to disappear? Trading volume alone cannot answer it, because volume records only executed demand, while displayed supply can also be added or quietly withdrawn before anyone trades. We therefore define, for a time interval $[t, t+\Delta)$, the dollar value of passive limit-order additions $L_t(\eta)$, cancellations $C_t(\eta)$, and marketable executions M_t, with additions and cancellations counted within a relative distance η of the midpoint,[10] and use the forward net-liquidity target, which cumulates the same additions, cancellations, and executions as Equation (4), pooled across cohorts and summed forward,

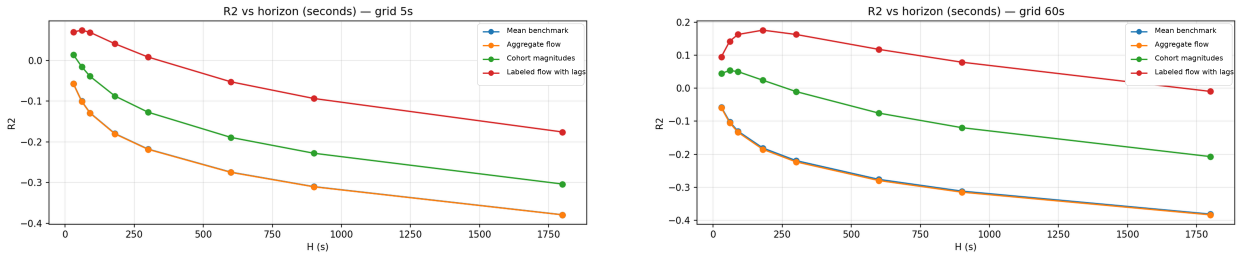
$$y_{t,H}(\eta) = \sum_{h=1}^{H/\Delta} L_{t+h\Delta}(\eta) - \sum_{h=1}^{H/\Delta} C_{t+h\Delta}(\eta) - \sum_{h=1}^{H/\Delta} M_{t+h\Delta}. \tag{11}$$

This object also serves as the outcome variable of the event studies in Section 4. From one-second aggregates we construct the cohort signed flows $q_{H,t}$ and $q_{M,t}$ (buys minus sells) and their multi-horizon histories. The comparison is intentionally parsimonious, with four nested models: an intercept-only benchmark, an aggregate-flow benchmark using current and lagged flow $q_{H,t} + q_{M,t}$, a magnitude benchmark using $|q_{H,t-\ell}|$ and $|q_{M,t-\ell}|$ to ask whether cohort activity matters even after sign is discarded, and the full labeled-flow model using $q_{H,t-\ell}$, $q_{M,t-\ell}$, and their difference for lags $\ell = 0, \ldots, L$. All models are ridge regressions on fixed calendar grids.[11] All four models use the same ridge estimator, the same grids, and the same fixed penalty, so differences in forecast accuracy can only come from the information in the features, not from differences in estimation technology. Test blocks are held out in calendar time and purged by at least the maximum forecast horizon on both sides, so no training label overlaps a test label and the reported R^2 measures the share of variation explained on data the model never saw. The R^2 is measured against the mean forecast of the training block, so a value of zero means no better than a constant and a negative value means worse than a constant.

Figure 4 reports out-of-sample R^2 across horizons on five-second and sixty-second grids for the four models. The full lagged labeled-flow design dominates the aggregate-flow controls. On the five-second grid the advantage is largest at short horizons and decays as the forecast horizon lengthens. On the sixty-second grid the labeled design remains superior and peaks at intermediate horizons. Two readings accompany the levels. At the longest horizons the out-of-sample R^2 can fall toward or below zero, which says that no model beats a constant far into the future, as expected for liquidity at these frequencies. The economically meaningful result is therefore the ranking rather than the level. The aggregate-flow model performs about as well as the mean benchmark, while

[10] Formally, $L_t(\eta) = \sum_{n:T_n \in [t,t+\Delta)} p_n x_n \, \mathbf{1}\{\text{add}_n,\ |p_n/m_{n^-} - 1| \le \eta\}$, $C_t(\eta)$ is defined identically for cancellations, and $M_t = \sum_{n:T_n \in [t,t+\Delta)} P_n x_n \, \mathbf{1}\{\text{match}_n\}$.

[11] The estimator is $\widehat{\beta}_h(\kappa) = \arg\min_\beta \sum_{t \in \mathcal{T}_{\text{train}}} \{y_{t,h} - X_t'\beta\}^2 + \kappa \|\beta\|_2^2$ with the predictors standardized and the penalty fixed at $\kappa = 1$ for all horizons and specifications, which avoids a tuning layer that could exploit serial dependence, and the reported statistic is the out-of-sample $R^2_{OS}(h) = 1 - \sum_{t \in \mathcal{T}_{\text{test}}} (y_{t,h} - X_t'\widehat{\beta}_h)^2 / \sum_{t \in \mathcal{T}_{\text{test}}} (y_{t,h} - \bar{y}_{\text{train},h})^2$. The five-second and sixty-second grids use the same construction; only the spacing of t and the feasible lag set differ.



(a) Five-second grid

(b) Sixty-second grid

Figure 4. Forecasting future net liquidity with labeled order-flow histories. The lagged labeled-flow specification uses cohort-specific signed flows and their differences. Aggregate-flow controls perform poorly relative to the labeled design, showing that human-machine labels contain incremental information about future liquidity.

the labeled model does not, so future book conditions are forecast by the decomposition of pressure across order-entry technology and by its recent path, not by aggregate buying and selling pressure alone.

3.6. Interpretation

The human-machine evidence delivers a simple message with nontrivial implications. Machines and humans do not just differ in speed. They differ in placement, timing, size discreteness, reaction speed, and residual price impact. Machines keep liquidity close to the spread and revise it rapidly as the state changes. Humans submit more discrete orders and are more likely to demand immediacy after directional flow. Technological heterogeneity is therefore a distributional property of liquidity rather than a background feature of the trading protocol.

The two-population Kyle model is the organizing device for this interpretation. In a one-population model, the market maker observes only aggregate signed flow and assigns a single price-impact coefficient to it. In the data, the same aggregate flow can be generated by human buying against machine quotes, machine buying against human quotes, same-direction trading by both groups, or offsetting cross-cohort flows. These states imply different future depth and different execution costs. The forecasting evidence confirms that the labels are financially informative because future liquidity is predicted by the labeled path of flow rather than by the sign or magnitude of aggregate flow alone.

The next section asks whether some of the states to which humans respond are strategic. The evidence above does not require manipulation, but it establishes why manipulation can matter. Displayed depth is fragile, best-quote liquidity is rapidly revised, order signs have memory, and human liquidity demand is state dependent. In such a market a fleeting change in displayed depth or a short burst of aggressive flow can alter not only prices, but also the allocation of execution costs across trader populations.

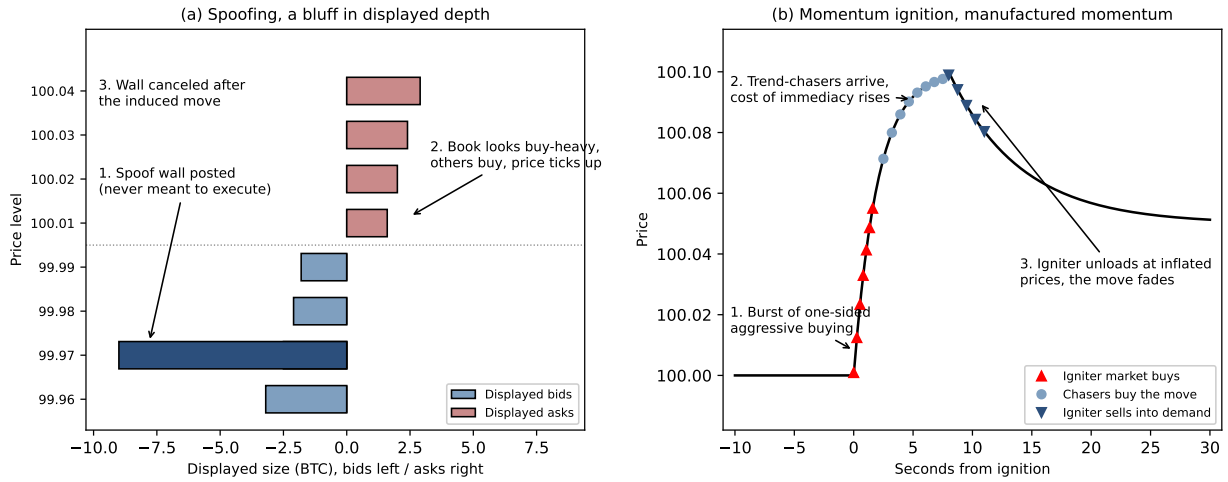


Figure 5. Stylized illustration of the two manipulation mechanisms. The example is artificial and drawn for intuition, not from data. Panel (a) shows spoofing. A large bid wall is posted below the best bid, the displayed book suddenly looks buy-heavy, other traders buy and the price ticks up, and the wall is canceled once the move has happened. Panel (b) shows momentum ignition. A burst of aggressive buying starts a price move, trend-chasing traders and algorithms buy into it and push the cost of immediacy up, and the igniter sells to them at inflated prices before the move fades. The empirical designs of this section detect the footprints of these patterns in the actual order book.

4. Market Manipulation

If a speed advantage converts public information into private advantage anywhere, it should be in the order-book states that surveillance teams monitor most closely. In the language of the paper's thesis, these are the states where human predictability meets machine speed. The preceding sections identify the conditions under which such states have economic force: displayed depth is small at the horizons where immediacy (trading now rather than waiting) is demanded, quotes are revised quickly, and human liquidity demand is state dependent (Section 3). We study the two classic mechanisms that operate through those conditions, and both are easy to state in plain words. A spoofer posts a large displayed order he never intends to execute. The wall is a bluff. It fakes supply or demand, moves the beliefs and trades of everyone watching the book, and is canceled once it has worked. An igniter fires a rapid burst of aggressive one-sided orders to start a price move, pulls trend-chasing traders and algorithms into the move, and trades against their demand at improved prices. Both strategies profit from the predictable reactions of other traders, which is why a market with a slow and predictable population is fertile ground for them. Figure 5 illustrates both mechanisms with a stylized example before the formal analysis begins.

We estimate episodes of likely manipulation from order-book states implied by surveillance guidance and microstructure models. These are economic footprints in the data feed, not legal findings about individual traders, since we observe neither intent nor accounts nor communications.

Spoofing begins with the cross-section of displayed depth and momentum ignition begins with

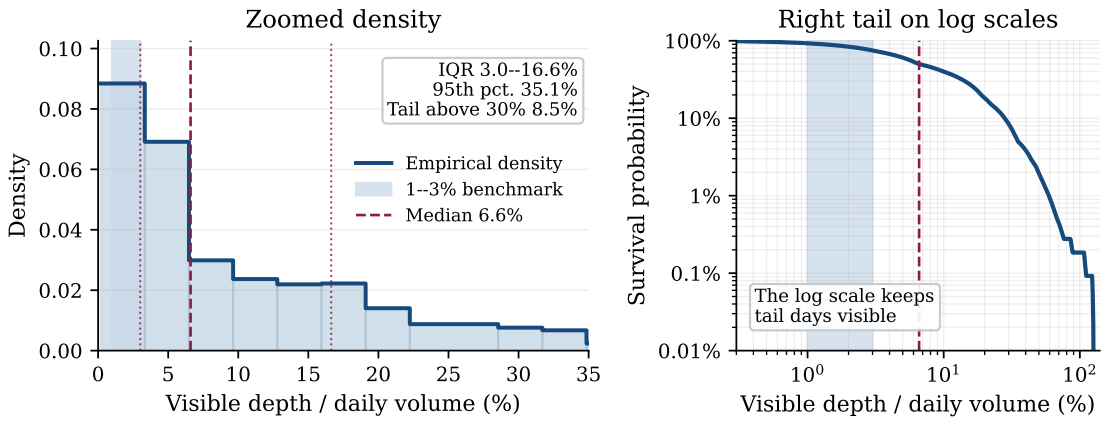


Figure 6. Visible depth within one percent of the midpoint relative to daily volume. The left panel reports the zoomed density and the right panel reports the survival function on logarithmic axes. The benchmark band marks the one-to-three percent range discussed in the limit-order-book literature. The figure separates the dense central region from the right tail and shows why static visible depth must be interpreted together with cancellation and replenishment.

a time-series burst of marketable order flow, but their common feature is the endogenous response of other traders. A spoof is valuable only if other traders react to displayed depth, and momentum ignition is valuable only if other traders chase the induced move. The cohort decomposition therefore turns a surveillance problem into a testable economic mechanism.

Why can fleeting book states matter at all in a deep market? The answer is a fact about scarcity, and it is the first of the stylized facts that the reconstructed book reproduces. For each day we aggregate the displayed dollar depth posted within one percent of the midpoint and compare it with the volume that trades that day. Figure 6 reports the result. Visible depth close enough to trade is typically one to three percent of daily turnover, in line with the benchmark range documented for equity order books, and the benchmark is useful because it turns a visual book into a capacity statistic. Even when the book appears deep in dollars, the visible liquidity close enough to trade is a small fraction of daily turnover. This is the canonical limit-order-book fact that makes the rest of the paper economically possible. A small amount of displayed depth can determine execution terms for a large amount of urgent order flow, especially when traders demand immediacy during volatile states. The fact is easy to miss on a trading screen that always shows a full ladder of quotes; measured against the volume that must trade through it, the visible cushion is thin, and a thin cushion can be moved, withdrawn, or manufactured within seconds. Every mechanism studied later operates through this scarcity.

4.1. The Order Book as a Depth Curve

The purpose of this subsection is a parsimonious representation of the entire displayed book. Both detectors below need a notion of what a normal book looks like, and the whole book at any moment is a high-dimensional object with thousands of price levels on each side. We compress it into two curves, how much is displayed at each distance from the best quotes and which side is heavier, and we summarize each curve by a small number of interpretable factors plus a localized residual. Different manipulation mechanisms then show up as movements of different state variables, which is what makes the representation useful. The construction adapts the framework of Filipović, Pelger, and Ye, abbreviated FPY throughout, which combines nonparametric curve estimation in a reproducing-kernel Hilbert space with smoothness-ordered factor extraction and was developed for interest-rate curves (Filipović et al., 2022, 2024). The analogy is statistical rather than based on no-arbitrage. A limit order book can be jagged for perfectly legitimate reasons, including queue priority, tick size, hidden liquidity, and inventory constraints, and the paper does not treat those kinks as infeasible. Smoothing instead acts as a filter. It separates the persistent displayed-supply state, which is what liquidity providers actually condition on, from millisecond quote flicker, and the local residuals are retained deliberately because wall-like distortions live precisely there.

Let $0 = d_0 < d_1 < \cdots < d_J$ denote depth-bin endpoints measured in ticks from the relevant best quote, let $u_j = \log(1 + d_j)$ be the log-depth coordinate, and let $b_t(u_j)$ and $a_t(u_j)$ be bid and ask dollar-depth densities on the log-depth grid, obtained from cumulative depth within d_j ticks. We work with liquidity and imbalance density curves

$$\ell_t(u_j) = b_t(u_j) + a_t(u_j), \qquad \iota_t(u_j) = b_t(u_j) - a_t(u_j), \tag{12}$$

transformed as

$$x_t^L(u_j) = \log\{1 + \ell_t(u_j)\}, \qquad x_t^I(u_j) = \operatorname{asinh}\{\iota_t(u_j)\}, \tag{13}$$

The liquidity curve ℓ_t records how much is displayed at each distance, and the imbalance curve ι_t records which side is heavier. The two transformations are standard variance stabilizers rather than arbitrary choices. Depth is heavy-tailed, so the logarithm compresses it, and imbalance is signed, so it uses the inverse hyperbolic sine. Each curve is then smoothed with a weighted kernel-ridge estimator whose penalty makes smooth shapes cheap and jagged shapes expensive, with an exponential tilt that concentrates accuracy near the touch, meaning close to the best quotes, where displayed depth matters most for immediate execution. The smoothing degree is chosen by blocked cross-validation, meaning the penalty is selected by how well the fitted curve predicts held-out depth observations in separate time blocks, and the selected values lie on broad flat regions of the validation surface, so nothing is knife-edge in the tuning. The number of factors is chosen afterwards from reconstruction criteria on the smoothed panel, and the two choices interact, since heavier smoothing removes rough directions and lowers the number of factors needed. There is no single optimal filter, because a filter by definition discards structure. A filter is good relative to a purpose, and ours must do three things. It must pass through the persistent displayed-supply state

that liquidity providers condition on, block millisecond quote flicker, and leave localized walls in the residual where the detector looks for them. The diagnostics in the Internet Appendix check exactly these three properties. Smoothing does not erase spoofing evidence, but rather, it defines the normal smooth state against which a localized wall is measured. A principal-components analysis on the raw unsmoothed depth bins would instead be dominated by quote flicker and grid jaggedness, which is why the filter comes first. Internet Appendix Sections IA.D.1 and IA.D.2 give the estimator, the exact bid-ask recovery mapping, and the error-propagation bounds, while Figure IA.16 shows the cross-validation diagnostics.

The smoothed curves are stacked across time and decomposed into factors. The idea in words is that each snapshot's curve is expressed as a combination of a few common shapes, and the score $F_{t,k} = \langle \widehat{x}_t, v_k \rangle$ measures how much of shape k today's curve contains. We retain two factors per curve. In the curve-PCA benchmark these two factors explain 99.96 percent of the variation in the smoothed liquidity curve and 99.70 percent in the smoothed imbalance curve, with the first factor accounting for 94.49 and 90.78 percent and the second adding 5.47 and 8.92 percent (Table IA.IV in the Internet Appendix). For each of the liquidity and the imbalance curve, the first factor is a level movement of the whole curve and the second is a slope that shifts weight between the touch and deeper levels. Together with the high-order residual share, the part that the smooth shapes cannot explain and where localized walls appear, this gives five interpretable state variables, the liquidity level and slope, the imbalance level and slope, and the residual. We compute the scores two ways, by curve PCA as a high-fit benchmark and by a conservative shrinkage specification in the smoothness-ordered kernel eigenbasis that suppresses rough directions easily generated by fleeting quote traffic. The event-time figures in this paper use the shrinkage scores, and the PCA benchmark serves to show how compressible the book is, with both giving the same qualitative picture. Figure 7 shows the estimated shapes. The leading liquidity and imbalance loadings are smooth and low order, the first close to a level movement and the second shifting weight between near-touch and deeper depth, so the factor names are earned by the data rather than imposed.

Internet Appendix Section IA.D.4 records the factor interpretation in baseline and manipulation states. The representation returns in Section 4.4, where it shows that spoofing, ignition, and news each move different parts of the book state.

4.2. Spoofing

Spoofing-like activity exploits the informational role of displayed depth. A large displayed sell wall reads as "someone wants to sell a lot," so buyers hold back, sellers press ahead, and algorithms revise fair value downward. A large displayed bid wall has the symmetric effect. If the wall is canceled after the induced movement, the book has been used as a signaling device rather than as a durable supply of liquidity. The bluff is cheap because a resting order costs nothing to post and can be withdrawn in milliseconds.

Why does forging this particular signal pay? A second stylized fact supplies the answer. We condition on days' most extreme imbalance states (above the daily 95th or below the 5th percentile

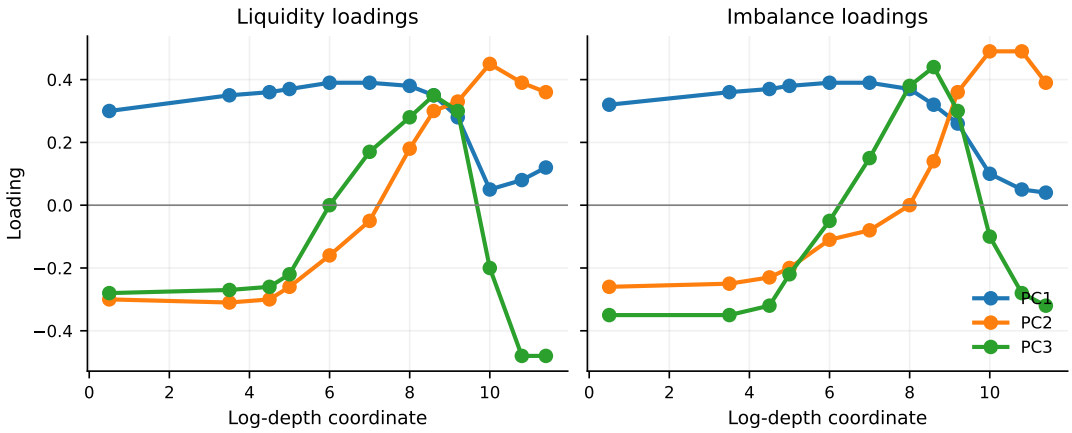


Figure 7. Empirical FPY loadings across depth. The leading liquidity and imbalance loadings have smooth low-order shapes across the log-depth grid. The first component is close to a level movement, the second shifts mass between near and deeper depth, and the third captures curvature. This geometry is the order-book analogue of level-slope-curvature factor structures in term-structure models.

of I_n) and trace the average future paths of the midpoint and of top-of-book depth in event time.[12] Figure 8 shows that extreme imbalance forecasts the direction of the midpoint and the subsequent decline in best-quote liquidity. The same state variable therefore predicts both price pressure and worsening execution terms, and this dual predictability is what gives the fact its force. A trader who reads the imbalance quickly enough sees both the next price move and the coming deterioration in trading conditions before they arrive, while a trader who cannot is systematically late to both. It is also why imbalance is the natural raw material for spoofing. A manufactured wall distorts exactly the signal that best predicts short-run prices and liquidity.

Any credible spoofing measure must solve one problem above all: large one-sided displayed depth is often legitimate market making, so the measure must ask how unusual the book is relative to what legitimate trading produces in the same market conditions. Our design does exactly this, following the logic of surveillance practice and the spoofing literature (Tao et al., 2022; Cartea et al., 2023): score the displayed book against a benchmark matched to the current state, and count only the excess. One caveat travels with the results and is stated where they are read: the forecast target includes the candidate depth's own lifecycle, so the projections combine the candidate's removal with responses elsewhere in the book. The main text uses a simple imbalance statistic built on the top-of-book imbalance of Section 2.3. Because financially meaningful depth often sits several ticks away from the best quotes in this market, we aggregate dollar depth over the first K ticks on each

[12] Formally, with $D_n^{\mathrm{BBO}} = A_n Q_n^a + B_n Q_n^b$ top-of-book notional depth and $\mathcal{E}_+ = \{n : I_n \geq q_{.95,d(n)}\}$, the plotted paths are $\widehat{\mu}_m^+(k) = |\mathcal{E}_+|^{-1} \sum_{n\in\mathcal{E}_+} (m_{n+k} - m_n)/m_n$ and $\widehat{\mu}_D^+(k) = |\mathcal{E}_+|^{-1} \sum_{n\in\mathcal{E}_+} (D_{n+k}^{\mathrm{BBO}} - D_n^{\mathrm{BBO}})/D_n^{\mathrm{BBO}}$; the negative paths use $\mathcal{E}_- = \{n : I_n \leq q_{.05,d(n)}\}$.

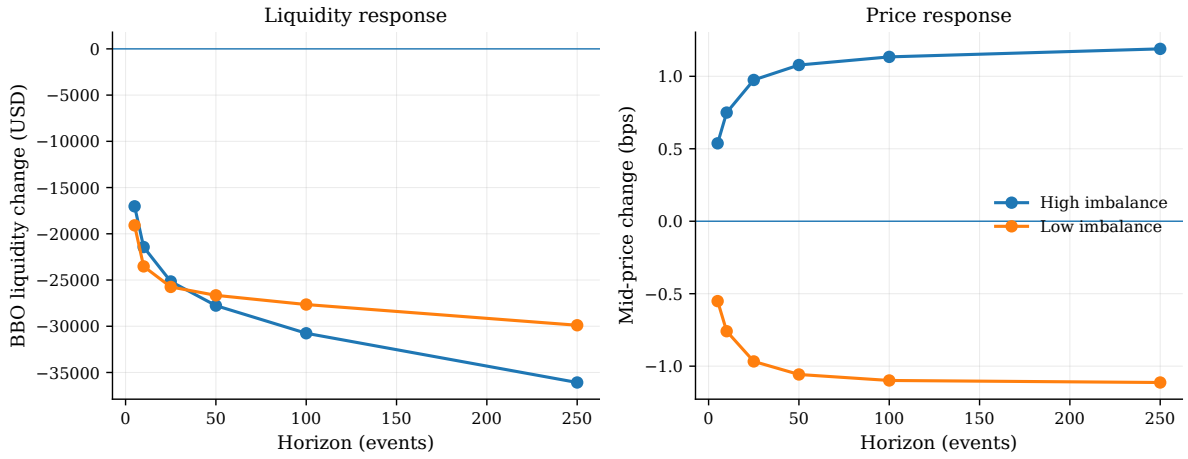


Figure 8. Event-time response to extreme imbalance. The figure plots average future top-of-book liquidity changes and midprice changes conditional on high and low imbalance states. Extreme imbalance predicts price movements in the direction of the imbalance and a subsequent reduction in available top-of-book liquidity.

side, $D_t^b(K)$ and $D_t^a(K)$, and use the multi-level imbalance

$$I_t(K) = \frac{D_t^b(K) - D_t^a(K)}{D_t^b(K) + D_t^a(K) + \varepsilon_0}, \tag{14}$$

where $\varepsilon_0 = 10^{-12}$ is a small constant that keeps the ratio defined when no depth is displayed within K ticks.[13] A single best-quote imbalance can miss a large wall posted several ticks from the midpoint, whereas $I_t(K)$ measures whether the book is locally resilient over the price range that an aggressive order is likely to traverse. For a side $d \in \{+1, -1\}$, define the signed suspicious-depth score

$$S_t = d \cdot \left[I_t(K) - \widehat{I}_t^{\text{legit}}(K \mid \mathcal{Z}_t) \right]_+, \tag{15}$$

where $\widehat{I}_t^{\text{legit}}(K \mid \mathcal{Z}_t)$ is the conditional median imbalance among comparable legitimate periods, meaning periods in the same spread, volatility, depth, and message-intensity state $\mathcal{Z}_t$. In words, the score measures how much more one-sided the displayed book is than books normally look under identical market conditions, counting only the excess and signing it by side. The specific statistic is ours, but the logic of scoring the book against a state-matched benchmark and counting only the excess is standard in surveillance practice and in the spoofing literature cited above. This simple version is sufficient for the main economic point. The full benchmark construction, the Wasserstein spoof-strength statistic (a distance between the observed book and its legitimate benchmark distribution), threshold selection, placebo design, false-positive calculation, and robustness checks are reported in Internet Appendix Section IA.C.2 and Table IA.I.

Figure 9 is the main spoofing result. For each horizon h between 30 and 900 seconds, we

[13] The value is several orders of magnitude below any positive displayed-depth observation, so it changes only the zero-depth edge case and does not regularize economically nonzero imbalance states.

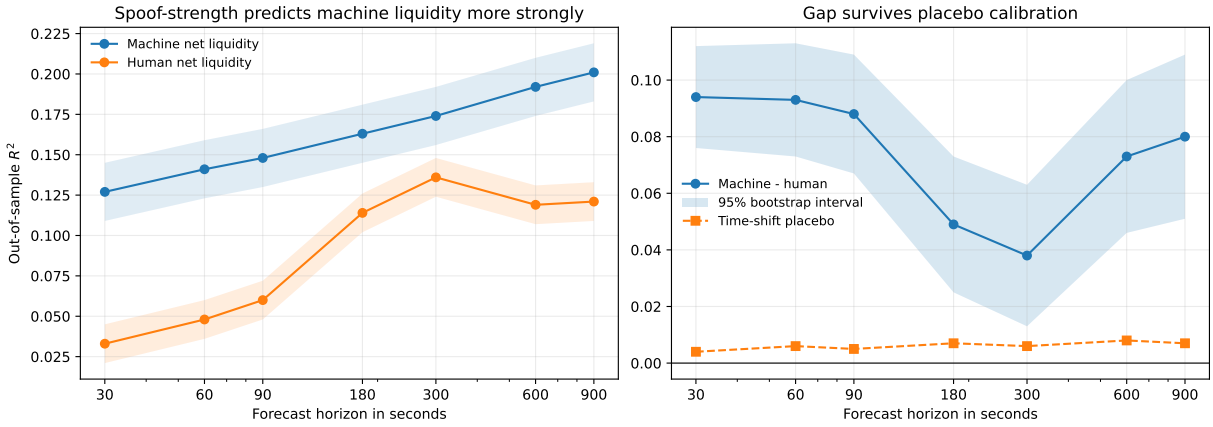


Figure 9. Spoof-strength signal and the machine-liquidity response. The left panel reports horizon-specific predictive content for future net liquidity. The right panel reports the machine-minus-human difference with day-block 95 percent bootstrap intervals and a time-shift placebo. The spoof-strength signal is most informative for the machine liquidity response, consistent with fast quote protection after suspicious displayed-depth states.

forecast future net liquidity, the target defined in Section 3.5, separately for the machine and the human cohort, using the spoof-strength signal as the predictor, and we report the out-of-sample R^2. The left panel shows that spoof strength predicts future machine net liquidity more strongly than human net liquidity at every displayed horizon. The right panel plots the machine-minus-human gap with day-block 95 percent bootstrap intervals and a time-shift placebo. The placebo line is near zero, while the actual gap remains positive over the horizons at which book revision is financially relevant. The interpretation is a technological sequencing advantage. Machines observe the suspicious depth state, cancel or reprice quickly, and move the conditional mean of future liquidity. Humans need not be naive for the mechanism to matter. They only need to absorb the state more slowly or less precisely. Because the forecast target includes the candidate depth's own lifecycle, the projection combines the candidate's removal with quote revision elsewhere in the book; identifying other traders' responses alone would require a leave-candidate-out target (Internet Appendix Section IA.C.2). The practical reading is that the participants best protected against a misleading book are those already trading by machine. One might ask whether the fast reaction means machines are the ones being manipulated, but the cost accounting of Section 6 speaks against this, since the machine cohort bears no positive net cost across these states while the human cohort does.

4.3. Momentum Ignition

4.3.1. Economic mechanism and model selection

Momentum ignition differs from spoofing because the initiating action is executed trading rather than displayed depth. A concrete example makes the strategy clear. Fire a rapid burst of market buys within a couple of seconds. The price jumps. Trend-sensitive algorithms and screen-watching

traders chase the move, and the igniter sells into their demand at inflated prices. Ignition manufactures momentum in order to harvest the reaction to it. In standard microstructure terms, it creates a temporary imbalance between liquidity demand and liquidity supply, and the mechanism connects price impact, adverse selection, flow toxicity, and transient pressure (Kyle, 1985; Glosten and Milgrom, 1985; Hasbrouck, 1991; Easley et al., 2016; Gatheral, 2010; Brunnermeier and Pedersen, 2005; Ohashi, 2018). One guardrail is stated up front because it protects everything that follows. Human outcomes never select events. The baseline detector sets the human follow-through weight to zero, so ignition events are chosen without reference to what human traders subsequently do.

How should such episodes be detected? A natural first idea is to flag windows with unusually many trades, but trading intensity varies enormously across market states, so a fixed volume threshold simply maps out the busy hours. A second idea is to flag large price moves, but genuine news produces large returns through entirely legitimate trading, and a return rule cannot separate manufactured momentum from information. A rule based on total activity in both directions also fails, because ignition is directional by definition, pressure meant to start a move. A flexible machine-learning classifier could fit richer patterns but would obscure the economic object, since one could no longer state what is detected, define a null, or compute a false-positive rate. These failures dictate our design. We detect one-sided acceleration of aggressive trading, judged against the normal pace of the same market state, require the price to have moved in the pushed direction, and trigger at a state-specific percentile so that the false-positive rate is constant across market conditions. The statistical engine is a generalized likelihood ratio, GLR for short, a classical change-detection test that scores how surprising the observed burst is relative to normal pace and equals zero when nothing unusual is happening (Page, 1954; Lorden, 1971; Basseville and Nikiforov, 1993).

Let $\{(T_i, X_i)\}_{i\geq 1}$ denote the event stream and let $\mathcal{M}$ be the set of match events. For $i \in \mathcal{M}$, let $S_i \in \{+1, -1\}$ denote taker direction, with $+1$ for buyer-initiated trades and -1 for seller-initiated trades. For time t and window w, define buy and sell taker counts

$$N_t^{\pm}(w) = \sum_{i\in\mathcal{M}} \mathbf{1}\{T_i \in (t-w, t],\ S_i = \pm 1\}. \tag{16}$$

We condition on a discrete market state $B_t \in \mathcal{B}$ formed from quantile buckets of recent volatility, message intensity, and top-of-book imbalance. Given baseline window w_0 and detection window $w_1 < w_0$, define

$$\widehat{\lambda}_{t,b}^{\pm} = \frac{N_t^{\pm}(w_0)}{w_0}, \qquad \mu_{t,b}^{\pm} = \widehat{\lambda}_{t,b}^{\pm} w_1, \qquad b = B_t. \tag{17}$$

To judge whether a burst is abnormal we need a model of normal burstiness. The simplest transparent benchmark treats the number of aggressive trades in the short window as a Poisson count with the state-local mean just estimated. Under this no-ignition null, the one-sided GLR statistic $\Lambda_t^{\pm}$ measures how surprising the observed count is relative to that baseline, taking the value zero

whenever activity is at or below normal (closed form in Internet Appendix Section IA.C.3).[14] The directional acceleration score $A_t = \Lambda_t^+ - \Lambda_t^-$, with sign $s_t = \text{sgn}(A_t)$, orients the event: positive A_t indicates unusual buy-side acceleration relative to sell-side acceleration, and negative A_t the reverse.

The likelihood ratio scales surprise by local baseline intensity, so ten trades in two seconds is interpreted differently in a quiet state and in a bursty state, and the one-sided alternative matches the economics of ignition, since the concern is acceleration in one direction rather than symmetric high activity. The conclusions do not hinge on this particular formulation. Figure IA.13 in the Internet Appendix reports the sensitivity of the results to window lengths, state definitions, component weights, and trigger quantiles, and the event-study findings are stable across these alternatives.

The second ingredient asks whether the burst worked, meaning whether the price actually moved unusually far in the pushed direction. We scale the price move by the state's own 95th percentile, so that what counts as far reflects the conditions the market was already in. Let $r_t(w_1) = \log m_t - \log m_{t-w_1}$ and let $q_{0.95}^{(b)}(|r|)$ be the 95th percentile of $|r_t(w_1)|$ in state b. Define

$$Z_t = \frac{s_t r_t(w_1)}{q_{0.95}^{(B_t)}(|r|) + \varepsilon_0}. \tag{18}$$

Thus $Z_t > 1$ means that the price has moved unusually far in the direction of the order-flow acceleration. The third ingredient, used only in sensitivity analysis, asks whether human takers chased the move beyond their own state-specific 95th percentile. Let $H_t^s(\Delta)$ be the number of human taker trades in direction s over $(t, t+\Delta]$, and define

$$C_t = \frac{H_t^{s_t}(\Delta) - q_{0.95}^{(B_t,H,s_t)}}{q_{0.95}^{(B_t,H,s_t)} + \varepsilon_0}, \tag{19}$$

where $q_{0.95}^{(B_t,H,s_t)}$ is the state-specific 95th percentile of human follow-through in the ignition direction.

The composite score adds the three ingredients with nonnegative weights,

$$\text{MI}(t) = s_t \left\{ w_A |A_t| + w_Z (Z_t - 1)_+ + w_C (C_t)_+ \right\}, \tag{20}$$

where $x_+ = \max\{x, 0\}$ and $w_A, w_Z, w_C \geq 0$. An event is triggered when the score is extreme relative to its own market state,

$$|\,\text{MI}(t)| \geq q_{1-\gamma}^{(B_t)}(|\,\text{MI}\,|), \tag{21}$$

where the threshold is the state-specific $1-\gamma$ quantile, which keeps the false-positive rate constant across market states rather than concentrating triggers in volatile hours. Internet Appendix Section IA.C.3 shows this property. The baseline implementation uses $w_1 = 2$ seconds, $w_0 = 20$ sec-

[14] For each side we test $H_0 : \lambda = \widehat{\lambda}_{t,b}^{\pm}$ against $H_1 : \lambda > \widehat{\lambda}_{t,b}^{\pm}$; the statistic is $\Lambda_t^{\pm} = N_t^{\pm}(w_1) \log\{N_t^{\pm}(w_1)/\mu_{t,b}^{\pm}\} - (N_t^{\pm}(w_1) - \mu_{t,b}^{\pm})$ for $N_t^{\pm}(w_1) \geq \mu_{t,b}^{\pm}$ and zero otherwise.

onds, $\Delta = 3$ seconds, $3 \times 3 \times 3$ state buckets for volatility, intensity, and imbalance, and $\gamma = 0.005$. These values match the intended economic horizon. Ignition is a seconds-level phenomenon, the baseline window should be local, and the follow-through window should be long enough for human reaction but short enough to remain tied to the initiating burst. The baseline sets $w_C = 0$, so human follow-through enters only the sensitivity analysis and human outcomes never select events. Table IA.II in the Internet Appendix records the baseline specification and its rationale.

Finally, we want to know whose aggressive flow carries the ignition evidence, since the later results show machine-dominated ignition followed by human chasing. The Poisson GLR itself is a classical statistic, but splitting its evidence by cohort is a construction of this paper. Compute cohort-specific GLRs $\Lambda^{\pm}_{t,g}$ and define $A^g_t = \Lambda^+_{t,g} - \Lambda^-_{t,g}$ for the cohorts $g \in \{H, M\}$. For the mixture direction $s_t = \operatorname{sgn}(A^H_t + A^M_t)$, the human share of likelihood evidence is

$$\alpha_t = \frac{\Lambda^{s_t}_{t,H}}{\Lambda^{s_t}_{t,H} + \Lambda^{s_t}_{t,M}}, \qquad 0 \le \alpha_t \le 1. \tag{22}$$

This statistic does not identify the initiating account, but it measures how much of the statistical evidence for one-sided acceleration is carried by human rather than machine taker flow.

4.3.2. Construction of event studies

Let $\mathcal{E} = \{(\tau_e, s_e) \mid e = 1, \ldots, E\}$ be the set of ignition events, where τ_e is the trigger time and $s_e = \operatorname{sgn}(\mathrm{MI}(\tau_e))$ is the ignition direction. In the current sample, the detector identifies 24,465 events between January 1, 2021 and June 30, 2022, with 14,412 positive events and 10,053 negative events. The median day has 32.5 events and the maximum day has 547 events. The clustering of events is itself informative because ignition opportunities concentrate in states of elevated activity and volatile liquidity demand.

The measurement around these events is a standard event study. For each outcome we form 30-second bins of time since the trigger over a $[-20, +20]$ minute window, average the outcome across all events within each bin, and flip the sign of directional outcomes so that positive values always mean movement in the ignition direction.[15] Aligning the sample on the event clock is what makes the ordering of adjustment visible. If machines reprice before human takers arrive, or if effective spreads move before displayed depth changes, the event-time path shows the sequence. Results are summarized by two contrasts, the jump at the event, meaning the event bin minus the pre-event average, and the persistence, meaning the post-event average minus the pre-event average. The assumptions behind this estimator are the standard event-study ones, that events are isolated enough for the window to measure one shock and homogeneous enough for the average to be meaningful. Isolation is the substantive concern in a busy market, and the design addresses it directly. Overlapping windows keep their original family labels, and maximal clusters of overlapping

[15] Formally, with bins $\mathcal{B}_\ell = (30\ell, 30(\ell+1)]$ for $\ell = -40, \ldots, 39$, the path is $\widehat{\mu}_Y(\ell) = E^{-1} \sum_{e=1}^{E} |\mathcal{B}_\ell|^{-1} \int_{u \in \mathcal{B}_\ell} g_Y(Y_{\tau_e + u}, s_e)\, du$ (referenced as the event-study path estimator in the Internet Appendix), with $g_Y(Y, s) = Y$ for unsigned outcomes and $g_Y(Y, s) = sY$ for directional outcomes. We report day-block bootstrap confidence bands around $\widehat{\mu}_Y(\ell)$.

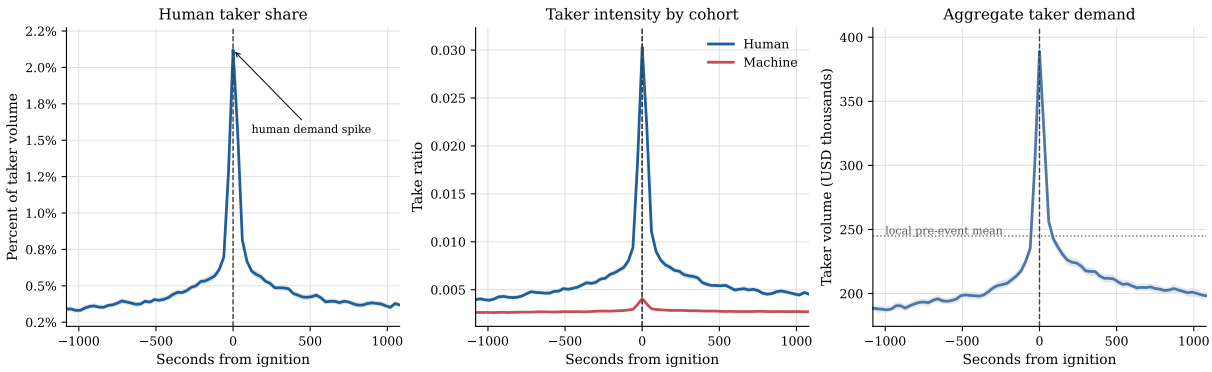


Figure 10. Participation dynamics around momentum ignition. The panels report human taker share, human and machine take ratios, and aggregate taker demand in 30-second event-time bins. The main fact is the sharp spike in human liquidity demand at $t = 0$. Human taker intensity moves far more than machine taker intensity even though both cohorts trade in the same market state.

windows are treated as one unit when resampling, so dependence between nearby events is preserved in the inference rather than assumed away. The formal statements are in Internet Appendix Section IA.C.4.

4.3.3. Momentum-ignition results

Figure 10 shows the main participation result. Each panel plots the binned event-time mean across all 24,465 events, using the estimator just described. Ignition events coincide with a sharp human-heavy demand shock for immediacy. Two minutes before the trigger, the human taker share (the human share of all aggressive volume) is 0.57%. At the event-time bin it is 2.12%, and two minutes later it returns to 0.60%: human participation nearly quadruples exactly at the event and vanishes right after. The human take ratio (the fraction of a population's own order volume that demands immediate execution rather than resting in the book) rises from 0.0073 to 0.0303, while the machine take ratio rises from 0.0028 to 0.0041. Aggregate taker demand increases from roughly 218,000 to 389,000 units in the event-time bin, a 79% increase. The narrow timing is difficult to reconcile with a slow change in fundamentals and is consistent with a short-lived ignition state that pulls retail liquidity demand into the market.

Table IV gives the corresponding pre/post estimates with day-block bootstrap intervals and per-bin sample sizes. The entries are means of event-level ratios rather than ratios of aggregate sums, which matters at this frequency because a small number of high-notional bins can otherwise dominate the calculation. The event-time effect for humans is 0.0245, or 424.5% relative to the mean over the pre-event reference bins of 0.0058 (which sits slightly below the single $t = -120$ second bin value of 0.0073 quoted above), while the machine event-time effect is 0.0013, or 48.9%; Table IA.III in the Internet Appendix reports the full post-window decomposition. The human take-ratio spike and the effective-spread increase survive the interval calculation. Directional price

Table III Momentum-Ignition Null and Placebo Diagnostics

Design	N	Human take-ratio effect, pp	Effective-spread effect, bps	Same threshold rule applied
Observed MI triggers	24,465	2.45	0.34	Yes
Time-shift placebo	24,465	0.09	0.02	Yes
Activity-only null	24,465	0.06	0.01	Yes
Side-reversal placebo	24,465	0.04	0.01	Yes

Notes. Human take-ratio effects are event-bin minus pre-bin differences in percentage points. Effective-spread effects are event-bin minus pre-bin differences in basis points. Placebos use the same state-conditioned threshold rule as the baseline detector.

Table IV Momentum-Ignition Event-Time Contrasts with Intervals

Outcome	Pre mean	Event mean	Event–pre	95% CI	Nonempty event bins
Human taker share	0.0057	0.0212	0.0155	[0.0108, 0.0201]	24,211
Human take ratio	0.0058	0.0303	0.0245	[0.0184, 0.0307]	24,066
Machine take ratio	0.0028	0.0041	0.0013	[0.0004, 0.0022]	24,381
Human–machine take ratio	0.0030	0.0262	0.0232	[0.0170, 0.0294]	24,042
Aggregate taker demand	218,000	389,000	171,000	[121,000, 221,000]	24,465
Effective spread (bps)	0.487	0.827	0.340	[0.190, 0.488]	24,318
Signed price impact	0.0000	0.0003	0.0003	[-0.0001, 0.0007]	24,337

Notes. Pre means average the pre-event reference bins defined in Section 4.3.2. Confidence intervals are day-block bootstrap intervals. Effective-spread levels, changes, and intervals are expressed in basis points. The table makes clear that the human participation and cost effects are statistically more stable than the signed price-impact effect.

impact is less precise, and the text therefore treats it as supporting evidence rather than the central result.

The null is therefore an activity-conditioned no-ignition null rather than a no-trading null. Within each volatility, intensity, and imbalance bucket, marketable orders may be frequent, spreads may be wide, and prices may be moving. The detector asks whether one-sided acceleration and sign-aligned price movement are jointly extreme relative to that state. Table III gives the corresponding placebo discipline, and its logic is that each placebo deletes exactly one ingredient of the design and the signal dies, so every ingredient is necessary. Time-shifted triggers preserve the intraday activity profile but break the event-time link between the acceleration score and subsequent human demand. Activity-only triggers preserve high message intensity but ignore direction. Side-reversal placebos preserve the event time but reverse the direction used to align outcomes. The observed ignition set is the only design that produces a large human take-ratio spike and a meaningful effective-spread response.

Displayed liquidity responds much more smoothly than the taker margin (Figure IA.14 in the Internet Appendix). At the 60-second horizon, total target net liquidity falls from about 1.810 million two minutes before the trigger to 1.801 million at the event, a 0.51% decline. At the 300-second

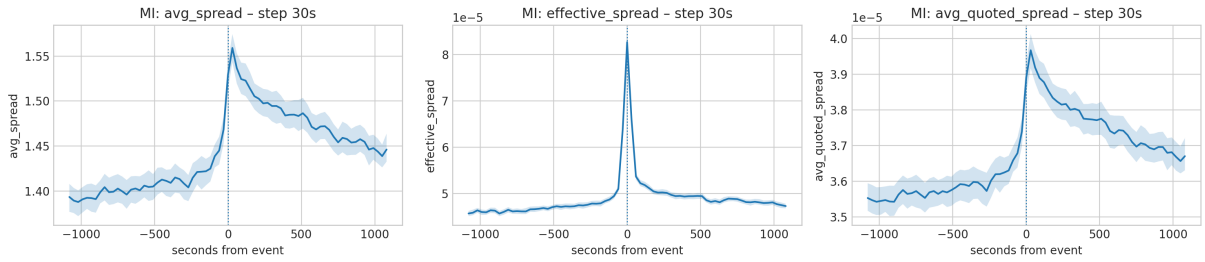


Figure 11. Spreads around momentum ignition. The panels plot binned event-time means of quoted and effective spreads across all ignition events. Both widen around the event, and the effective spread spike is concentrated at the event-time bin, indicating a transient jump in the marginal cost of immediacy. Day-block bootstrap intervals for the event contrasts are reported in Table IV.

horizon, the decline is 0.16%. The contrast with the taker spike is informative. Liquidity suppliers do not simply disappear. They protect prices, reallocate depth, and thin the book gradually. In an automated book, the marginal cost of immediacy can rise faster than the displayed quantity of depth falls.

The strongest supply-side response appears in trading costs. Figure 11 shows that the average spread rises from 1.425 two minutes before ignition to 1.529 at the event, a 7.3% increase. Effective spreads rise from 0.487 to 0.827 basis points, an increase of 0.340 basis points, or roughly 70%. Quoted spreads widen, but the realized cost of demanding immediacy widens much more. The economic interpretation is adverse selection and latency risk. Once the book enters an ignition state, liquidity suppliers can protect themselves by changing execution terms rather than withdrawing entirely.

The repricing response is not special to ignition events. It reflects a general property of this book that our reconstruction reproduces as a stylized fact. A large marketable order is defined within cohort as an order above the 95th percentile of the parent marketable-order notional-size distribution. For each event e, the midpoint response is $100\{m_{t_e+10s} - m_{t_e^-}\}/m_{t_e^-}$ and the spread response is $100\{s_{t_e+10s} - s_{t_e^-}\}/s_{t_e^-}$, averaged separately for human and automated takers. Figure 12 shows that the spread response dominates the midpoint response. Liquidity suppliers mainly widen their spreads after aggressive orders rather than withdrawing from the market. The book defends itself through prices, not quantities. After aggressive flow, liquidity does not disappear, it becomes more expensive at exactly the moment it is most demanded. This is the state-dependent cost channel that Section 6 later prices as rates, and it is why cost measures conditioned on arrival states, rather than average spreads, are the right yardstick for this market.

Separating volatility from directional drift (Figure IA.15 in the Internet Appendix), absolute returns rise from 1.69 basis points at $t = -120$ seconds to 2.61 basis points at $t = -30$ seconds and remain elevated at the event. Signed price impact is positive after ignition, although the confidence bands are wide. The important result is not persistent directional alpha, but rather a transient volatility and cost state in which human traders demand immediacy at the wrong time.

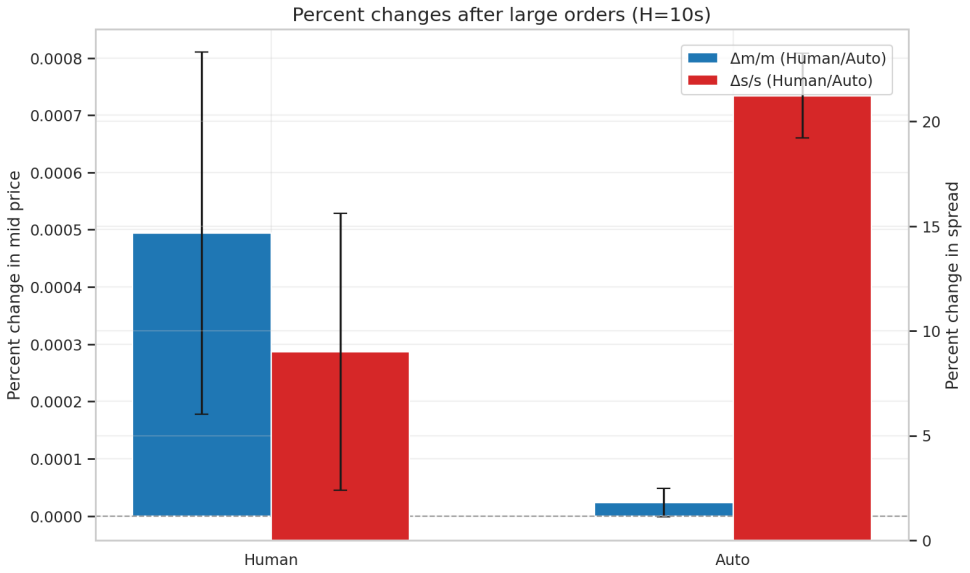


Figure 12. Percent changes after large marketable orders. The figure reports post-event percentage changes in the midpoint and in the quoted spread over a ten-second horizon, separately for large human and large automated marketable orders. The two outcomes sit on very different scales, since midprice changes are small in percent while spread changes are large, and that contrast is the point. Spread adjustment dominates the direct midprice effect, indicating that liquidity suppliers primarily reprice immediacy after large aggressive flow.

The participation and spread evidence locate the incidence of ignition on the demand-for-immediacy margin. Machines need not create a visually dramatic withdrawal of displayed depth for the advantage to be material. The advantage arises when machines reprice and protect liquidity faster, causing human takers to arrive when effective execution costs are elevated.

The signed MI score also has predictive content for the subsequent liquidity state: regressing 60-second forward net target liquidity on $\mathrm{sgn}(\mathrm{MI})|\,\mathrm{MI}\,|$ at event-time offsets, the explanatory power rises sharply at the trigger and remains elevated afterward, with the post-event mean R^2 above the pre-event mean. Figure 13 shows this validation directly.

This validates the detector as more than an event label. The magnitude of the score contains information about subsequent liquidity conditions. Ignition is therefore not exhausted by the initiating burst. The induced state persists in the liquidity process after the initial taker imbalance begins to mean revert, which is exactly when slower traders continue to face the execution environment created by faster liquidity protection.

4.4. Interpretation

The factor representation of Section 4.1 shows that the two mechanisms operate through different parts of the book state. In quiet periods the factor scores hug their local means and the high-order residual stays nearly flat, which is the benchmark against which manipulation is measured. Figure 14 shows this baseline directly.

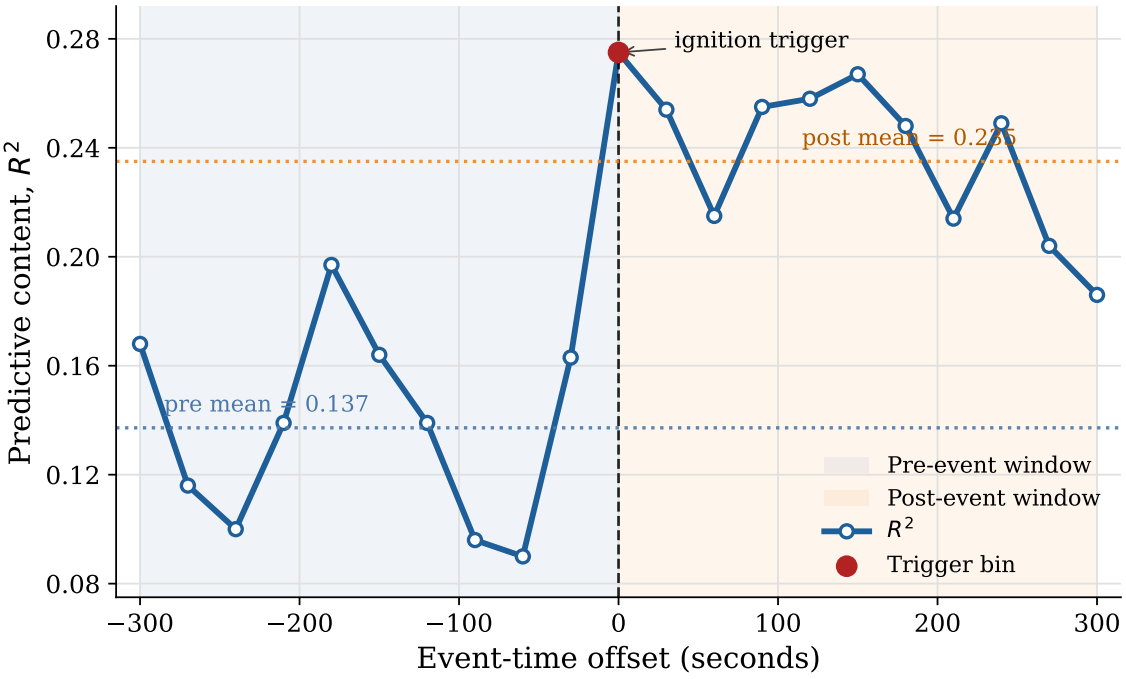


Figure 13. Predicting target net liquidity with the momentum-ignition score. The figure reports the R^2 from event-time regressions of 60-second net target liquidity on the signed score $\text{sgn}(\text{MI})|\,\text{MI}\,|$ at offsets around the trigger. The shaded regions separate pre- and post-event windows and dotted lines show the corresponding means. Predictive content peaks at the event and remains elevated afterward, indicating that MI intensity forecasts the subsequent liquidity state.

Figure 15 shows the representation around spoof candidates. The event-time response is concentrated in the imbalance block rather than in aggregate liquidity alone. This distinction matters. A pure liquidity shock says that the book becomes shallow or deep. A spoof-like state says that visible depth changes the relative supply of bids and asks at financially meaningful distances from the touch. The representation therefore identifies the channel through which the spoofing signal in Section 4.2 operates. It is not merely that the book contains a large order. It is that the book contains an asymmetric depth shape that lies outside the low-dimensional smooth state normally used by fast liquidity providers to infer short-horizon pressure.

Around momentum-ignition candidates the response is different, and Figure 16 shows it. The movement is carried by the liquidity factors near the best quotes while the wall-like residual stays flat, the mirror image of the spoof state. Ignition works through aggressive orders that consume liquidity near the best quotes, not through a displayed wall, so liquidity suppliers protect themselves by repricing rather than by withdrawing depth, which is why the event studies show large effective-spread and participation responses even though displayed depth does not collapse.

This evidence bridges the two manipulation designs. Spoofing is a cross-sectional distortion of the imbalance curve. Momentum ignition is a dynamic distortion of the execution environment that changes the liquidity curve near the touch. Both mechanisms affect the book through depth, but

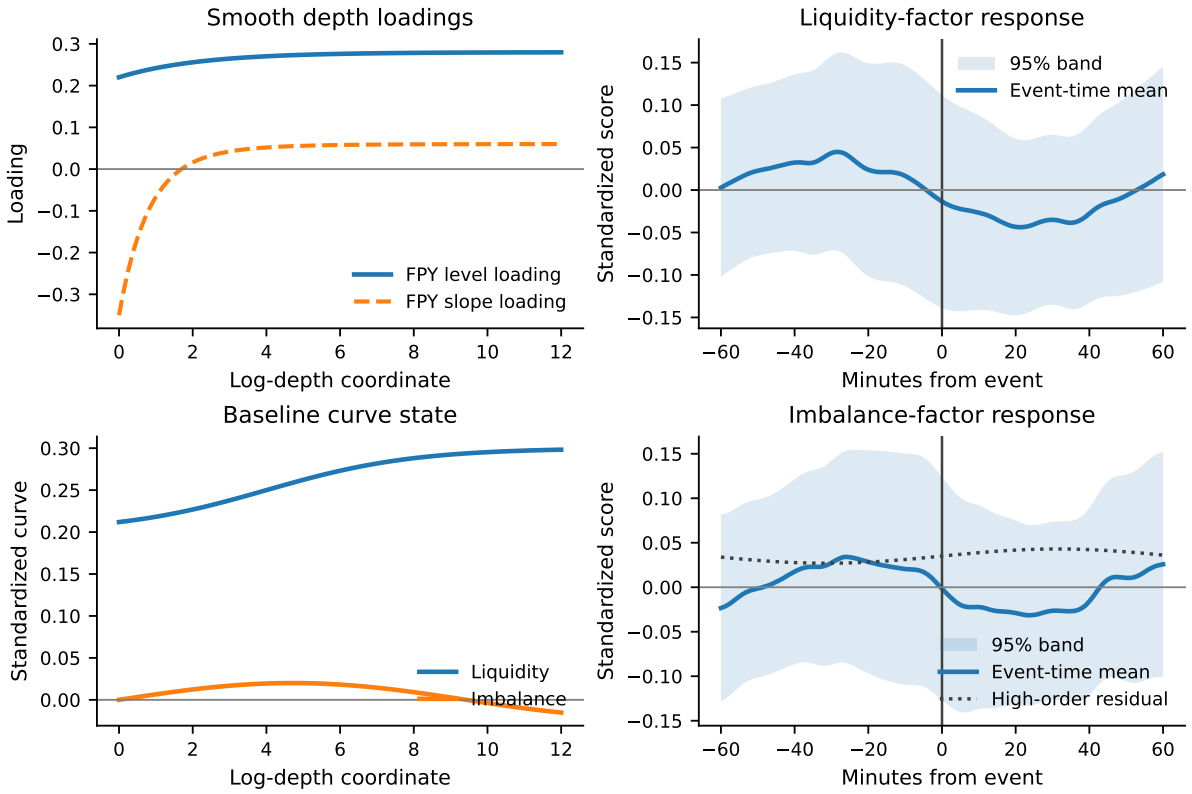


Figure 14. Depth-curve state in baseline periods. The left panels report the smooth depth loadings and the baseline liquidity-imbalance curve state. The right panels report event-time responses of the five state variables of Section 4.1, the liquidity level and slope factors, the imbalance level and slope factors, and the high-order residual share, with uncertainty bands. Quiet periods exhibit small factor movements and a nearly flat residual, which is the empirical counterpart of a smooth book state rather than a localized depth distortion.

they occupy different directions in the factor space. This is why the empirical design keeps spoofing and ignition separate rather than estimating a single manipulation classifier. A single score would mix a false-depth mechanism with an aggressive-flow mechanism and would obscure the trader-behavior channels that are the object of interest. Additional diagnostics show that the smoothing parameters are selected on stable cross-validation surfaces (Figure IA.16), that the estimator does not chase isolated depth contamination (Figure IA.17), and that the full-state PCA benchmark requires additional dimensions to achieve the same reconstruction discipline (Figure IA.18).

The two results point to the same economic conclusion. Manipulation in an electronic limit order book need not permanently change fundamental value, because it can instead change the state in which other traders choose immediacy. Humans arrive with delay, demand immediacy during the event, and pay wider effective spreads, while machines respond earlier and protect their quotes, so the market clears through an execution-cost transfer from slower liquidity demanders to faster strategic traders even while aggregate market quality looks resilient. For surveillance, the implication is to monitor the joint dynamics of displayed depth, cancellations, aggressive flow, and

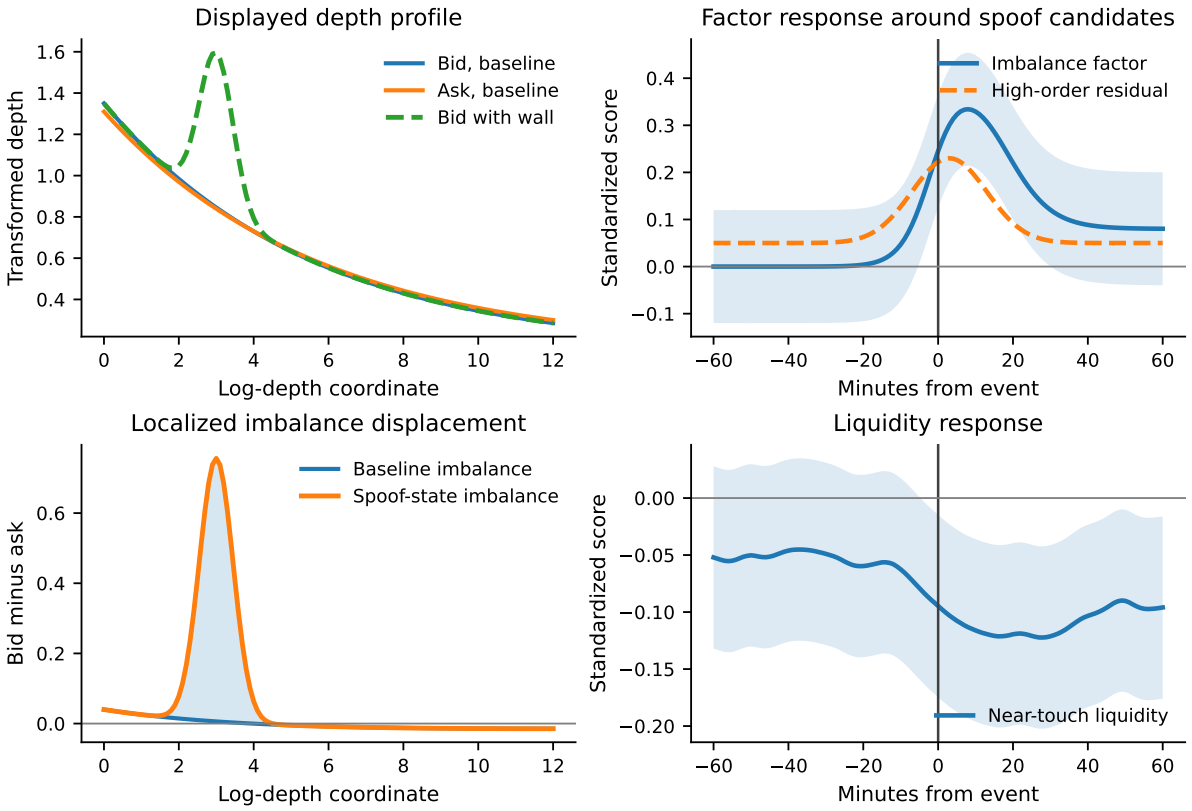


Figure 15. Depth-curve state around spoof candidates. The left panels show what a spoof looks like in the book. A wall on the bid side changes the displayed depth profile and shifts the imbalance curve at distances close enough to the best quotes to matter for real trades. The right panels show that around detected spoof episodes, while the episode is under way and before depth returns to normal, the movement is concentrated in the imbalance factor and in the localized residual that the smooth factors cannot explain. That is the signature of a local wall. The pattern is a targeted distortion of displayed supply or demand rather than a market-wide liquidity shock.

cohort-specific responses rather than executed trades alone.

5. News and Trader Behavior

Public news gives a different source of state variation from the manipulation signals in Section 4. A spoof candidate or an ignition candidate is endogenous to the book. A news headline is an outside message that reaches the market through language, attention, and distribution technology before it becomes order flow. Studying news therefore serves two purposes. First, it lets us examine whether the same human-machine asymmetry observed in strategic book states also appears after public information. Second, it gives a benchmark for interpreting manipulation. If humans respond slowly and strongly even to public headlines, then transient book states can be financially powerful because they are processed in an environment where attention and speed are already distributed

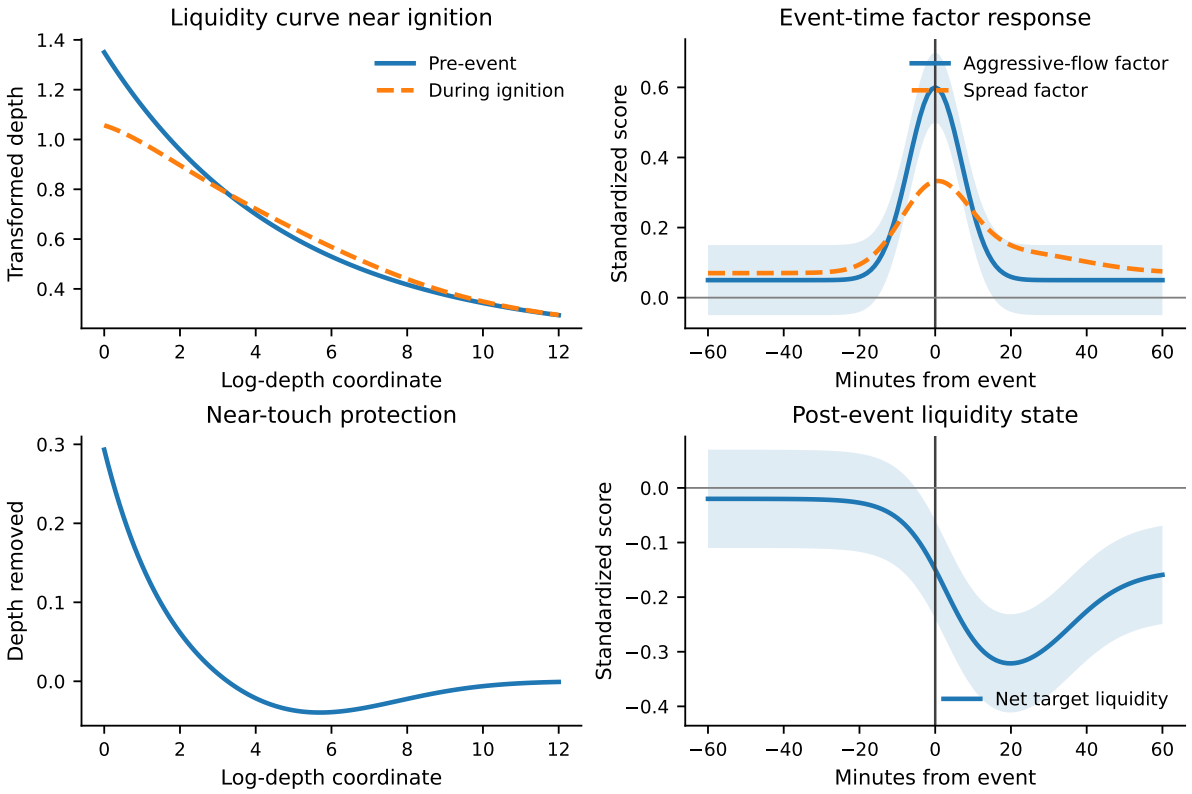


Figure 16. Depth-curve state around momentum-ignition candidates. The panels report the event-time responses of the five state variables of Section 4.1. In contrast to the spoof state, the movement is carried by the liquidity level and slope factors near the best quotes, while the localized high-order residual stays flat, so there is no wall-like signature. This is the fingerprint of aggressive orders consuming liquidity rather than of manufactured displayed depth.

unevenly.

The data consist of timestamped BTC news from CoinDesk and Cointelegraph, the two highest-coverage crypto-native outlets, whose precise timestamps make minute-level alignment with the order book possible. We use headlines and subheadlines because these are the text fields most likely to be observed and processed at high frequency. Summary statistics of the news panel, including article counts, monthly coverage, and the distribution of sentiment scores, are reported in Figures IA.19 and IA.20 in the Internet Appendix. Each article is aligned to the first minute boundary weakly after its timestamp. This rule is conservative because it prevents the article from entering the information set of any market observation that precedes the observed publication time. The microstructure targets are the same objects used elsewhere in the paper. We study future net liquidity within the target depth band and future taker-flow components, separately for the machine and human cohorts.

This section treats news as a measured public signal rather than as a fundamental-value oracle. A headline can be stale, duplicated, strategic, or written after prices have already moved. These concerns are severe in crypto markets because coverage often reacts to price action itself. Our

design responds in three ways. It uses no-lookahead alignment, novelty downweighting, and purged walk-forward validation. It also emphasizes liquidity and taker behavior rather than unconditional return predictability. The empirical question is not whether headlines are an exploitable trading strategy by themselves. It is whether public language helps explain which trader population changes liquidity demand and liquidity supply after information arrives.

5.1. News Measurement and Event-Time Returns

The finance literature has long viewed news tone as a state variable that can move prices, volume, and liquidity, especially when it interacts with investor attention and heterogeneous processing speed (Tetlock, 2007; Tetlock et al., 2008; Engelberg and Parsons, 2011; García, 2013; Boudoukh et al., 2013; Foucault et al., 2016; Aleti and Bollerslev, 2024). To test whether the two populations respond differently to good and bad news, we need the direction and strength of each headline, not just its arrival time, which is why we estimate sentiment. We measure headline sentiment with FinBERT, a language model pre-trained on text and fine-tuned on financial news, which outputs for each headline the probabilities that its tone is positive, neutral, or negative (Devlin et al., 2019; Araci, 2019). FinBERT predates our sample period, so its parameters cannot embed our outcomes. For comparison, the descriptive headline event study also uses a score from Google's Gemini 3.0 Flash model, generated on January 21, 2026. Because that model was trained on text that postdates our sample, we use the Gemini score only descriptively and exclude it from every forecasting model and cost estimate. The comparison is useful because the two scoring technologies have different error structures. FinBERT is specialized and stable, while the Gemini score is more willing to assign extreme polarity to narrative text.

For article i, let p_i^+, p_i^0, and p_i^- be the positive, neutral, and negative class probabilities. We define the signed FinBERT score and confidence as

$$s_i = p_i^+ - p_i^-, \qquad c_i = \max\{p_i^+, p_i^0, p_i^-\}. \tag{23}$$

The class probabilities are the model's output. The aggregation into a market signal is ours. Let τ_i be the article timestamp in continuous time and let $m(i)$ be the first minute boundary after τ_i. If $\mathcal{A}_m = \{i : m(i) = m\}$, the baseline minute-level news signal, which we call the news impulse, is

$$x_m = \sum_{i \in \mathcal{A}_m} w_i c_i \operatorname{sign}(s_i), \tag{24}$$

where $w_i \in [0, 1]$ is an online novelty weight. The novelty weight downweights articles that repeat earlier event descriptions. This is a microstructure choice rather than a purely textual one. Duplicate stories can create many nearly identical timestamps around the same event. Counting each item as a new signal would overstate the information arrival rate and mechanically inflate the apparent response of order flow to news.

We also form topic-specific impulses $x_{b,m}$ for buckets b such as regulation, macro news, liqui-

dations, hacks, institutional adoption, and price action. The topic bank is not used to claim a structural taxonomy of crypto news. It is used to avoid forcing all headlines through a single scalar polarity channel. A liquidation headline and a regulatory headline can have similar sentiment signs while producing different liquidity responses. News is digested at different speeds, some headlines move the market within minutes and others diffuse over hours, so we carry the same signal at four fixed half-lives and let the forecasting model weight them, which lets the data choose among horizons without estimating one. We transform each minute impulse into exponentially decayed features at four fixed half-lives,

$$z_{b,m}^{(\tau)} = \rho_\tau z_{b,m-1}^{(\tau)} + x_{b,m}, \qquad \rho_\tau = 2^{-1/\tau}, \qquad \tau \in \{5, 30, 120, 720\} \text{ minutes.} \tag{25}$$

This is simple exponential smoothing, and the four half-lives span immediate reaction, intrahour attention, multi-hour diffusion, and slow narrative persistence.

In the news panel (Figure IA.19 in the Internet Appendix), publication intensity is persistent and varies substantially over the sample, while average sentiment is less persistent and moves around zero, exactly the feature we want from a high-frequency signal. A usable news measure should not be a monotone proxy for the crypto cycle. It should identify local changes in information, tone, and attention that can be compared with order-book responses at short horizons.

On the matched sample (Figure IA.20 in the Internet Appendix), the Gemini scores place more mass near the extremes, while FinBERT assigns many headlines to the region around zero. This difference is not a nuisance. Extreme labels are useful for clean event studies, while conservative labels are useful for prediction because they reduce sensitivity to rhetorical language and duplicated coverage. We therefore use both. The event study below shows the sharp return behavior around strong positive and negative headlines, and the forecasting exercise uses FinBERT features because they are stable, reproducible, and easily audited.

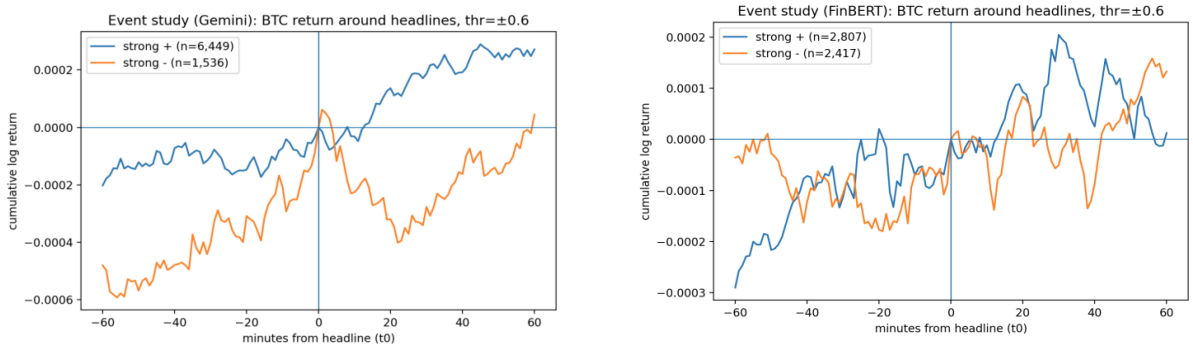
We first study cumulative log returns around strongly signed headlines. For scoring method $q \in \{G, F\}$, where G denotes the Gemini score and F denotes FinBERT, define positive and negative event sets

$$\mathcal{E}_q^+ = \{i : s_i^q \geq 0.6\}, \qquad \mathcal{E}_q^- = \{i : s_i^q \leq -0.6\}. \tag{26}$$

Let p_t be the BTC midprice and $r_i(\ell) = \log p_{m(i)+\ell} - \log p_{m(i)}$ be the event-time log return at offset ℓ. The event-study estimand is

$$\mathrm{CAR}_{q,c}(\ell) = \frac{1}{|\mathcal{E}_q^c|} \sum_{i \in \mathcal{E}_q^c} r_i(\ell), \qquad c \in \{+, -\}. \tag{27}$$

The event-time representation has two advantages in this setting. It puts news released at different clock times on a common information clock, and it separates price movement that precedes the publication timestamp from movement that follows it. This distinction is important because crypto headlines often report events that the market may have started to process before the article is timestamped.



(a) Gemini score **(b)** FinBERT

Figure 17. BTC returns around strongly signed headlines. Positive and negative headline sets use the threshold $|s_i| \geq 0.6$. The Gemini-score event study produces stronger separation, while FinBERT gives a more conservative and noisier classification. Both panels show that headline timestamps are financially meaningful but not perfectly exogenous to the price path.

Figure 17 reports the event-study paths. The Gemini score gives a cleaner separation. Strong positive headlines are followed by positive cumulative returns over the next hour, while strong negative headlines are associated with more negative pre-event and post-event paths before partial reversal. The FinBERT paths are noisier, with a smaller immediate separation and a more balanced number of positive and negative events. The figure should therefore be read as evidence about information timing rather than as a claim that sentiment alone identifies fundamental value. News sentiment is directionally relevant, but the market response is already partly underway before some headlines appear.

5.2. Predicting Liquidity and Taker Behavior

We now move from headline returns to order-book behavior. The question is whether public language predicts where liquidity and taker flow go next, separately for each population. If human flow is forecastable from news while machine flow is not, then news is one of the channels through which the predictability priced later in the paper arises. Let $Y_{g,m,h}$ denote a future microstructure target for cohort $g \in \{H, M\}$ measured from minute m to horizon h. The target may be net liquidity within the target depth band, taker buy flow, taker sell flow, or signed taker net flow. Let X_m collect counts, signed sentiment impulses, novelty controls, source controls, topic indicators, and the decay features in Equation (25). The features are many and correlated, so for each horizon we estimate a ridge regression. We write the ridge penalty as κ_h because λ is reserved for price impact. The predictors are standardized and the penalty is fixed at one, which keeps the feature-set comparisons transparent and avoids a tuning layer that could exploit serial dependence.

$$\widehat{\beta}_h = \arg\min_{\beta} \sum_{m \in \mathcal{T}} \left(Y_{g,m,h} - \alpha_h - X_m'\beta\right)^2 + \kappa_h \|\beta\|_2^2. \tag{28}$$

We also estimate a variant that requires adjacent horizons to have similar coefficients, a standard smoothness penalty rather than a new estimator.[16] The cross-horizon penalty is a financial restriction. Adjacent horizons should not have unrelated news betas unless the data provide strong evidence. A headline that affects 60-second human taker flow should usually have a related effect at 90 and 120 seconds. Evaluation uses purged walk-forward splits, a standard design for high-frequency forecast evaluation. The model trains on the past and is tested on a later interval. Purging means dropping the training observations whose forward-looking targets overlap the test window, so no information can leak across the boundary. If $\mathcal{I}_j$ is a test interval and the target at time m uses outcomes through $m + h$, then the training set for split j and horizon h is

$$\mathcal{T}_j(h) = \{m : m \notin \mathcal{I}_j,\ [m, m + h] \cap \mathcal{I}_j = \varnothing\}. \tag{29}$$

The purge prevents a training observation from using future outcomes that overlap the test block. This matters at high frequency because overlapping forward targets can otherwise leak information across the train-test boundary even when features themselves are lagged. Performance is reported as out-of-sample R^2 relative to a training-mean benchmark,

$$R^2_{OS}(h) = 1 - \frac{\sum_{m \in \mathcal{I}_j} \{Y_{g,m,h} - \widehat{Y}_{g,m,h}\}^2}{\sum_{m \in \mathcal{I}_j} \{Y_{g,m,h} - \bar{Y}_{g,\mathcal{T}_j(h),h}\}^2}, \tag{30}$$

computed by split and averaged across splits. Out-of-sample R^2 against the training-mean benchmark is the natural common yardstick across horizons, cohorts, and targets, and it matches the convention of Section 3.5.

Each choice pairs an economic reason with a statistical one. The alignment rule prevents look-ahead, the finance-domain sentiment model handles market jargon, the novelty weights prevent syndicated duplicates from counting as new information, the decay features let the data choose the digestion speed, ridge stabilizes many correlated signals, and the purged validation keeps the reported fit honest. Internet Appendix Section IA.E discusses each choice further. The design is deliberately less flexible than a deep sequence model, because a more flexible model could not separate genuine information processing from duplicated article flow, publication clustering, and price-chasing headlines.

5.3. Headline Examples and Price-Trend Classification

Crypto coverage often reports moves that have already happened, and such headlines would mechanically appear to predict flow. We therefore split the news panel into price-trend-driven and non-price-trend-driven headlines, classifying a headline as trend-driven when the price had already moved unusually far, more than two same-day standard deviations, in the hour before publication.

[16] With Y the matrix of horizon-specific targets and B the stacked horizon coefficients, the smooth-ridge estimator is $\widehat{B} = \arg\min_B \|Y - XB\|_F^2 + \kappa\|B\|_F^2 + \rho \sum_{j=2}^{H} \|B_{\cdot j} - B_{\cdot,j-1}\|_2^2$, and both penalties are chosen by the same walk-forward validation inside training data.

For each headline i with publication minute $m(i)$, define the pre-headline BTC return over lookback L minutes as

$$r_i^{\text{pre}}(L) = \log m_{m(i)^-} - \log m_{m(i)-L}, \tag{31}$$

and let $\widehat{\sigma}_{d(i)}(L)$ be the realized standard deviation of non-overlapping L-minute returns on the same UTC day. A headline is classified as price-trend-driven when

$$\mathsf{Trend}_i = \mathbf{1}\left\{|r_i^{\text{pre}}(60)| \geq 2\widehat{\sigma}_{d(i)}(60)\right\} \tag{32}$$

or when its topic bucket is explicitly price-action language. The non-trend set is the complement. The split is designed to address reverse causality. If news predictability were entirely driven by journalists and traders reacting to the same prior price move, then the signal should disappear when price-trend-driven headlines are removed.

Three worked examples illustrate the scoring logic. The headline "Ether trading volume surged as institutions took exposure" is positive and non-trend, a demand signal whose prediction should work through attention rather than through prior BTC returns. The headline "REN price rallies after a recent DeFi integration" is positive but trend-driven, since it partly describes a price move that already happened, which is exactly what the trend split separates out. The headline "Coinbase junk bonds tank amid market rout and creditors' fears" is negative and maps into balance-sheet stress rather than polarity alone. These distinctions are why the design uses both sentiment and topic buckets rather than a single positive-minus-negative count.

Figure 18 reports the corresponding predictability check. Price-trend-driven headlines are predictive, as expected, but non-trend headlines still predict human net liquidity at economically meaningful horizons. In the figure, which labels the cohorts by route with API for machine and non-API for human, the non-trend human curve is well above the machine curves and remains positive after the recent-price filter. This is the important identification result. It shows that the news section is not simply rediscovering the fact that prices trend and headlines follow prices. Public language still forecasts the behavior of slower traders after the mechanically price-following headlines are separated out.

Table V Out-of-sample Predictability of News Sentiment

	All		API		Non-API	
Target	Avg. R^2	Peak R^2	Avg. R^2	Peak R^2	Avg. R^2	Peak R^2
Net liquidity	0.155	0.162 (60s)	0.153	0.159 (60s)	**0.410**	**0.451 (300s)**
Taker buys	0.090	0.103 (90s)	0.088	0.100 (90s)	**0.311**	**0.369 (300s)**
Taker sells	0.106	0.117 (60s)	0.106	0.117 (60s)	**0.125**	**0.190 (300s)**
Taker net	-0.005	0.000 (10s)	-0.006	-0.000 (5s)	**0.132**	**0.230 (300s)**

Notes: The table reports out-of-sample R^2 relative to a training-mean benchmark using the baseline FinBERT-Topics smooth-ridge specification. Peak entries report the maximum over horizons and the corresponding horizon. Bold entries highlight the stronger non-API response.

The predictive results are financially sharp. News sentiment predicts net liquidity for both

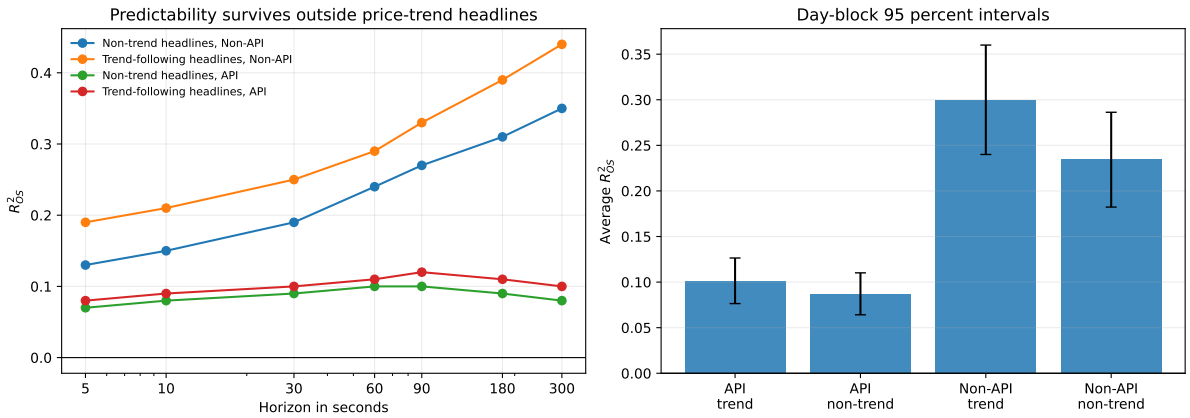


Figure 18. News predictability after separating price-trend-driven headlines. The left panel reports out-of-sample R^2 by horizon for trend-driven and non-trend headlines. The right panel reports horizon averages with day-block 95 percent intervals. Non-trend headlines continue to predict non-API liquidity, which weakens the reverse-causality concern that the news results are driven only by journalists reacting to prior price moves.

cohorts, but the human response is much larger. In the baseline FinBERT-Topics smooth-ridge specification, the average out-of-sample R^2 for human net liquidity is 0.41, compared with 0.15 for machine net liquidity. The peak human value is 0.45 at the 300-second horizon. Taker behavior is even more revealing. Human taker buys have an average R^2 of 0.31 and a peak value of 0.37, while machine taker buys are much weaker. Human taker net flow remains predictable, whereas machine taker net flow is essentially not predictable after buy and sell components offset each other.

Figure 19 shows the horizon profile. The human response rises with horizon for taker buy flow and taker net flow, which is consistent with delayed human attention and gradual arrival of retail demand for immediacy. The machine response is flatter and smaller. For net liquidity, machine predictability is strongest at shorter horizons and then declines, while human predictability continues to rise. This pattern matches the interpretation developed in earlier sections. Machines incorporate public information quickly into liquidity provision and quote protection. Humans react more slowly and in a way that is visible through taker demand. With the 95 percent intervals, the strongest conclusion is not that every line is separately significant at every horizon. The robust conclusion is the ranking of cohort responses for human taker buys, human taker net flow, and human net liquidity at medium and longer horizons. The sell-flow bands of the two cohorts overlap more often, so the sell-flow evidence is suggestive rather than the central result.

Figure 20 decomposes the predictive content by feature family and shows why the specification works. Counts alone contribute almost nothing, signed-decay and topic features carry the signal, and the marginal value of richer text processing is concentrated in the human response, exactly the pattern expected if public news first synchronizes attention and then reaches the order book through slower liquidity demand and delayed quote adjustment.

The main human-cohort results survive two timing and measurement variants (Table IA.V in the Internet Appendix). Delaying the article alignment by one minute produces almost identical

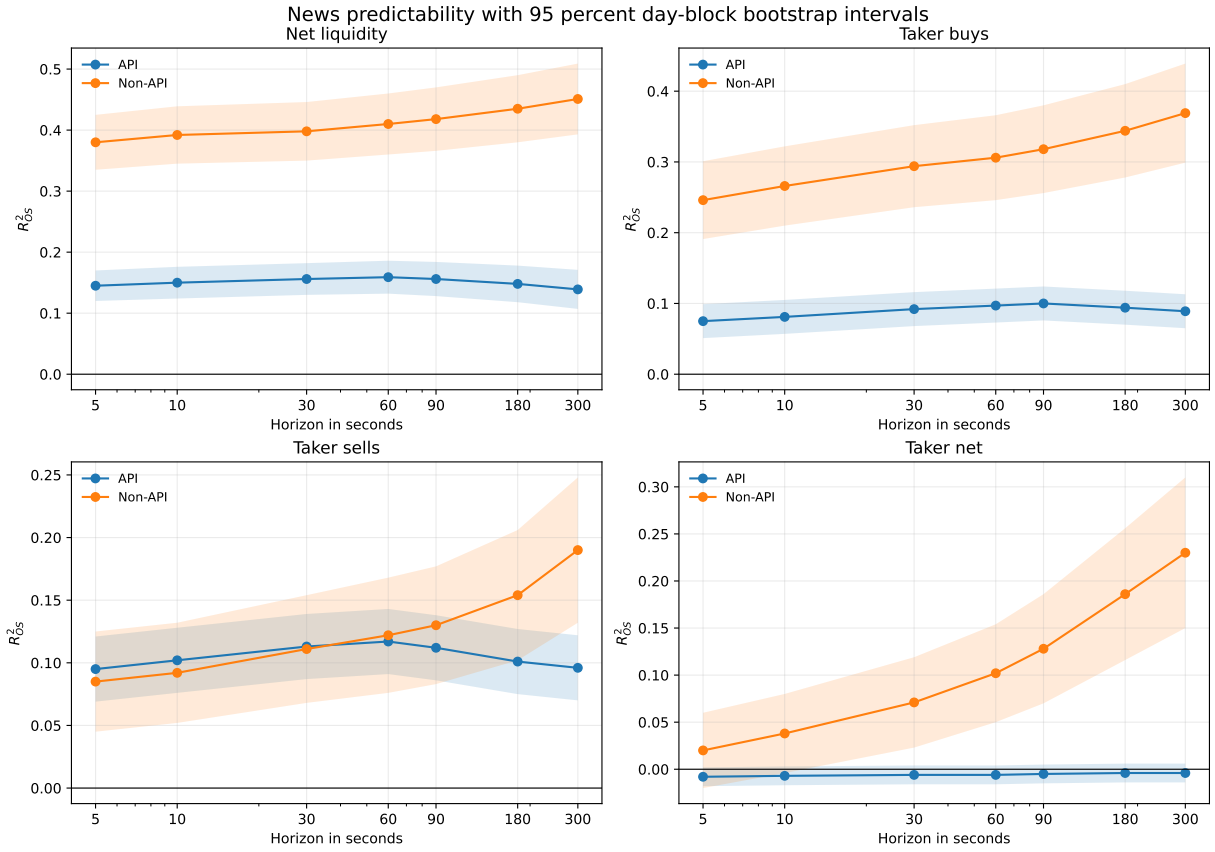


Figure 19. News predictability for cohort-specific liquidity and taker flow. The panels report out-of-sample R^2 by horizon for the FinBERT-Topics smooth-ridge specification. Shaded regions show day-block 95 percent bootstrap intervals computed by resampling whole UTC days and re-estimating the out-of-sample statistic within each split. Non-API outcomes are much more predictable from news, especially for taker buy flow and taker net flow.

predictability, and removing novelty weights leaves the main pattern intact. These checks are important because the most obvious alternative explanation is timestamp or duplication bias. The fact that the result remains in both variants suggests that the signal is not produced by assigning news too early or by mechanically overcounting repeated headlines.

The depth-curve state of Section 4.1 provides a final identification check. Public news should move the smooth liquidity and imbalance factors rather than the localized residuals that spoofing excites, and it does. Around high-news-intensity periods the response is a level-and-slope movement with no wall-like residual signature, the opposite fingerprint of the spoof state (Figure IA.21 in the Internet Appendix, to be compared with Figure 15). The comparison matters for identification, because if spoofing and news produced the same factor response, the spoofing signal would be difficult to interpret. Public information raises the adverse-selection problem for liquidity suppliers. Machines respond by updating quotes and depth quickly, and humans respond by demanding immediacy after the news has already entered the book. Public information does not eliminate human-machine asymmetry, it reveals it.

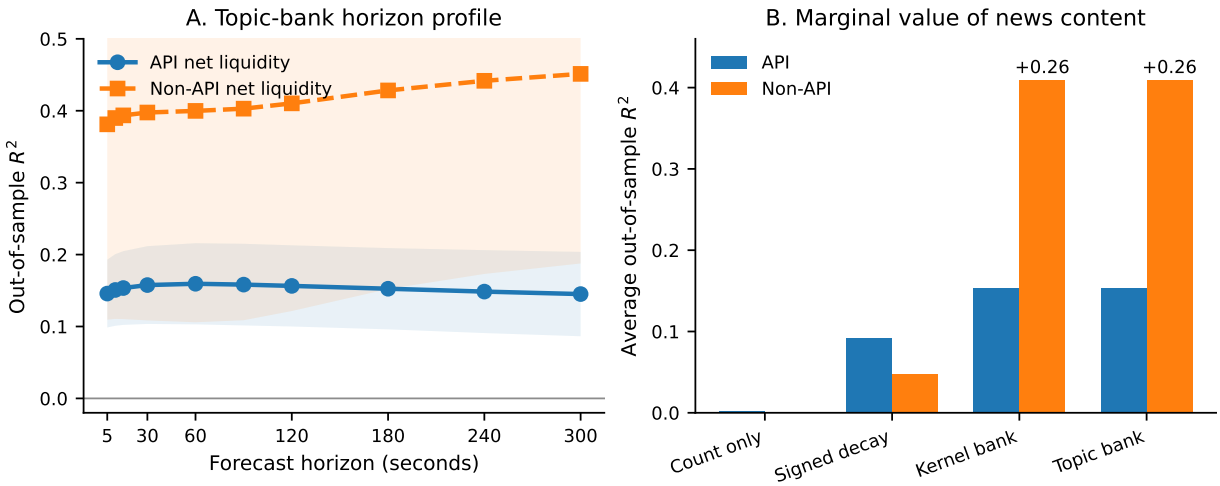


Figure 20. News content and net-liquidity predictability. The left panel reports the out-of-sample R^2 horizon profile for the FinBERT topic-bank specification. The right panel compares average out-of-sample R^2 across feature designs. Content-rich signals dominate count-only news variables, and the improvement is concentrated in human net liquidity.

5.4. Cross-Exchange Information Shocks and the Human-Machine Reaction

Earlier subsections of this section study Coinbase trader reactions to scheduled news events. Those events are clean but rare. Here we turn to a more frequent source of public news that arrives at Coinbase from outside the venue, namely large one-second returns on Binance. For this exercise we supplement the Coinbase feed with one-second BTCUSDT spot-price bars from Binance's official historical archive over the same dates. Binance is the largest venue for BTC by volume over our sample period, and it is used only as an external price reference.

BTC trades on many centralized venues, and Binance is the largest by volume over our sample period. Most prior work on the dynamics linking these venues is at the daily horizon or focused on arbitrage (Makarov and Schoar, 2020; Aleti and Mizrach, 2021; Griffin and Shams, 2020). We look at the venue-level analogue of the cross-population lead-lag function $\Gamma_{H,M}(\ell)$ defined in Section 3.3. Let r_t^V denote the one-second log return on venue $V \in \{B, C\}$, where B = Binance BTCUSDT and C = Coinbase BTC–USD. Define

$$\Gamma_{B,C}(\ell) \;=\; \mathrm{Corr}\big(r_t^B,\, r_{t+\ell}^C\big)\,, \qquad \ell \in \{-5, \ldots, +5\} \text{ seconds}, \tag{33}$$

using the same sign convention as the cross-population $\Gamma_{H,M}$ in Section 3.3: positive ℓ means Coinbase lags Binance.

Estimated on the full January 2021 through June 2022 sample, the function has $\Gamma_{B,C}(0) = 0.47$, $\Gamma_{B,C}(+1) = 0.28$, and $\Gamma_{B,C}(-1) = 0.06$. The asymmetry is sharp (Figure IA.22 in the Internet Appendix plots the full cross-correlation function). Sub-second leads are not resolvable at this sampling rate. A real lead shorter than one second is mapped onto $\ell = 0$, and a lead longer than one second is mapped onto $\ell = +1$. The gap between $\Gamma_{B,C}(+1) = 0.28$ and $\Gamma_{B,C}(-1) = 0.06$ echoes

the cohort-level pattern of automated flow lagging human flow inside Coinbase (Section 3.3): on this venue, the news moving one-second returns has usually shown up at the leading exchange first. The asymmetry is what makes Binance useful for identification. It supplies, hundreds of thousands of times in our sample, a moment at which relevant public information has just arrived for every Coinbase trader simultaneously, an external event clock for reaction speed that no within-venue variable can provide. We use it below to define high-frequency news shocks and to ask which population reacts first.

These shocks complement the news analysis at a much higher frequency, since headlines operate over minutes while cross-exchange moves operate over seconds. The event seconds provide a public event clock external to Coinbase route labels, not exogenous variation in information. We ask whether the machine and human cohorts on Coinbase react differently to the same outside shock, and through which part of the order book.

5.4.1. Extreme Binance seconds as an external event clock

We need moments at which relevant public information has just arrived for every Coinbase trader at once, and extreme one-second Binance returns provide them roughly 940,000 times in our sample. Index seconds by t. A *top-event* second is one in which the Binance log return is in the top 1% of its empirical distribution, and a *bottom-event* second is one in the bottom 1%. Denote the two sets $\mathcal{T}^+$ and $\mathcal{T}^-$, each with approximately 470,000 events.

For each event second t_i, the *local control set* $\mathcal{C}_i$ is the set of non-event seconds in $[t_i - W,\, t_i - B) \cup (t_i + B,\, t_i + W]$, with $W = 900$ seconds (a $\pm$15-minute window) and $B = 60$ seconds (a one-minute buffer on either side of the event). Other event seconds inside that window are removed from $\mathcal{C}_i$. The window is narrow enough that intraday seasonality and the prevailing volatility regime change little between event and control, and the buffer keeps the event itself out of the control mean.

The identifying assumption is the standard event-study assumption that, absent the Binance shock, cohort-g activity on Coinbase in event second t_i would equal its mean over $\mathcal{C}_i$. Because we separate the two cohorts in the next step, this assumption must hold per cohort, not in aggregate. The asymmetry of $\Gamma_{B,C}$ in equation (33) is what makes the assumption plausible: Coinbase order flow in the second *before* the Binance return realizes does not contain enough information about that return to materially shift the Binance tail classification.

The event-time path of cohort activity from $t - 5$ to $t + 5$ is informative about whether this assumption is consistent with the data. Figure 21 plots, for each cohort and each near-book activity measure, the cross-event mean of dollar-notional activity. Because baseline API volume is roughly two orders of magnitude larger than baseline non-API volume, the figure uses a dual vertical axis (API on the left axis, in millions; non-API on the right axis, in thousands) so the percentage response in each cohort is visible.

Three features of Figure 21 matter for the identifying assumption and for what the estimator captures. First, the pre-event window is roughly flat between $t - 5$ and $t - 3$ in both cohorts on all four measures, with a small upward drift at $t - 2$ and $t - 1$ on the with-the-move and taker

measures; with one-second sampling, a Binance return classified in second t is partly in progress by the end of $t-1$, so the cleanest pre-event window is $t-5$ to $t-3$. Second, both cohorts jump sharply at $t=0$ on the with-the-move measures, with roughly similar percentage changes but dollar levels that differ by about a hundredfold; we therefore compare standardized responses rather than raw dollar changes. Third, the reaction is concentrated at $t=0$ and partly decays over $t+1$ to $t+5$, so the contemporaneous estimator captures the peak of a short-lived response rather than a permanent level shift. Off-direction activity is informative in the same way: maker activity on the recovering side rises slightly at $t=0$, while taker activity against the move falls by roughly a quarter to a third, so taking concentrates on the with-the-move side in the event second. Figure IA.23 in the Internet Appendix records the underlying dollar magnitudes.

5.4.2. A matched-control comparison of cohort reactions

The design in simple words is a matched-control event study. For each extreme Binance second we compare each cohort's Coinbase activity in that second with the same cohort's average in nearby quiet seconds, and we then compare the two cohorts' responses. The rest of this subsection makes that precise. We work in side-decomposed components of the cohort-level primitives of Section 3.1. Let $L^{(b)}_{g,t}$ and $L^{(s)}_{g,t}$ denote the buy-side and sell-side notional of passive limit orders posted by cohort $g \in \{\text{API}, \text{NAPI}\}$ at second t, where API is the machine cohort and NAPI the human cohort in the notation of this subsection, restricted to orders within 0.25% of the contemporaneous midprice; these are the side-decomposed components of $L_{g,t}$ defined in Section 3.1. Side-decomposed cancellations $C^{(b)}_{g,t}, C^{(s)}_{g,t}$, maker-side matches $\widetilde{M}^{(b)}_{g,t}, \widetilde{M}^{(s)}_{g,t}$ (column sums of X_t over taker cohort), and taker-side matches $M^{(b)}_{g,t}, M^{(s)}_{g,t}$ (row sums of X_t over maker cohort) are defined analogously. Let Y stand for any one of these side-decomposed per-second activity measures, in dollar notional.

For each event i, each cohort $g \in \{\text{API}, \text{NAPI}\}$, and each measure Y, write the local mean and standard deviation as $\mu^g_i(Y)$ and $\sigma^g_i(Y)$, computed over $\mathcal{C}_i$. The per-event standardized contrast is

$$z^g_i(Y) \;=\; \frac{Y^g(t_i) \;-\; \mu^g_i(Y)}{\sigma^g_i(Y)}, \tag{34}$$

i.e., cohort-g activity in the event second, expressed in units of cohort-g standard deviations from the cohort-g local mean. This is not a dollar effect but a responsiveness measure. We use it specifically because dollar levels are not comparable across cohorts (the API cohort has much larger baseline volume than the non-API cohort, as documented in Section 3.1), so a dollar difference would be dominated by the level gap rather than the reaction.

The per-event difference-in-differences is

$$d_i(Y) \;=\; z^{\text{API}}_i(Y) \;-\; z^{\text{NAPI}}_i(Y), \tag{35}$$

read as "how much more the machine cohort responds than the human cohort, in Z-score units." For cohorts with thin baseline activity (in particular non-API taker activity, which averages a few

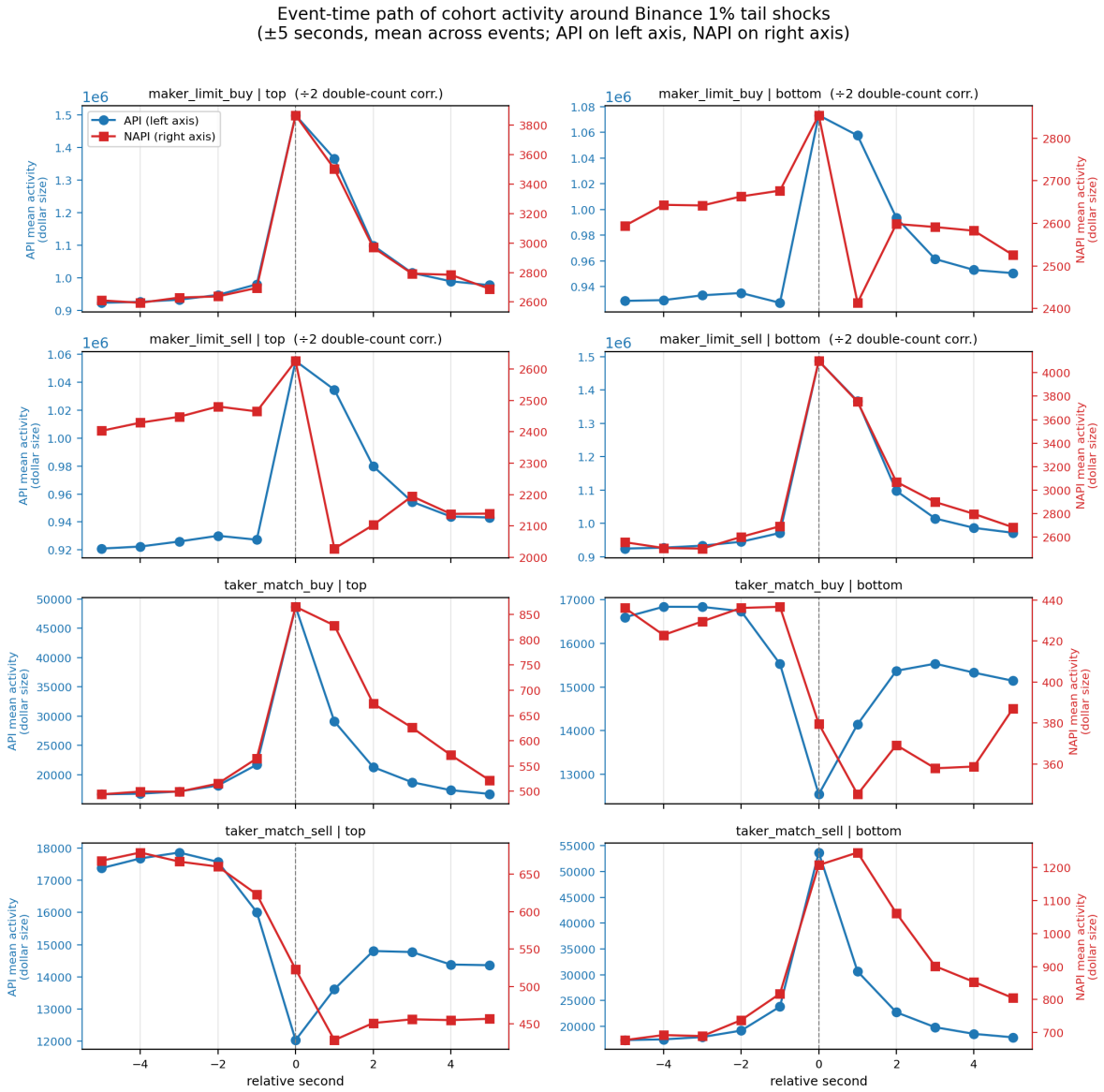


Figure 21. Event-time path of mean Coinbase activity in dollar notional in the eleven seconds around top- and bottom-percentile Binance returns ($t = 0$ is the event second). Rows are activity measures; columns are top versus bottom tail. The API cohort is on the left axis (millions of dollars) and the non-API cohort is on the right axis (thousands of dollars), to make the percentage response visible in both cohorts.

hundred dollars per second), the local control standard deviation $\sigma_i^g(Y)$ can be very small on quiet control windows, and dividing by it produces a per-event $z_i^g(Y)$ in the hundreds. These outliers come from how quiet the control was rather than from the cohort actually reacting hard. To keep them from dominating the cross-event mean, we clip each per-event z-score at $|z| \leq 50$ before

forming the difference and averaging:

$$\hat{D}(Y) \;=\; \frac{1}{N}\sum_{i=1}^{N}\Big[\,\mathrm{clip}_{50}\big(z_i^{\mathrm{API}}(Y)\big) \;-\; \mathrm{clip}_{50}\big(z_i^{\mathrm{NAPI}}(Y)\big)\,\Big], \qquad \mathrm{clip}_c(x) \;=\; \min\big(\max(x,-c),\,c\big), \quad (36)$$

taken over the $N \approx 470{,}000$ events in each tail. At most 0.6% of events get clipped in any cell. Across clipping thresholds in $\{10, 25, 50, 100\}$ the estimates keep their sign and their economic ordering (Table IA.VI in the Internet Appendix). Cancellation and maker-side match estimates barely move, taker-side match estimates remain large and positive at every threshold, and the maker-limit magnitudes decline as the cap is relaxed, so the honest summary is stability of sign and ordering rather than invariance of every coefficient. The uncapped raw mean is not stable, because for lightly traded taker activity a handful of quiet control windows produce per-event z-scores in the hundreds and pull the average toward whichever cohort happens to be near its noise floor. Confidence intervals on $\hat{D}(Y)$ come from resampling the clipped per-event differences with replacement ($B = 2000$ replicates) and reading off the 2.5th and 97.5th percentiles. As a non-parametric cross-check we also report the event-level median $\mathrm{med}\big(d_i(Y)\big)$ and the fraction of events with $d_i(Y) > 0$.

5.4.3. Machines respond first and on the side of the move

Equation (36) gives four headline numbers, one for each near-book measure on the with-the-move side of the relevant tail. The machine cohort responds more than the human cohort on every cell, and the gap is larger on the taker side than on the maker side: $\hat{D}(Y) \approx 0.85$ standard deviations for limit-order placement and $\hat{D}(Y) \approx 1.4$ standard deviations for aggressive taking. On the against-the-move side, the maker gap is roughly half as large and the taker gap is essentially zero. Figure 22 reports the maker estimates as a reaction-gap figure, and Table IA.VI in the Internet Appendix reports the full per-event distribution behind every cell, whose clipped means and unclipped medians agree in sign and approximate magnitude.

The clip matters most on the taker cells, because non-API taker activity in dollar terms is small enough that on some control windows σ_i^{NAPI} is near zero, and the per-event z_i^{NAPI} blows up.[17] Putting the maker, taker, and cancel cells together, the machine cohort's faster response shows up in fresh quoting and aggressive taking on the with-the-move side, and not in pulling quotes or in off-direction activity.

[17]The uncapped raw means are -0.26 on `taker_match_buy` in top-tail events and -3.75 on `taker_match_sell` in bottom-tail events, both with cross-event standard deviations in the hundreds. The clip removes a handful of events whose z_i^{NAPI} is in the hundreds or thousands while leaving the median and the clipped mean essentially unchanged. The cancel-side cells never had this problem, and the bootstrap CI on the unclipped mean includes zero on the Z-score scale despite a large raw-dollar effect, so cancellations move roughly in lockstep across cohorts.

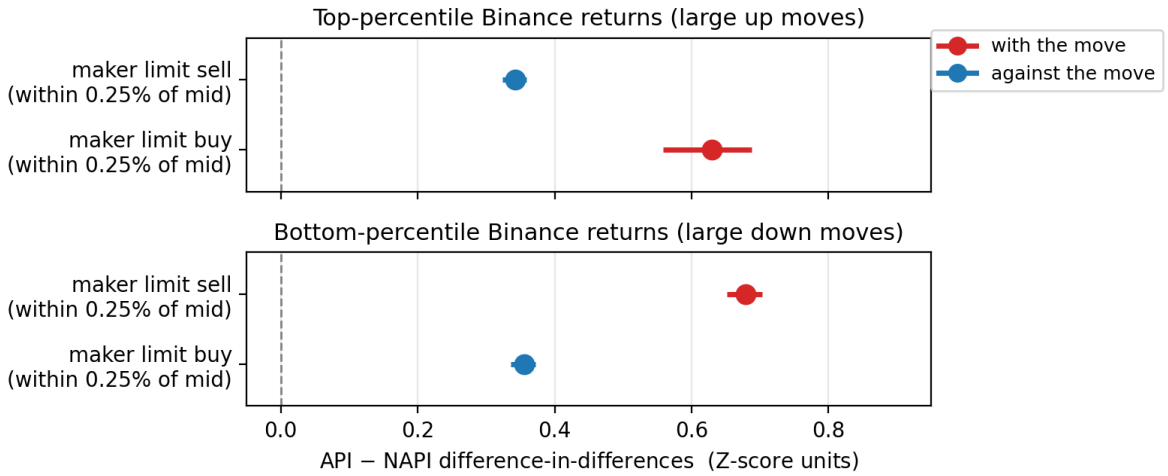


Figure 22. Cohort reaction gaps $\hat{D}(Y)$ for Z-scored per-side maker limit-order activity within 0.25% of the midprice, around top- and bottom-percentile Binance returns. Each point is an estimate and each bar is a 95% bootstrap confidence interval over events. Red points mark the side that is "with the move" for that tail, and blue points mark the opposite side. Positive values mean the machine cohort responds more than the human cohort. The figure uses the raw (unclipped) bootstrap mean, and the clipped estimator of equation (36) moves the with-the-move estimates from 0.63 to 0.88 and from 0.68 to 0.84, with at most 0.25% of events trimmed.

5.4.4. What the cross-exchange evidence adds

We do not claim that cross-exchange shocks are equivalent to the scheduled-news events earlier in this section. They are a different object, more frequent, less obviously tied to a single news release, and identifiable only because Coinbase is not the price-discovery venue at this horizon. Two things this analysis adds. First, a sample big enough ($N \approx 470{,}000$ per tail) to ask whether a cohort gap survives a per-event selectivity check, which the scheduled-event samples are too small for. Second, a setting where "with the move" and "against the move" have a clear meaning, since the move has already happened on the leading venue when the second begins.

The story is consistent across the three channels we can split apart. In the seconds that Binance moves, the machine cohort on Coinbase posts more fresh quotes and takes more aggressively on the side the move favors, the human cohort does the same thing but at much smaller proportional sizes, and neither cohort behaves differently on cancellations or on the off-direction side. The gap is bigger on the taker channel than on the maker channel, which is where the cross-population literature would put it, since aggressive taking is where the information advantage shows up most directly. We read this as evidence that the speed and information advantages of automated traders translate into a different reaction to outside news, with the human cohort going the same way but more slowly and at smaller proportional size.

6. The Cost of Technological Asymmetry

This section shifts the focus from mechanism to incidence. It is the pricing step of the paper's argument, since if slow and predictable trading is exploitable, the cost must appear here. Spoofing-like depth states, ignition-like episodes, and news-driven attention shocks are brief, but they matter when they arrive exactly as investors demand immediacy or leave stale passive interest in the book. The relevant question is not whether the venue is expensive on average. It is who pays when the book enters these states and how large the payment is on the exposure actually at risk. The exercise is therefore disciplined cost accounting rather than a complete welfare calculation. The order book does not reveal private values, routing opportunities, inventory constraints, or the equilibrium trades that would occur under another market design; it does reveal, with unusual precision, actual execution prices, contemporaneous midquotes, passive fills before adverse price moves, and the volatility state that makes it costly to carry inventory through the event window.

The main estimate is a matched-control difference-in-differences estimate of realized execution harm at the event-window level. For each detected spoofing, momentum-ignition, or news window, we compare the post-event change in execution costs with the same change in non-event windows from the same day matched on spread, depth, volatility, volume, order-flow imbalance, and time of day. This is the right first object because the empirical claim is not that the whole venue becomes inefficient. It is that a small set of brief states changes the terms of trade for the human cohort. The resulting rate must therefore be evaluated on the exposure that is actually at risk. In plain terms, we ask what a dollar of human trading cost during a detected event compared with what the same dollar would have cost in normal windows matched on market state. The answer up front is 9.14 basis points gross and 5.58 basis points net on exposed human notional, against a matched-normal benchmark cost of 3.42 basis points, so trading inside these states costs more than triple the normal rate. Repeating the identical construction for machines, with machine costs in the numerator and machine post-event taker exposure in the denominator, yields 1.99 gross and -0.50 net basis points. Event windows are therefore not costless for machines, but the gross burden is 4.6 times larger for humans, and only the human cohort retains a positive net burden after gains and losses are netted. Normalizing by all venue notional instead yields 0.37 basis points, because most trading occurs outside the adverse states. Throughout, investor costs are reported as notional-weighted basis-point rates rather than absolute totals. One basis point is one hundredth of one percent.

The positive-part convention is also part of the incidence definition. If, for instance, a human taker pays an abnormal spread and a fast counterparty earns that spread, the taker's loss is not offset for investor-protection purposes by the counterparty's gain. The gross measure asks how much harm is borne by accounts exposed to adverse execution states; because $x \mapsto x^+$ is nonlinear, it is an incidence total for exposed accounts rather than a net transfer, and it is deliberately an investor-incidence measure rather than a social-surplus measure. The net measure is the sign-preserving aggregate that remains after gains and losses within the window are netted, and it addresses the natural objection that the calculation would otherwise ignore gains to counterparties. We report

both, with gross as the primary investor-incidence measure.

6.1. Cost Estimands and Model Choice

Let e index an event with timestamp t_e, and let $f(e) \in \{\text{spoofing}, \text{momentum ignition}, \text{news}\}$ denote its family. Throughout this section $g \in \{H, M\}$ continues to index the trader cohort, and f indexes the event family. The pre-event and post-event windows are $\mathcal{W}_e^- = [t_e - W, t_e)$ and $\mathcal{W}_e^+ = [t_e, t_e + W)$, with $W = 60$ seconds. For trade i, let p_i be the execution price, v_i the quantity, $m(t_i)$ the contemporaneous midquote, and $q_i \in \{+1, -1\}$ the taker sign. Define the relative taker shortfall

$$r_i^T = q_i \frac{p_i - m(t_i)}{p_i}. \tag{37}$$

For a trade set I, let $D(I) = \sum_{i \in I} p_i v_i$ and $\omega_i = p_i v_i / D(I)$. The notional-weighted taker cost rate is

$$TC(I) = 10^4 \sum_{i \in I} \omega_i r_i^T. \tag{38}$$

Passive fills are priced by a markout.[18] Fix $\tau = 10$ seconds and define

$$r_i^M(\tau) = \begin{cases} \{p_i - m(t_i + \tau)\}/p_i, & \text{maker buy}, \\ \{m(t_i + \tau) - p_i\}/p_i, & \text{maker sell}. \end{cases} \tag{39}$$

The maker cost rate is $MC(I) = 10^4 \sum_{i \in I} \omega_i r_i^M(\tau)$. Let TC_e^-, TC_e^+, MC_e^-, and MC_e^+ denote pre- and post-event rates. The local pre-window changes are

$$\Delta TC_e = TC_e^+ - TC_e^-, \qquad \Delta MC_e = MC_e^+ - MC_e^-. \tag{40}$$

The matched-control estimator subtracts the same post-minus-pre change in K non-event control windows $c(e,1), \ldots, c(e,K)$:

$$\begin{aligned} \Delta TC_e^{\text{DiD}} &= \Delta TC_e - \frac{1}{K} \sum_{k=1}^{K} \Delta TC_{c(e,k)}, \\ \Delta MC_e^{\text{DiD}} &= \Delta MC_e - \frac{1}{K} \sum_{k=1}^{K} \Delta MC_{c(e,k)}. \end{aligned} \tag{41}$$

The identifying assumption is that, absent the event, event windows would have experienced the same post-minus-pre drift in cost rates as their matched controls, conditional on the matching variables. Internet Appendix Section IA.G.2 states this conditional-comparability condition formally. Section 6.3 reports estimates under both the matched-control counterfactual and the local pre-window baseline.

[18] Taker outcomes are priced at execution because the taker chose the moment of trade, so the slippage paid at that moment is the taker's cost of immediacy. Post-trade price drift enters through the maker side of the same trades, which avoids counting the same price path twice.

For each event, define gross and net rate changes

$$h_e^{\text{gross}} = (\Delta TC_e^{\text{DiD}})^+ + (\Delta MC_e^{\text{DiD}})^+, \tag{42}$$

$$h_e^{\text{net}} = \Delta TC_e^{\text{DiD}} + \Delta MC_e^{\text{DiD}}. \tag{43}$$

For an event family $\mathcal{E}_j$, the reported basis-point estimands are notional-weighted averages,

$$H_j^a = \frac{\sum_{e\in\mathcal{E}_j} D_e^+ h_e^a}{\sum_{e\in\mathcal{E}_j} D_e^+}, \qquad a \in \{\text{gross}, \text{net}\}. \tag{44}$$

Gross harm is monotone in harmful changes to either execution component. Net harm can be smaller, and can be negative, because it offsets gains against losses within the window. The two rates answer different questions. Gross is the investor-incidence measure, while net describes the transfer-adjusted aggregate.

For the cohort comparison, every object is indexed by the cohort $g \in \{H, M\}$. The cohort-specific estimand is

$$H_{j,g}^a = \frac{\sum_{e\in\mathcal{E}_j} D_{e,g}^+ h_{e,g}^a}{\sum_{e\in\mathcal{E}_j} D_{e,g}^+}, \qquad a \in \{\text{gross}, \text{net}\}. \tag{45}$$

Both the numerator and the denominator therefore change with the cohort, so the machine estimate is a full reconstruction from machine taker slippage, machine maker markouts, and machine post-event taker exposure under the same event and matched-control definitions, not a rescaling of the human rate.

The choice among candidate cost estimators matters because each measures something different, and Internet Appendix Section IA.G.1 records the full comparison. Implementation shortfall and signed slippage are the closest trade-level measures because they price actual executions relative to the contemporaneous midquote (Perold, 1988; Cartea et al., 2015; O'Hara, 2015). Effective spreads and realized spreads are transparent market-quality rates, but unweighted rates do not reveal how much exposure is concentrated in adverse states. Markouts capture passive adverse selection, which is essential here because a stale human maker can experience a costly fill even when average taker slippage is modest. Kyle-λ is useful as a mechanism diagnostic, but it is not the headline rate estimator because a short-window impact slope can be weakly identified and does not incorporate the realized exposure that traded. A structural welfare model would require private valuations and equilibrium order-submission responses under counterfactual rules. This paper instead estimates notional-weighted incidence in observed event windows.

6.2. Normal-Period Benchmarks and Denominators

Table VI reports taker-slippage benchmarks for human trades. These are means of trade-level slippage rates, so they answer a per-trade execution question. Momentum ignition is the only event type with a large average taker-slippage spike. Spoofing and news are much more muted on this margin, which is why the cost accounting includes passive markouts and event tails rather than

Table VI Human Taker-Slippage Benchmarks: Event Windows versus Matched Normal Windows

Event type	Normal (bps)	Event (bps)	Abnormal (bps)	Peak 0–10s (bps)	Exposure share (%)
Spoofing	2.84	2.85	0.01	0.01	46.9
Momentum ignition	2.91	2.99	0.08	0.50	21.9
News	2.72	2.73	0.01	0.01	31.2
Any event	2.82	2.89	0.07	0.50	100.0

Notes. Normal windows are matched on pre-event spread, top-of-book depth, realized volatility, notional, order-flow imbalance, and UTC time of day. Abnormal slippage is event slippage minus the matched normal benchmark. Exposure share is each event family's share of post-event human taker notional. The table reports means of trade-level human taker slippage rates.

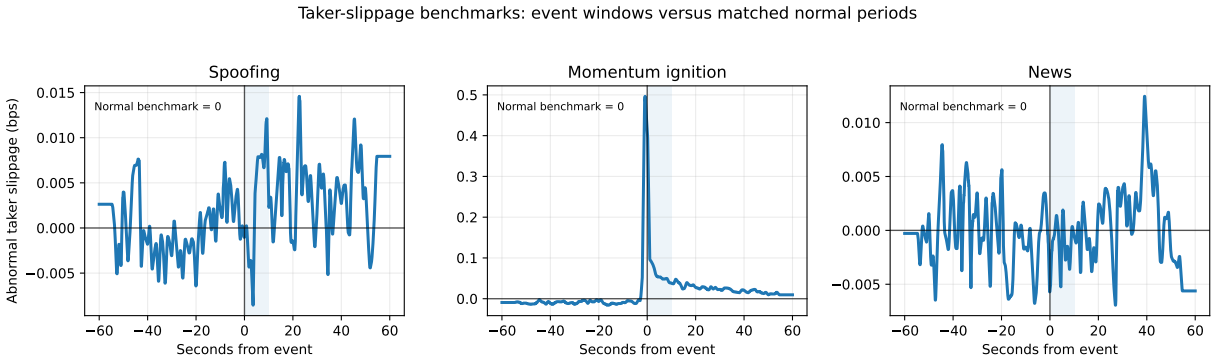


Figure 23. Taker-slippage benchmarks around event windows. Each panel reports abnormal taker slippage in basis points relative to matched normal windows. The horizontal zero line is the normal-period benchmark, the vertical line is event time, and the shaded region marks the first ten seconds after the event. Momentum ignition generates the clear event-time taker-slippage spike. Spoofing and news are more visible in markouts, repeated small deteriorations, and right-tail episodes.

relying only on average taker slippage. Table IX quantifies this division of channels.

The denominator matters. The measured gross harm is 0.37 basis points when mechanically spread over all executed venue notional, but 9.14 basis points when normalized by post-event human taker notional inside detected event windows. That exposure represents only 4.0% of total venue notional, yet it is the exposure that bears the timing risk measured by the design. Table VII and Figure 24 make this concentration explicit.

Table VII places the 9.14-basis-point estimate against a matched normal-window benchmark computed with the same notional weights. The event-window and normal-window rates are therefore directly comparable. The event-window gross rate is 12.56 basis points, the matched normal rate is 3.42 basis points, and their difference is 9.14 basis points.

Table VIII reports the cohort comparison using a cohort-specific normalization. For each population, the numerator contains that population's matched-control taker-slippage and maker-

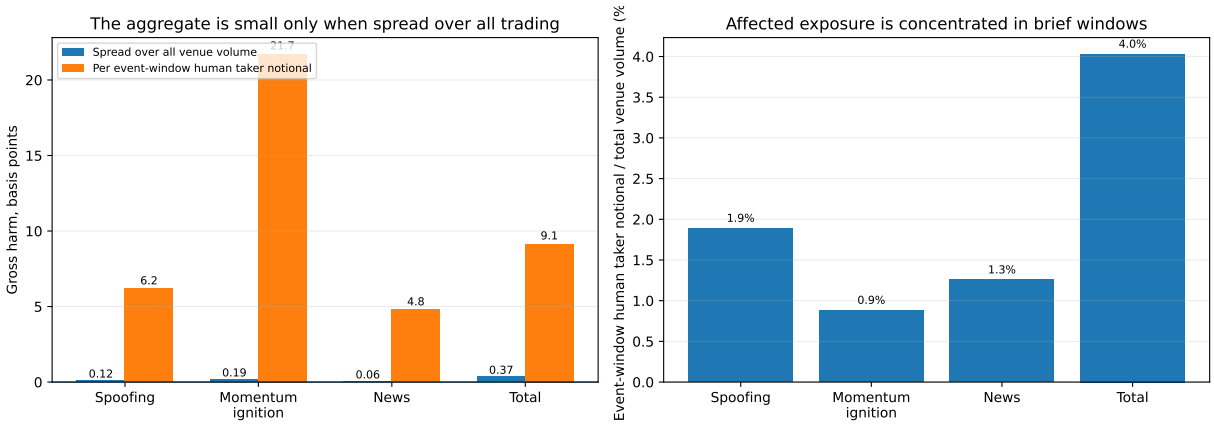


Figure 24. Denominators for investor-cost interpretation. The left panel reports gross harm in basis points when spread over all venue volume and when scaled by event-window human taker notional. The right panel reports the share of total venue volume represented by event-window human taker notional. The aggregate cost is small only under the all-volume denominator. It is materially larger on the exposure actually bearing the timing risk.

Table VII Same-Denominator Gross-Cost Benchmark

Event type	Exposure share (%)	Matched normal (bps)	Event window (bps)	Abnormal DiD (bps)
Spoofing	46.9	3.10	9.27	6.17
Momentum ignition	21.9	4.40	26.10	21.70
News	31.2	2.90	7.68	4.78
Any event	100.0	3.42	12.56	9.14

Notes. Rates are notional-weighted averages over post-event human taker exposure. The abnormal column equals the event-window gross rate minus the matched normal-window rate under the difference-in-differences construction. Exposure shares sum to the pooled event-window denominator.

markout costs, while the denominator is that population's own post-event taker exposure. The machine denominator is therefore not the human denominator rescaled by relative activity. This distinction is essential because machines trade much more notional inside the same event families.

The comparison answers two distinct questions. First, adverse event windows raise the machine gross rate by 1.99 basis points, so these states are not costless for the fast cohort. Second, the human gross rate is 7.16 basis points larger and 4.61 times as high. The sign-preserving comparison is sharper still. Human net abnormal cost is 5.58 basis points, whereas the machine net estimate is slightly negative. The distributional result therefore does not arise because all traders face a generally expensive state. It is concentrated on the slower cohort even when each population is measured on its own exposure.

Table VIII Cohort-Specific Matched-Control Cost Rates

Cohort	Matched normal (bps)	Event window (bps)	Gross abnormal (bps)	Net abnormal (bps)	Gross ratio relative to machine
Human	3.42	12.56	9.14	5.58	4.61
Machine	0.42	2.41	1.99	-0.50	1.00

Notes. The human row is the paper's canonical matched-control estimate. The machine row applies the same trade-level cost definitions, matched-control subtraction, component-level positive-part convention, and pooled normalization to machine-specific taker and maker outcomes, using machine post-event taker exposure in the denominator. The machine gross estimate has a 14-day moving-block-bootstrap 95% interval of [1.78, 2.22] basis points, and the machine net estimate has an interval of [-0.78, -0.27] basis points. The gross ratio is computed from unrounded estimates, $9.14/1.982972 = 4.61$.

6.3. Realized Execution Harm

This subsection converts the detections into realized per-dollar costs on actually executed human volume, through two channels. The taker channel prices aggressive human orders that pay spread and impact during events, and the maker channel prices resting human orders that are filled just before prices move against them. Table X reports the primary matched-control estimates. Gross abnormal harm is 9.14 basis points across 118,097 detected event windows. The sign-preserving net estimate is 5.58 basis points. The corresponding machine estimates are 1.99 basis points gross and -0.50 basis points net, so the human gross burden is 4.6 times the machine burden when each population is measured on its own exposure. The positive net rate shows that the result is not generated solely by the refusal to offset gains and losses. Gross is larger because it measures losses borne by exposed accounts, while net is smaller because gains to counterparties offset part of those losses.

The gross rate operates through two distinct channels. Table IX separates taker slippage from maker markouts using the component-level positive-part identity. For spoofing, only 0.35 of the 6.17 abnormal basis points comes from taker slippage, while 5.81 basis points comes from adverse maker markouts. For news, the corresponding contributions are 0.31 and 4.48 basis points. Momentum ignition is different. Taker slippage contributes 14.57 of its 21.70 abnormal basis points, consistent with a mechanism that pulls human takers into an expensive liquidity state.

Accumulated over calendar time, the pooled matched-control gross rate rises persistently to 9.14 basis points, with large steps around spring 2021 and late 2021 (Figure IA.24 in the Internet Appendix). The path is not a one-day artifact, although tail days matter disproportionately.

The distribution is sharply right-tailed. The top 1% of event contributions accounts for 52.6% of the pooled gross rate, while the top ten days account for 15.6%. The cost is therefore not a smooth surcharge on every trade. It is a state-contingent deterioration that becomes economically important when human traders demand immediacy or leave stale passive exposure during brief intervals of severe book stress.

Figure IA.26 in the Internet Appendix compares the estimators per event: the local baseline and the matched-control DiD agree on the ranking of harmful episodes, while Kyle-λ is nearly

Table IX Basis-Point Decomposition of Matched-Control Gross Harm

Event type	Taker slippage (bps)	Maker markout (bps)	Total gross (bps)	Taker share (%)	Dominant channel
Spoofing	0.35	5.81	6.17	5.7	Passive markout
Momentum ignition	14.57	7.14	21.70	67.1	Taker slippage
News	0.31	4.48	4.78	6.5	Passive markout
Total	3.46	5.69	9.14	37.8	Mixed

Notes. Each channel is the notional-weighted average of the positive matched-control DiD rate within the event family. Event-family rows use the corresponding human event-window notional. The total row uses pooled human event-window notional. Channel contributions add exactly, up to rounding, to the gross rate in each row. The event-family counts sum to 118,163; 66 ignition detections lacking a complete pre-window, post-window, ten-second markout, or valid human cost component are excluded, leaving 118,097 cost-eligible windows. Overlapping windows from different families retain their family labels, and maximal clusters of overlapping windows are resampled jointly for inference.

Table X Matched-Control Execution Harm in Basis Points

Event type	N	Gross DiD (bps)	Net DiD (bps)	Local baseline gross (bps)	Matched normal (bps)	Event window (bps)
Spoofing	70,826	6.17	3.46	5.96	3.10	9.27
Momentum ignition	24,465	21.70	15.12	21.04	4.40	26.10
News	22,872	4.78	2.06	5.08	2.90	7.68
Total	118,097	9.14	5.58	8.99	3.42	12.56

Notes. Gross and net DiD correspond to Equations (42)–(44). Local baseline uses the pre-event window without matched-control subtraction. Matched normal and event-window rates use identical notional weights. The total row uses pooled human event-window notional.

orthogonal to realized harm. That is informative rather than embarrassing. In a short window, realized harm can arise through widened spreads, toxic passive fills, or a surge in post-event human notional even when a linear flow-impact slope is poorly estimated.

6.4. Market-Phase Heterogeneity

This subsection asks whether the cost machinery runs differently in bull and bear markets, and the headline answer is that bear markets hurt through frequency rather than severity. To split the sample we need a bull-bear classification that does not rely on hand-picked dates, and a split based only on one-day returns is not adequate because a sustained selloff can contain frequent positive-return days. We therefore estimate a latent market-phase model from a vector that distinguishes short-horizon direction from medium-horizon trend and stress. Let P_t be the Coinbase BTC–USD daily close and define

$$Z_t = (r_{t,5},\ r_{t,20},\ d_{t,60},\ \sigma_{t,20})', \qquad r_{t,h} = \log P_t - \log P_{t-h}, \tag{46}$$

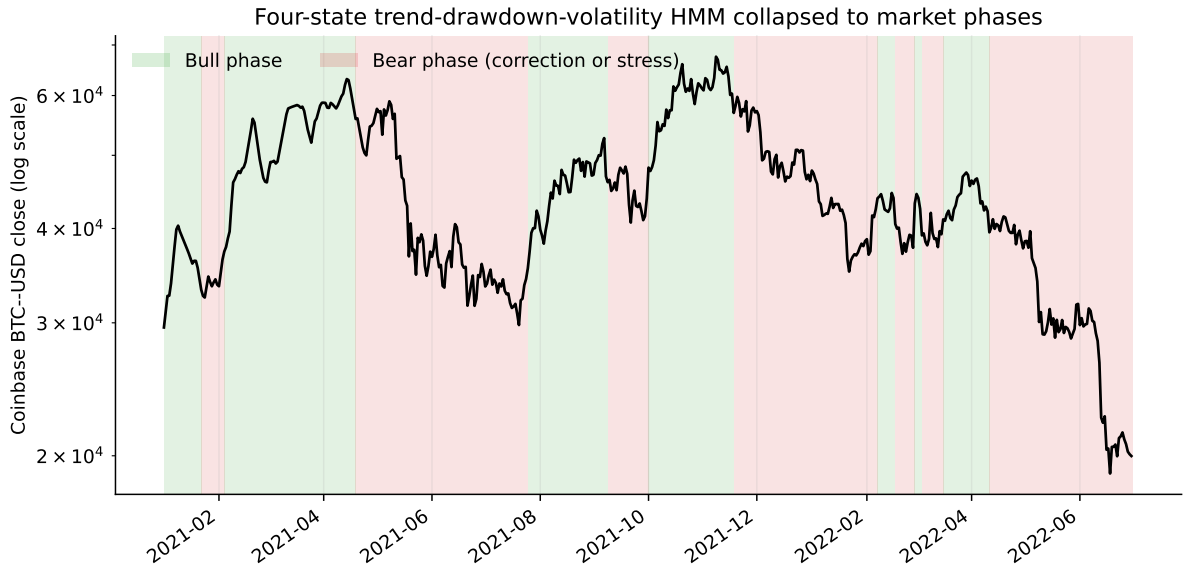


Figure 25. Coinbase BTC–USD market phases from the four-state hidden Markov model. Green regions combine strong-advance and steady-advance states. Red regions combine correction and stress states. The classifier uses multi-horizon returns, trailing drawdown, and realized volatility, which prevents temporary positive-return days from relabeling a sustained selloff as a bull market.

where $d_{t,60} = \log P_t - \max_{0 \le j \le 59} \log P_{t-j}$ is the drawdown from the trailing 60-day high and $\sigma_{t,20}$ is 20-day realized volatility. The features are centered by their sample medians and scaled by median absolute deviations. A diagonal-Gaussian hidden Markov model is estimated with persistent transitions in the tradition of regime-switching models (Hamilton, 1989). We compare specifications with two, three, and four states using two criteria, BIC and an expanding blocked out-of-sample log score, because BIC alone can be unreliable for hidden Markov order selection. Both criteria select four states, and we use the four-state partition throughout. The fits with two and three states enter only the model selection, so we do not report cost results under them. The two advancing states are combined into a bull phase, while correction and stress are combined into a bear phase. Internet Appendix Section IA.G.5 reports the full selection exercise, state parameters, and filtered-probability diagnostics.

Figure 25 shows the resulting market phases. Unlike a return-only two-state model, the classifier identifies the prolonged 2021 drawdowns and the 2022 decline as bear periods even when individual days have positive returns. The smoothed classification is used for descriptive partitioning. A filtered classification based only on information available through date t agrees with it on 97.8% of sample days, indicating that the main phase boundaries are not artifacts of ex post smoothing.

For event family $j \in \{\text{spoofing}, \text{momentum ignition}, \text{news}\}$ and phase $q \in \{\text{bull}, \text{bear}\}$, define the phase contribution

$$H_{j,q}^{\text{gross}} = \frac{\sum_{e \in \mathcal{E}_j} D_e^+ h_e^{\text{gross}} \mathbf{1}\{S_{d(e)} = q\}}{\sum_{e \in \mathcal{E}_j} D_e^+}. \tag{47}$$

Table XI Execution-Harm Rates across Bull and Bear Market Phases

Event type	Bull contribution (bps)	Bear contribution (bps)	Bear share of harm	Bear/bull active-day ratio [95% CI]
Spoofing	2.37	3.79	61.5%	1.10 [0.88, 1.41]
Momentum ignition	9.69	12.02	55.4%	0.99 [0.74, 1.33]
News	2.08	2.71	56.6%	1.00 [0.77, 1.34]
Total	3.88	5.26	57.5%	–

Notes. Event-family rows normalize by the corresponding event-family human exposure. The total row uses pooled human event-window notional. The active-day ratios are invariant to this normalization. Confidence intervals use a 14-day circular moving-block bootstrap with 5,000 replications. Bear phases comprise 58.2% of sample days.

The dated series is reconciled before partitioning. Every bull-bear pair therefore sums to the corresponding event-family rate in Table IX. The pooled bull and bear contributions sum to 9.14 basis points.

Bear phases contribute 5.26 of the pooled 9.14 basis points, or 57.5% of measured harm, but they also account for 58.2% of sample days. The aggregate split therefore does not establish that an otherwise identical event is more costly in a bear market. The bear-to-bull active-day ratios are 1.10 for spoofing, 0.99 for momentum ignition, and 1.00 for news, and every moving-block-bootstrap interval contains one. Conditional cost severity is therefore statistically similar across phases.

Momentum ignition displays a separate incidence margin. The detector finds 7,641 events over 228 bull days and 16,824 events over 318 bear days, corresponding to 33.51 and 52.91 events per day.[19] The bear-to-bull event-rate ratio is 1.58. Thus, the larger bear-period momentum-ignition total reflects a substantially higher frequency of ignition states rather than a reliably larger measured cost conditional on an active day. The phase analysis is descriptive rather than causal. The latent state summarizes recent prices and volatility and may proxy for leverage, liquidity demand, or participation. Its value is to distinguish when harm accumulates from how severe a given active-cost day is.

6.5. Volatility, Risk Capital, and Participation Capacity

Execution harm prices trades, but it is not the whole burden. Events raise short-run volatility, volatility consumes risk capital, and a trader with a fixed risk budget must shrink position size when the risk per dollar rises. This subsection prices that second channel. Let $r_s = \Delta \log m_s$ be one-second midquote returns inside a 60-second window, and define the standard high-frequency volatility measures, realized variation, bipower variation, and jump variation (Barndorff-Nielsen

[19]The phase day counts partition all 546 calendar days. The three unprocessed days of footnote 3 contain no detected events, so including them in the denominators can only slightly understate the per-day frequencies and does not affect the interpretation of the ratio.

Table XII Incremental Expected-Shortfall Harm from Event Volatility

Event type	N	Mean	Median	P95	P99
		Basis points per event			
Spoofing	70,826	2.17	0.00	10.20	38.00
Momentum ignition	24,465	5.49	0.65	26.40	87.50
News	22,872	2.89	0.00	13.70	50.20
Total	118,097	2.99	0.00	14.10	58.90

Notes. Expected-shortfall harm uses $\alpha = 0.99$, $h = 300$ seconds, and the RV-DiD counterfactual. Every entry is an event-level relative loss rate. The table deliberately omits cross-event aggregation.

and Shephard, 2004, 2006),

$$RV = \sum_{s=1}^{W} r_s^2, \qquad BV = \frac{\pi}{2} \sum_{s=2}^{W} |r_s||r_{s-1}|, \qquad JV = (RV - BV)^+. \tag{48}$$

For event e, the volatility shock is the matched-control DiD change in RV. For confidence level α and holding horizon $h = 300$ seconds,

$$\sigma_{e,h} = \sqrt{\frac{RV_e}{W} h}. \tag{49}$$

Expected shortfall is the average relative loss in the worst outcomes, here the worst 1%. We compute it under a local Gaussian approximation, treating short-horizon returns as normal with the variance just estimated for the window. This is a transparent first-order device, and because short-horizon crypto returns have fatter tails than the normal distribution, it understates the true expected shortfall, so the reported risk layer is conservative. The expected-shortfall rate is

$$ES_\alpha^{\text{bps}}(\sigma_{e,h}) = 10^4 \sigma_{e,h} \frac{\varphi(z_\alpha)}{1-\alpha}, \qquad z_\alpha = \Phi^{-1}(\alpha). \tag{50}$$

The event-level risk-capital harm is

$$H_e^{\text{ES,bps}} = \left[ES_\alpha^{\text{bps}}(\sigma_{e,h}^+) - ES_\alpha^{\text{bps}}(\sigma_{e,h}^{\text{cf}}) \right]^+. \tag{51}$$

We set $\alpha = 0.99$ and $h = 300$ seconds. The resulting statistic is a scale-free tail-loss rate in basis points, directly comparable across events without introducing an arbitrary reference exposure.

Volatility also affects participation. Suppose an investor or liquidity supplier has an expected-shortfall budget B and must choose notional N satisfying $ES_\alpha(N, \sigma_h) \le B$. Since expected shortfall is linear in N and σ_h, the maximum admissible notional is

$$N^{\max}(\sigma_h) = \frac{B}{\kappa_\alpha \sigma_h}, \qquad \kappa_\alpha = \frac{\varphi(z_\alpha)}{1-\alpha}. \tag{52}$$

Table XIII Composition of Representative Right-Tail Harm Days

Date	MI share (%)	Spoofing share (%)	News share (%)	Interpretation
2021-05-19	11.4	44.0	44.6	Spoofing and news co-occur, and the largest episode is maker-markout intensive.
2021-04-22	79.2	17.6	3.1	Ignition dominates through large taker-slippage episodes.
2021-05-12	82.1	6.3	11.6	A single ignition episode contributes a large share of the daily rate.
2021-11-03	97.6	1.5	0.8	The right-tail day is almost entirely an ignition-state immediacy shock.
2021-10-06	11.4	14.3	74.4	News dominates and passive orders bear most of the adverse markout.

Notes. The table reports event-family shares for selected days from the right tail of matched-control gross harm. It emphasizes composition rather than absolute magnitude. Shares may differ from 100% by 0.1 percentage point because of rounding.

The positive participation-capacity loss is

$$L_e^{\text{cap}} = \left[1 - \frac{N^{\max}(\sigma_{e,h}^{+})}{N^{\max}(\sigma_{e,h}^{\text{cf}})}\right]^{+} = \left[1 - \frac{\sigma_{e,h}^{\text{cf}}}{\sigma_{e,h}^{+}}\right]^{+}. \tag{53}$$

This channel transforms volatility into lower feasible depth. A risk-constrained maker can leave the spread unchanged and still reduce size when event volatility raises the capital required to support inventory. Figure IA.30 in the Internet Appendix plots the capacity losses and the leveraged liquidation amplification.

6.6. Tail Days, Co-Occurrence, and Liquidation Stress

Because harm is so heavily right-tailed, the natural final questions are whether the bad states cluster in time and whether they coincide with the forced-liquidation pressure that the industry independently recognizes as dangerous. The distribution is extremely concentrated. The top 1% of events account for 52.6% of aggregate gross execution harm. The top ten days account for 15.6% of aggregate harm despite representing only 1.8% of calendar days with nonzero harm (Figure IA.29 in the Internet Appendix). Average spreads and average volume are therefore incomplete investor-protection statistics. The economically relevant state is the tail, where strategic order-book conduct, public information, and volatility interact.

Finally, we use known BTC liquidation dates as an external validation sample for human forced-trading stress. The liquidation analysis is not added to the headline basis-point estimate because liquidation windows are a separate event definition. It is useful because it isolates days on which retail leverage and forced execution are known to be salient. Across 543 processed trading days, the sample contains 16 known liquidation days and 158 detected liquidation-like stress windows. To locate liquidation-style pressure we build two composite stress scores. Each component is

an independent symptom of one-sided forced flow, and the equal-weighted sum of standardized components is the simplest transparent aggregate, with no fitted weights to overfit. Directional stress is measured with

$$S_t^- = z(-r_t) + z(F_t^{\text{sell}}) + z(-I_t) + z(s_t) + z(-D_t) + z(RV_t), \tag{54}$$

$$S_t^+ = z(r_t) + z(F_t^{\text{buy}}) + z(I_t) + z(s_t) + z(-D_t) + z(RV_t). \tag{55}$$

Here r_t is the midquote return, F_t is taker notional, I_t is signed flow imbalance, s_t is the spread, D_t is depth, and RV_t is realized variance. The sell-side score S_t^- captures long-liquidation pressure, while S_t^+ captures short-squeeze pressure. The detected windows are nearly balanced, with 80 forced-sell and 78 forced-buy windows, but forced sells dominate the most severe tail. Median duration is about eight minutes, and the longest window is roughly one hour. Figure IA.31 in the Internet Appendix plots the detected windows on known liquidation dates. Forced-trading stress is clustered and right-tailed, the same distributional structure observed in the manipulation and news cost estimates.

6.7. Interpretation

The central property of the cost estimates is concentration, and concentration is the finding rather than a caveat. The cohort comparison adds a second dimension of concentration, since the same event families cost humans 9.14 gross basis points and machines 1.99. The cost of technological asymmetry is not spread thinly over all trading. It concentrates on the human dollars that trade inside the adverse states, where it more than triples the normal cost of trading, and it is right-tailed, so average market-quality statistics miss it by construction. That is the paper's distributional thesis realized in prices. The estimates measure incidence in the observed market and do not prove that a different regulatory regime would have produced a different equilibrium.

The market-phase partition adds an incidence result rather than a new causal estimand. Bear phases contain 57.5% of measured harm and 58.2% of sample days, while conditional active-day severity is statistically indistinguishable across phases. Momentum-ignition states are nevertheless 1.58 times as frequent per day in bear phases, which helps explain when the aggregate burden accumulates.

The estimates are conservative along several dimensions. They use one venue, a finite event set, a short 60-second accounting window, and no extrapolation to hidden costs from unfilled orders, delayed execution, or cross-venue spillovers. They also do not value the welfare loss from reduced participation after adverse events. Three internal-validity threats deserve explicit mention, together with their likely direction. Detector false positives dilute the event set with normal windows and, under the positive-part convention, bias gross harm upward. The short 60-second accounting window misses delayed execution costs and biases harm downward. Matched controls drawn from the

same day can sit inside clustered stress periods, which absorbs part of the event effect into the counterfactual and also biases the estimates downward. The reported gross and net measures, the side-by-side baseline and matched-control estimators, and the per-event distributions are designed to make these forces visible rather than compress them into a single aggregate rate. At the same time, the gross convention is intentionally not a social-surplus measure. It is an investor-incidence measure. That distinction is the point. The paper asks whether market design predictably reallocates execution quality and risk capacity away from slower investors in the moments when those investors most need liquidity. An institutional ablation in Internet Appendix Section IA.H, built around Coinbase's March 2022 maker–taker fee revision, provides an external check on the cohort margin: the human proxy contracts on both sides of the market when the venue changes the direct price of immediacy, with no detectable rotation toward passive supply.

7. Conclusion

This paper studies a modern limit order book in which the same public state is not the same economic state for all traders. Its thesis is a chain. Human trades are slow and predictable, that combination is exploitable, and we price what it costs. Human and machine participants observe the book, but they do not transform the book into action at the same speed or with the same control over cancellation, queue position, and routing. That difference matters for liquidity measurement. It also matters for manipulation, news, and investor costs.

The contributions can be grouped into three layers. The first is measurement. Order-resolved data allow the paper to reconstruct the book, separate human and machine order flow, and show that canonical limit-order-book facts survive in crypto but acquire a new distributional meaning when the trader population is decomposed. The second is mechanism. Spoofing-like states operate through displayed-depth distortion, while momentum ignition operates through aggressive-flow acceleration and follow-through. The two mechanisms are related but not the same, and the FPY depth-curve representation makes their difference visible in the geometry of the book. The third is incidence. News, spoofing, and ignition do not merely predict returns or volumes. They change who demands immediacy and who protects liquidity. The direct cohort comparison makes the incidence statement quantitative. Gross execution cost is 9.14 basis points for humans and 1.99 for machines, and after netting only the human burden remains positive.

The phase analysis sharpens the incidence result. Bear markets do not make each adverse event detectably more severe, but they contain substantially more momentum-ignition states, so technological costs accumulate more often when the broader market is already under stress.

The central lesson is that average liquidity is not enough. A market can have tight spreads in ordinary time and still impose meaningful costs on slower traders in the short states when liquidity is most valuable. In that sense, technological asymmetry is part of the price of the trade itself, and it is the slower side of the market that pays that price, precisely in the moments when trading matters most.

REFERENCES

Deepali Aggarwal and Zarnash Khan. The Spoofing Puzzle: Deciphering Market Manipulation. Master's thesis, Stockholm Business School, Stockholm, 2024. URL https://su.diva-portal.org/smash/get/diva2:1875413/FULLTEXT01.pdf.

Yacine Aït-Sahalia and Celso Brunetti. High frequency traders and the price process. *Journal of Econometrics*, 217(1):20–45, July 2020. ISSN 0304-4076. doi: 10.1016/j.jeconom.2019.11.005. URL https://www.sciencedirect.com/science/article/pii/S0304407619302428.

Yacine Ait-Sahalia, Jianqing Fan, Lirong Xue, and Yifeng Zhou. How and When are High-Frequency Stock Returns Predictable? (4095405), April 2022. doi: 10.2139/ssrn.4095405. URL https://papers.ssrn.com/abstract=4095405.

Saketh Aleti and Tim Bollerslev. News and Asset Pricing: A High-Frequency Anatomy of the SDF. *The Review of Financial Studies*, page hhae019, May 2024. ISSN 0893-9454. doi: 10.1093/rfs/hhae019. URL https://doi.org/10.1093/rfs/hhae019.

Saketh Aleti and Bruce Mizrach. Bitcoin spot and futures market microstructure. *Journal of Futures Markets*, 41(2):194–225, 2021. ISSN 1096-9934. doi: 10.1002/fut.22163. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/fut.22163.

Dogu Araci. FinBERT: Financial sentiment analysis with pre-trained language models. *arXiv preprint arXiv:1908.10063*, 2019.

Kerry Back, Kevin Crotty, and Tao Li. Identifying Information Asymmetry in Securities Markets. *The Review of Financial Studies*, 31(6):2277–2325, June 2018. ISSN 0893-9454. doi: 10.1093/rfs/hhx133. URL https://doi.org/10.1093/rfs/hhx133.

Brad M. Barber and Terrance Odean. Trading is hazardous to your wealth: The common stock investment performance of individual investors. *The Journal of Finance*, 55(2):773–806, 2000. doi: 10.1111/0022-1082.00226.

Brad M. Barber, Yi-Tsung Lee, Yu-Jane Liu, and Terrance Odean. Just how much do individual investors lose by trading? *The Review of Financial Studies*, 22(2):609–632, 2009.

Ole E. Barndorff-Nielsen and Neil Shephard. Power and bipower variation with stochastic volatility and jumps. *Journal of Financial Econometrics*, 2(1):1–37, 2004.

Ole E. Barndorff-Nielsen and Neil Shephard. Econometrics of testing for jumps in financial economics using bipower variation. *Journal of Financial Econometrics*, 4(1):1–30, 2006.

Jean-Noel Barrot, Ron Kaniel, and David Sraer. Are retail traders compensated for providing liquidity? *Journal of Financial Economics*, 120(1):146–168, April 2016. ISSN 0304-405X. doi: 10.1016/j.jfineco.2016.01.005. URL https://www.sciencedirect.com/science/article/pii/S0304405X16000064.

Silvia Bartolucci and Andrei Kirilenko. A model of the optimal selection of crypto assets. *Royal Society Open Science*, 7(8):191863, August 2020. doi: 10.1098/rsos.191863. URL https://royalsocietypublishing.org/doi/full/10.1098/rsos.191863.

Michele Basseville and Igor V. Nikiforov. *Detection of Abrupt Changes: Theory and Application*. Prentice Hall, 1993.

Bruno Biais, Pierre Hillion, and Chester Spatt. An empirical analysis of the limit order book and the order flow in the paris bourse. *The Journal of Finance*, 50(5):1655–1689, 1995. doi: 10.1111/j.1540-6261.1995.tb05192.x.

Jean-Philippe Bouchaud, Julius Bonart, Jonathan Donier, and Martin Gould. *Trades, Quotes and Prices: Financial Markets Under the Microscope*. Cambridge University Press, Cambridge, 2018. ISBN 978-1-

107-15605-0. doi: 10.1017/9781316659335. URL https://www.cambridge.org/core/books/trades-quotes-and-prices/029A71078EE4C41C0D5D4574211AB1B5.

Jacob Boudoukh, Ronen Feldman, Shimon Kogan, and Matthew Richardson. Which news moves stock prices? a textual analysis. NBER Working Paper 18725, National Bureau of Economic Research, 2013.

Antonio Briola, Silvia Bartolucci, and Tomaso Aste. Deep Limit Order Book Forecasting. (arXiv:2403.09267), June 2024. doi: 10.48550/arXiv.2403.09267. URL http://arxiv.org/abs/2403.09267.

Jonathan Brogaard, Terrence Hendershott, and Ryan Riordan. High-Frequency Trading and Price Discovery. *The Review of Financial Studies*, 27(8):2267–2306, August 2014. ISSN 0893-9454. doi: 10.1093/rfs/hhu032. URL https://doi.org/10.1093/rfs/hhu032.

Markus K. Brunnermeier and Lasse Heje Pedersen. Predatory Trading. *The Journal of Finance*, 60(4): 1825–1863, 2005. ISSN 1540-6261. doi: 10.1111/j.1540-6261.2005.00781.x. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1540-6261.2005.00781.x.

David Byrd. Learning Not to Spoof. In *Proceedings of the Third ACM International Conference on AI in Finance*, pages 139–147, November 2022. doi: 10.1145/3533271.3561767. URL http://arxiv.org/abs/2306.06087.

Álvaro Cartea, Sebastian Jaimungal, and José Penalva. *Algorithmic and High-Frequency Trading*. Cambridge University Press, August 2015. ISBN 978-1-107-09114-6.

Álvaro Cartea, Patrick Chang, José Penalva, and Harrison Waldon. Algorithmic Collusion and a Folk Theorem from Learning with Bounded Rationality. (4293831), December 2022. doi: 10.2139/ssrn.4293831. URL https://papers.ssrn.com/abstract=4293831.

Álvaro Cartea, Patrick Chang, and Gabriel García-Arenas. Spoofing and Manipulating Order Books with Learning Algorithms. (4639959), November 2023. doi: 10.2139/ssrn.4639959. URL https://papers.ssrn.com/abstract=4639959.

Álvaro Cartea, Patrick Chang, and Rob Graumans. Anonymity, Signaling, and Collusion in Limit Order Books. (5080700), January 2025a. doi: 10.2139/ssrn.5080700. URL https://papers.ssrn.com/abstract=5080700.

Álvaro Cartea, Samuel N Cohen, Robert Graumans, Saad Labyad, Leandro Sánchez-Betancourt, and Leon van Veldhuijzen. Statistical Predictions of Trading Strategies in Electronic Markets. *Journal of Financial Econometrics*, 23(2):nbae025, January 2025b. ISSN 1479-8409. doi: 10.1093/jjfinec/nbae025. URL https://doi.org/10.1093/jjfinec/nbae025.

Alain P. Chaboud, Benjamin Chiquoine, Erik Hjalmarsson, and Clara Vega. Rise of the Machines: Algorithmic Trading in the Foreign Exchange Market. *The Journal of Finance*, 69(5):2045–2084, 2014. ISSN 1540-6261. doi: 10.1111/jofi.12186. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/jofi.12186.

CME Group. Rule 575: Disruptive practices prohibited, 2021. Market surveillance rulebook guidance.

Jean-Edouard Colliard and Thierry Foucault. Trading fees and efficiency in limit order markets. *The Review of Financial Studies*, 25(11):3389–3421, 2012. doi: 10.1093/rfs/hhs089.

Lin William Cong, Ye Li, and Neng Wang. Tokenomics: Dynamic adoption and valuation. *The Review of Financial Studies*, 34(3):1105–1155, 2021. doi: 10.1093/rfs/hhaa089.

Rama Cont, Arseniy Kukanov, and Sasha Stoikov. The Price Impact of Order Book Events. *Journal of Financial Econometrics*, 12(1):47–88, January 2014. ISSN 1479-8409. doi: 10.1093/jjfinec/nbt003. URL https://doi.org/10.1093/jjfinec/nbt003.

Rama Cont, Mihai Cucuringu, and Chao Zhang. Cross-impact of order flow imbalance in equity markets.

Quantitative Finance, 23(10):1373–1393, October 2023. ISSN 1469-7688. doi: 10.1080/14697688.2023.2236159. URL https://doi.org/10.1080/14697688.2023.2236159.

Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. BERT: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics*, pages 4171–4186, 2019.

David Easley, Marcos Lopez de Prado, and Maureen O'Hara. Discerning information from trade data. *Journal of Financial Economics*, 120(2):269–285, May 2016. ISSN 0304-405X. doi: 10.1016/j.jfineco.2016.01.018. URL https://www.sciencedirect.com/science/article/pii/S0304405X16000246.

David Easley, Marcos López de Prado, Maureen O'Hara, and Zhibai Zhang. Microstructure in the Machine Age. *The Review of Financial Studies*, 34(7):3316–3363, July 2021. ISSN 0893-9454. doi: 10.1093/rfs/hhaa078. URL https://doi.org/10.1093/rfs/hhaa078.

David Easley, Maureen O'Hara, Songshan Yang, and Zhibai Zhang. Microstructure and Market Dynamics in Crypto Markets. (4814346), April 2024. doi: 10.2139/ssrn.4814346. URL https://papers.ssrn.com/abstract=4814346.

Gregory W. Eaton, T. Clifton Green, Brian S. Roseman, and Yanbin Wu. Retail trader sophistication and stock market quality: Evidence from brokerage outages. *Journal of Financial Economics*, 146(2):502–528, November 2022. ISSN 0304-405X. doi: 10.1016/j.jfineco.2022.08.002. URL https://www.sciencedirect.com/science/article/pii/S0304405X22001726.

Zoltán Eisler, Jean-Philippe Bouchaud, and Julien Kockelkoren. The price impact of order book events: Market orders, limit orders and cancellations. *Quantitative Finance*, 12(9):1395–1419, September 2012. ISSN 1469-7688. doi: 10.1080/14697688.2010.528444. URL https://doi.org/10.1080/14697688.2010.528444.

Joseph E. Engelberg and Christopher A. Parsons. The causal impact of media in financial markets. *Journal of Finance*, 66(1):67–97, 2011.

Damir Filipović, Markus Pelger, and Ye Ye. Stripping the discount curve: A robust machine learning approach. Research Paper 22-24, Swiss Finance Institute, 2022. URL https://ssrn.com/abstract=4058150.

Damir Filipović, Markus Pelger, and Ye Ye. Shrinking the term structure. Working Paper 32472, National Bureau of Economic Research, 2024. URL https://www.nber.org/papers/w32472.

Thierry Foucault, Johan Hombert, and Ioanid Rosu. News Trading and Speed. (2188822), January 2016. doi: 10.2139/ssrn.2188822. URL https://papers.ssrn.com/abstract=2188822.

Neil Gandal, J. T. Hamrick, Tyler Moore, and Tali Oberman. Price manipulation in the bitcoin ecosystem. *Journal of Monetary Economics*, 95:86–96, 2018. doi: 10.1016/j.jmoneco.2017.12.004.

Diego García. Sentiment during recessions. *Journal of Finance*, 68(3):1267–1300, 2013.

Corey Garriott and Ryan Riordan. The price impacts of informed investors. (4866008), June 2024. doi: 10.2139/ssrn.4866008. URL https://papers.ssrn.com/abstract=4866008.

Jim Gatheral. No-dynamic-arbitrage and market impact. *Quantitative Finance*, 10(7):749–759, 2010.

Lawrence R. Glosten. Is the electronic open limit order book inevitable? *The Journal of Finance*, 49(4):1127–1161, 1994. doi: 10.1111/j.1540-6261.1994.tb02450.x.

Lawrence R. Glosten and Paul R. Milgrom. Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. *Journal of Financial Economics*, 14(1):71–100, 1985. doi: 10.1016/0304-405X(85)90044-3.

Ruslan Goyenko, Bryan T. Kelly, Tobias J. Moskowitz, Yinan Su, and Chao Zhang. Trading Volume Alpha. (4802345), May 2024. doi: 10.2139/ssrn.4802345. URL https://papers.ssrn.com/abstract=4802345.

John M. Griffin and Amin Shams. Is bitcoin really untethered? *The Journal of Finance*, 75(4):1913–1964, 2020. doi: 10.1111/jofi.12903.

Anri Gu, Yongzhao Wang, Chris Mascioli, Mithun Chakraborty, Rahul Savani, Theodore L. Turocy, and Michael P. Wellman. The Effect of Liquidity on the Spoofability of Financial Markets. In *Proceedings of the 5th ACM International Conference on AI in Finance*, pages 239–247, Brooklyn NY USA, November 2024. ACM. ISBN 979-8-4007-1081-0. doi: 10.1145/3677052.3698634. URL https://dl.acm.org/doi/10.1145/3677052.3698634.

James D. Hamilton. A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica*, 57(2):357–384, 1989.

Joel Hasbrouck. Measuring the information content of stock trades. *The Journal of Finance*, 46(1):179–207, 1991. doi: 10.1111/j.1540-6261.1991.tb03749.x.

Joel Hasbrouck and Gideon Saar. Low-latency trading. *Journal of Financial Markets*, 16(4):646–679, 2013. doi: 10.1016/j.finmar.2013.05.003.

Terrence Hendershott, Charles M. Jones, and Albert J. Menkveld. Does Algorithmic Trading Improve Liquidity? *The Journal of Finance*, 66(1):1–33, 2011. ISSN 1540-6261. doi: 10.1111/j.1540-6261.2010.01624.x. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1540-6261.2010.01624.x.

Nicholas Hirschey. Do High-Frequency Traders Anticipate Buying and Selling Pressure? *Management Science*, 67(6):3321–3345, June 2021. ISSN 0025-1909. doi: 10.1287/mnsc.2020.3608. URL https://pubsonline.informs.org/doi/10.1287/mnsc.2020.3608.

IOSCO. Technological challenges to effective market surveillance, 2013. Final report.

Jongho Kang, Jangkoo Kang, and Kyung Yoon Kwon. Market versus limit orders of speculative high-frequency traders and price discovery. *Research in International Business and Finance*, 63:101794, December 2022. ISSN 0275-5319. doi: 10.1016/j.ribaf.2022.101794. URL https://www.sciencedirect.com/science/article/pii/S0275531922001805.

Ron Kaniel, Gideon Saar, and Sheridan Titman. Individual investor trading and stock returns. *The Journal of Finance*, 63(1):273–310, 2008. doi: 10.1111/j.1540-6261.2008.01316.x.

Andrei A. Kirilenko and Gui Lamacie. Latency and Asset Prices. (2546567), January 2015. doi: 10.2139/ssrn.2546567. URL https://papers.ssrn.com/abstract=2546567.

Shimon Kogan, Igor Makarov, Marina Niessner, and Antoinette Schoar. Are cryptos different? Evidence from retail trading. *Journal of Financial Economics*, 159:103897, September 2024. ISSN 0304-405X. doi: 10.1016/j.jfineco.2024.103897. URL https://www.sciencedirect.com/science/article/pii/S0304405X2400120X.

Petter N. Kolm, Jeremy Turiel, and Nicholas Westray. Deep Order Flow Imbalance: Extracting Alpha at Multiple Horizons from the Limit Order Book. (3900141), August 2021. doi: 10.2139/ssrn.3900141. URL https://papers.ssrn.com/abstract=3900141.

Hans R. Künsch. The jackknife and the bootstrap for general stationary observations. *The Annals of Statistics*, 17(3):1217–1241, 1989.

Albert S. Kyle. Continuous auctions and insider trading. *Econometrica*, 53(6):1315–1335, 1985. URL https://www.jstor.org/stable/1913210.

Tao Li, Donghwa Shin, and Baolian Wang. Cryptocurrency Pump-and-Dump Schemes. (3267041), February 2021. doi: 10.2139/ssrn.3267041. URL https://papers.ssrn.com/abstract=3267041.

Gary Lorden. Procedures for reacting to a change in distribution. *Annals of Mathematical Statistics*, 42(6):1897–1908, 1971.

Tim Loughran and Bill McDonald. When is a liability not a liability? textual analysis, dictionaries, and 10-Ks. *Journal of Finance*, 66(1):35–65, 2011.

A. Craig MacKinlay. Event studies in economics and finance. *Journal of Economic Literature*, 35(1):13–39, 1997.

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

Katya Malinova and Andreas Park. Subsidizing liquidity: The impact of make/take fees on market quality. *The Journal of Finance*, 70(2):509–536, 2015. doi: 10.1111/jofi.12230.

Albert J. Menkveld. High Frequency Traders and Market Structure. (1986892), January 2014. doi: 10.2139/ssrn.1986892. URL `https://papers.ssrn.com/abstract=1986892`.

Adamantios Ntakaris, Martin Magris, Juho Kanniainen, Moncef Gabbouj, and Alexandros Iosifidis. Benchmark dataset for mid-price forecasting of limit order book data with machine learning methods. *Journal of Forecasting*, 37(8):852–866, 2018. ISSN 1099-131X. doi: 10.1002/for.2543. URL `https://onlinelibrary.wiley.com/doi/abs/10.1002/for.2543`.

Maureen O'Hara. High frequency market microstructure. *Journal of Financial Economics*, 116(2):257–270, May 2015. ISSN 0304-405X. doi: 10.1016/j.jfineco.2015.01.003. URL `https://www.sciencedirect.com/science/article/pii/S0304405X15000045`.

Y. Ohashi. Momentum ignition price manipulation and pricing mechanisms, 2018. SSRN working paper.

E. S. Page. Continuous inspection schemes. *Biometrika*, 41(1/2):100–115, 1954.

André F. Perold. The implementation shortfall: Paper versus reality. *Journal of Portfolio Management*, 14 (3):4–9, 1988.

Gabriel Peyré and Marco Cuturi. *Computational Optimal Transport*, volume 11 of *Foundations and Trends in Machine Learning*. Now Publishers, 2019.

Mehmet Sağlam, Ciamac C. Moallemi, and Michael G. Sotiropoulos. Short-term trading skill: An analysis of investor heterogeneity and execution quality. *Journal of Financial Markets*, 42:1–28, January 2019. ISSN 1386-4181. doi: 10.1016/j.finmar.2018.12.002. URL `https://www.sciencedirect.com/science/article/pii/S138641811630338X`.

Xuan Tao, Andrew Day, Lan Ling, and Samuel Drapeau. On detecting spoofing strategies in high-frequency trading. *Quantitative Finance*, 22(8):1405–1425, August 2022. ISSN 1469-7688. doi: 10.1080/14697688.2022.2059390. URL `https://doi.org/10.1080/14697688.2022.2059390`.

Paul C. Tetlock. Giving content to investor sentiment: The role of media in the stock market. *The Journal of Finance*, 62(3):1139–1168, 2007. doi: 10.1111/j.1540-6261.2007.01232.x.

Paul C. Tetlock, Maytal Saar-Tsechansky, and Sofus Macskassy. More than words: Quantifying language to measure firms' fundamentals. *The Journal of Finance*, 63(3):1437–1467, 2008. doi: 10.1111/j.1540-6261.2008.01362.x.

Trading Technologies. Tt trade surveillance: Guide to trade surveillance models, 2025. Surveillance model documentation.

Pedro Tremacoldi-Rossi. Manipulating Algorithmic Markets. (5025393), November 2024. doi: 10.2139/ssrn.5025393. URL `https://papers.ssrn.com/abstract=5025393`.

Cédric Villani. *Optimal Transport: Old and New*, volume 338 of *Grundlehren der mathematischen Wissenschaften*. Springer, Berlin, 2009.

Brian M. Weller. Does Algorithmic Trading Reduce Information Acquisition? *The Review of Financial*

Studies, 31(6):2184–2226, June 2018. ISSN 0893-9454. doi: 10.1093/rfs/hhx137. URL https://doi.org/10.1093/rfs/hhx137.

Isaac D. Wright, Matthew Reimherr, and John Liechty. A machine learning approach to classification for traders in financial markets. *Stat*, 11(1):e465, 2022. ISSN 2049-1573. doi: 10.1002/sta4.465. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/sta4.465.

Zihao Zhang, Stefan Zohren, and Stephen Roberts. DeepLOB: Deep Convolutional Neural Networks for Limit Order Books. *IEEE Transactions on Signal Processing*, 67(11):3001–3012, June 2019. ISSN 1053-587X, 1941-0476. doi: 10.1109/TSP.2019.2907260. URL http://arxiv.org/abs/1808.03668.

Internet Appendix

for "Human vs. Machine: Cryptocurrency Microstructure and Market Participants"

The Internet Appendix collects derivations, proofs, diagnostics, and additional results. Sections, tables, figures, and equations carry the prefix IA; cross-references without the prefix point to the main text.

Contents of the Internet Appendix

IA.A Additional Stylized Facts **74**

IA.A.1 Estimator Notation for Profile and Event-Time Figures 74

IA.A.2 Canonical Facts Summary 74

IA.B Additional Data and Cohort-Validation Diagnostics **78**

IA.B.1 Secondary Route Diagnostics: Definitions and Estimates 80

IA.C Market Manipulation: Derivations and Additional Results **82**

IA.C.1 Modeling Choices and Literature Anchors 82

IA.C.2 Spoofing detector: benchmark construction, perturbations, and false positives 83

IA.C.3 Momentum-ignition detector: likelihood ratio, thresholds, and inference 87

IA.C.4 Event-study estimand and consistency 89

IA.C.5 Parameter sensitivity of the manipulation detectors 90

IA.C.6 Additional Momentum-Ignition Results 92

IA.D FPY Depth-Curve Construction and Diagnostics **93**

IA.D.1 Grid, Transforms, Kernel-Ridge Estimator, and FPY Norm 93

IA.D.2 Exact recovery of bid and ask depth from liquidity and imbalance 94

IA.D.3 Additional FPY reconstruction diagnostics 97

IA.D.4 Depth-Factor Interpretation 101

IA.E News Measurement, Evaluation, and Additional Results **102**

IA.E.1 Measurement: alignment, decay features, and smooth-ridge estimator 102

IA.E.2 Evaluation: purged walk-forward design and event-study consistency 103

IA.E.3 Additional News Panels, Examples, and Feature Comparisons 103

IA.F Cross-Exchange Shocks: Additional Results **107**

IA.G Investor Cost Accounting: Derivations and Additional Results **110**

IA.G.1 Cost identities, basis-point aggregation, and impact diagnostic 110

IA.G.2 Matched-Control DiD Identification 111

IA.G.3 Volatility channel: expected shortfall, capacity, and liquidation risk 112

IA.G.4 Additional Cost Tables and Figures 114

IA.G.5 Market-Phase Classification and Regime Robustness 119

IA.H Fee-Schedule Change Impact **121**
IA.H.1 Setting, data, and empirical design 121
IA.H.2 Average post-period differences 124
IA.H.3 Dynamic specification and identification limits 124
IA.H.4 Implications for the main analysis 125

IA.A. Additional Stylized Facts

The main text keeps only the stylized facts that carry the human-machine argument into Sections 3 and 4. This appendix records the remaining diagnostic plots. They suggest that the cryptocurrency market tracks the standard anatomy of a high-frequency limit order book.

IA.A.1. Estimator Notation for Profile and Event-Time Figures

The stylized-fact figures in Section 2.4 and in this appendix use the following estimators. All clock-time figures are computed on Coinbase BTC-USD using Coordinated Universal Time (UTC). For a clock bin b of length Δ and day d, let $X_{d,b}$ denote the statistic in that bin. A profile line is

$$\widehat{\mu}_b = \frac{1}{|\mathcal{D}_b|} \sum_{d \in \mathcal{D}_b} X_{d,b}, \qquad \widehat{q}_b(\alpha) = \inf \left\{ x : \frac{1}{|\mathcal{D}_b|} \sum_{d \in \mathcal{D}_b} \mathbf{1}\{X_{d,b} \le x\} \ge \alpha \right\}, \tag{IA.1}$$

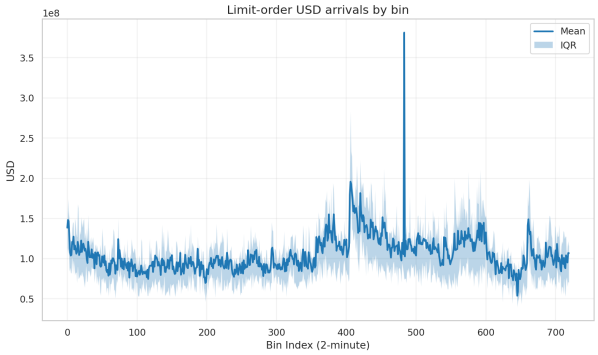
with interquartile bands $[\widehat{q}_b(0.25), \widehat{q}_b(0.75)]$. For an event-time statistic centered on event set $\mathcal{E}$, the plotted path is

$$\widehat{\mu}_Y(k \mid \mathcal{E}) = \frac{1}{|\mathcal{E}|} \sum_{e \in \mathcal{E}} Y_{e,k}, \tag{IA.2}$$

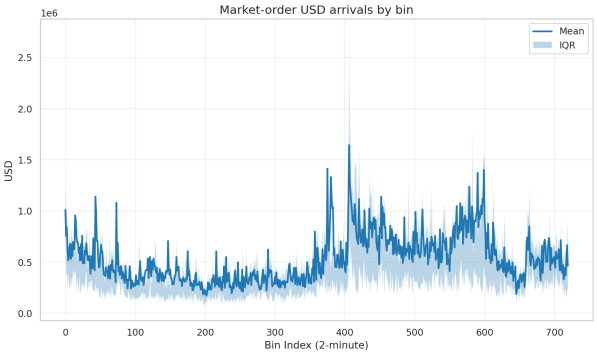
where $Y_{e,k}$ is the transformed outcome k book updates or k seconds after event e, depending on the figure.

IA.A.2. Canonical Facts Summary

The comparison between the canonical limit-order-book facts and the BTC-USD evidence can be stated in words. The canonical facts are those emphasized in the modern limit-order-book literature, especially Bouchaud et al. (2018), and the BTC-USD evidence is based on the event-level book and the figures in this appendix, with displayed depth meaning executable displayed depth in the reconstructed book rather than inferred hidden liquidity. First, visible depth is small relative to immediate demand. Top-of-book capacity is a small fraction of daily turnover, and depth within one percent of the midpoint is right-skewed and state dependent, so the relevant liquidity object is not static displayed volume but executable depth after cancellations, queue updates, and replenishment. Second, buy and sell statistics are approximately symmetric over long horizons. Bid and ask depth distributions are similar and aggressive-flow ratios are broadly comparable across sides, although aggregate symmetry does not imply symmetry across the human and machine cohorts. Third, activity and spreads have daily profiles. Activity, arrivals, and jump probabilities follow strong UTC profiles despite the absence of an exchange open or close, so the crypto analogue of the intraday profile is a global handoff and attention cycle rather than a local trading-day U-shape. Fourth, most activity occurs close to or inside the spread. New limit orders concentrate at the best quotes and nearby ticks, and large best-quote additions arrive at sub-second horizons, reflecting queue priority and immediate execution risk. Fifth, order sizes, queue sizes, and slippage

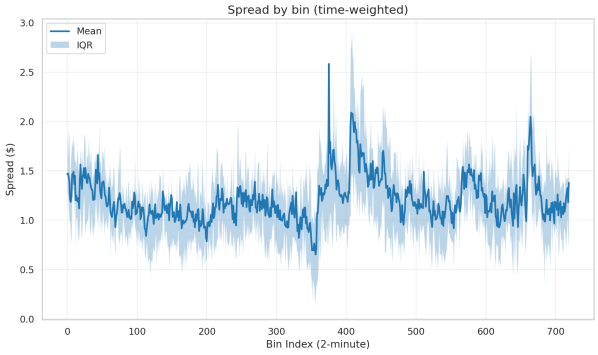


(a) Limit-order arrivals

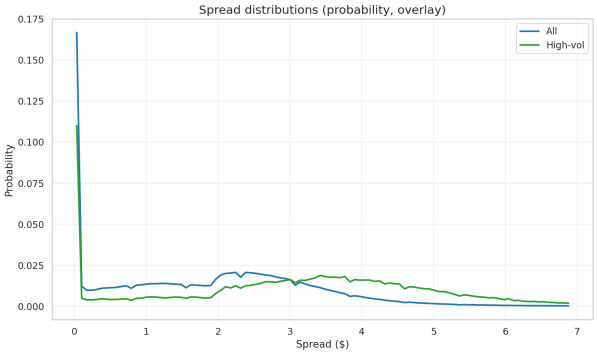


(b) Marketable-order arrivals

Figure IA.1. Intraday arrival profiles. The solid line is the cross-day mean and the shaded region is the interquartile range across UTC days. Passive message traffic is large throughout the day, while aggressive flow is more bursty and concentrated in global trading windows.



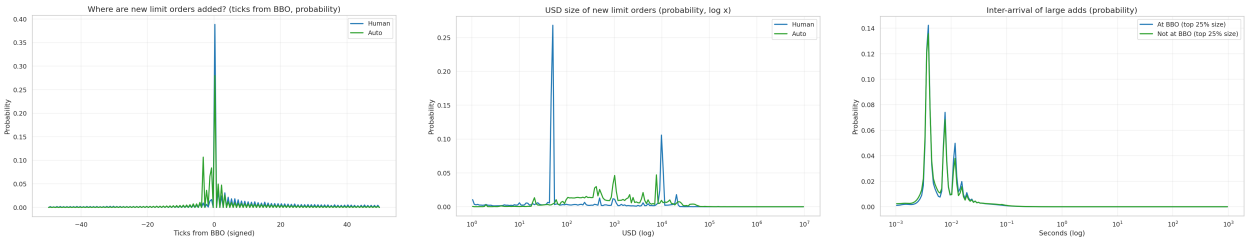
(a) Two-minute spread profile



(b) Full sample versus high-volatility states

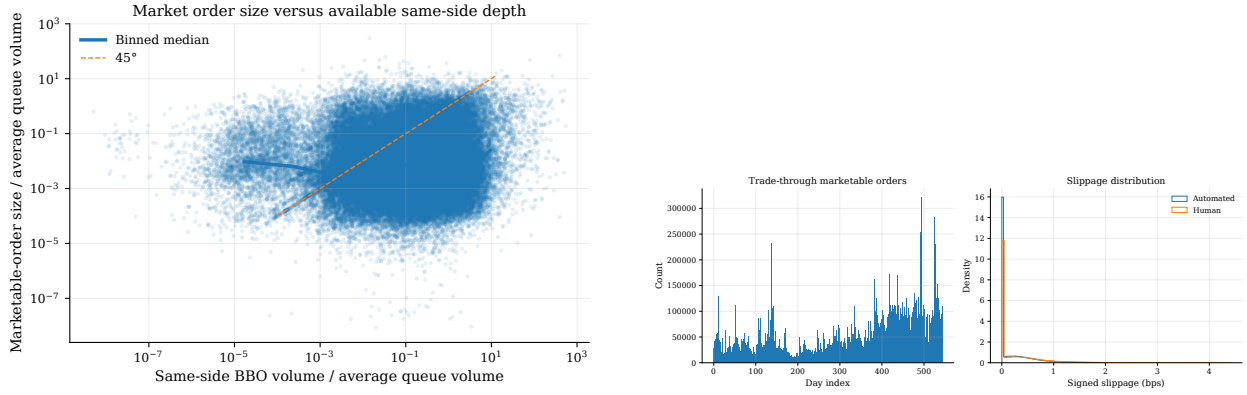
Figure IA.2. Spread profiles and regimes. The spread is state dependent rather than constant background noise. High-volatility states shift probability mass toward wider spreads, consistent with liquidity suppliers protecting quotes when adverse-selection risk rises.

have heavy tails, so average liquidity is a poor sufficient statistic and tail states drive execution costs and strategic opportunities. Sixth, marketable-order signs have long memory. Event-time sign autocorrelation decays slowly and remains positive over long lags, so flow persistence must be modeled explicitly and aggregation can mistake order splitting for information.



(a) Distance from BBO **(b)** Order-size distribution **(c)** Inter-arrival times

Figure IA.3. Passive placement and order-size heterogeneity. Most passive activity occurs near the best quotes, but human and automated order sizes differ. Human order sizes show more round-number mass, while automated sizes are more diffuse across small and intermediate notionals.



(a) Size versus same-side depth **(b)** Trade-throughs and slippage

Figure IA.4. Marketable order size, trade-throughs, and slippage. Most aggressive orders are small relative to same-side depth, but tail orders walk the book and generate slippage. This is why the cost analysis focuses on right-tail incidence rather than only average spreads.

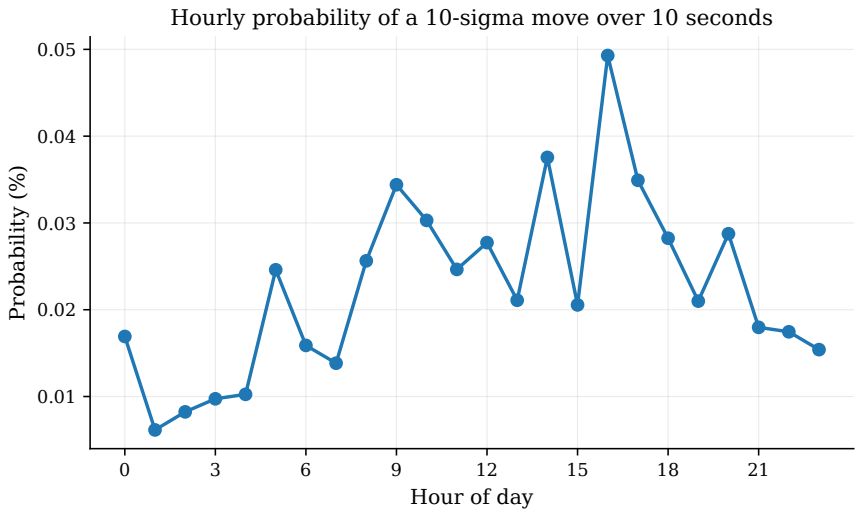


Figure IA.5. Hourly probability of a ten-sigma move over ten seconds. Extreme short-horizon price moves are rare but unevenly distributed across the UTC day, consistent with time-varying attention, liquidity, and information arrival.

IA.B. Additional Data and Cohort-Validation Diagnostics

Section 2.2 makes the cohort argument with order-level cancellation horizons, block-resampled uncertainty, simultaneous quantile bands, stability checks, and explicit contamination exercises. This appendix reports the underlying descriptive distributions and three secondary diagnostics. These plots make the scale and shape of the cohort differences visible, but they are not substitutes for the formal comparisons in the label-validation analysis of the main text. In particular, the panels have very different observation counts, and several objects are aggregate message-stream statistics rather than trader-level behavioral parameters.

Cancellation horizons. For a unique canceled order i, define its observed resting time as

$$T_i^{\text{cancel}} = T_i^{\text{done}} - T_i^{\text{open}}. \tag{IA.3}$$

Figure IA.6 plots the pooled distribution of T_i^{cancel} separately by route on a logarithmic horizontal axis. API-routed cancellations are concentrated at subsecond horizons, whereas UI-associated cancellations place substantially more mass at tens and hundreds of seconds. The separation is visible across the body and the right tail of the distribution, which is why the main text emphasizes quantile ratios rather than a single threshold such as one second. The vertical axes are route specific because the API sample is much larger. Bar heights should therefore not be compared across panels; the informative object is the location and shape of each distribution along the common time axis.

Cancellation-message intensity relative to executions. For route r on UTC date d, let $C_{r,d}$ be the number of canceled `done` messages and let $M_{r,d}$ be the number of `match` messages. The plotted statistic is

$$\text{CMR}_{r,d} = \frac{C_{r,d}}{M_{r,d}}. \tag{IA.4}$$

Figure IA.7 shows that API activity generates many more cancellation messages per match message on a typical day. This pattern is consistent with intensive automated quote maintenance and with the shorter cancellation clock documented above. The object is nevertheless deliberately excluded from the paper's identifying evidence. A match is an execution message rather than a unique completed order, one incoming order can generate several matches, and terminal messages for completed fills are suppressed in the processed feed. Accordingly, $\text{CMR}_{r,d}$ is a message-intensity ratio. It is neither a cancellation probability nor a direct estimate of unfilled-order propensity.

The aggregate event clock. Let $T_{r,j}$ be the timestamp of the jth event carrying route label r, and define the same-route inter-arrival time

$$\Delta T_{r,j} = T_{r,j} - T_{r,j-1}. \tag{IA.5}$$

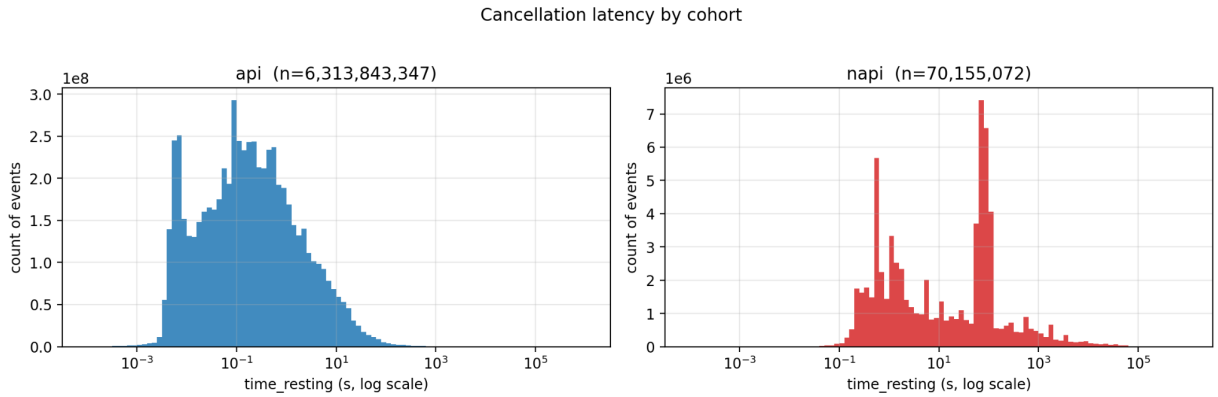


Figure IA.6. Cancellation latency by order-entry route. The panels report elapsed time from an order's `open` message to its canceled `done` message for unique orders. The horizontal axis is logarithmic, and each panel has its own vertical scale. The figure is a descriptive counterpart to the block-inference and quantile-ratio results in Section 2.2.

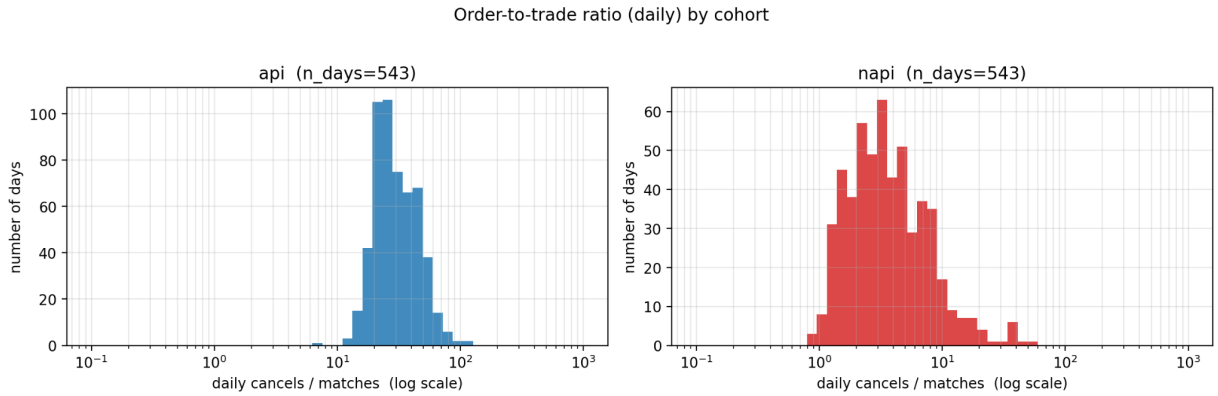


Figure IA.7. Daily cancellation-message to match-message intensity by order-entry route. The plotted quantity is $C_{r,d}/M_{r,d}$ on a logarithmic horizontal axis. It is reported as a descriptive indicator of quote-update intensity and is not interpreted as an order-level cancellation rate.

Figure IA.8 shows a sharp millisecond-scale mode for API events and a much slower distribution for UI-associated events. This finding complements the excess-dispersion statistic in Equation (IA.7). It establishes that the two route populations occupy different aggregate message clocks even before conditioning on event type. It does not establish the reaction time of an individual account, since the public feed does not provide a persistent trader identifier and consecutive same-route events can be submitted by different participants.

Intraday composition. For route r and UTC hour h, define within-route event and executed-notional shares by

$$\pi^N_{r,h} = \frac{N_{r,h}}{\sum_{j=0}^{23} N_{r,j}}, \qquad \pi^V_{r,h} = \frac{V_{r,h}}{\sum_{j=0}^{23} V_{r,j}}, \tag{IA.6}$$

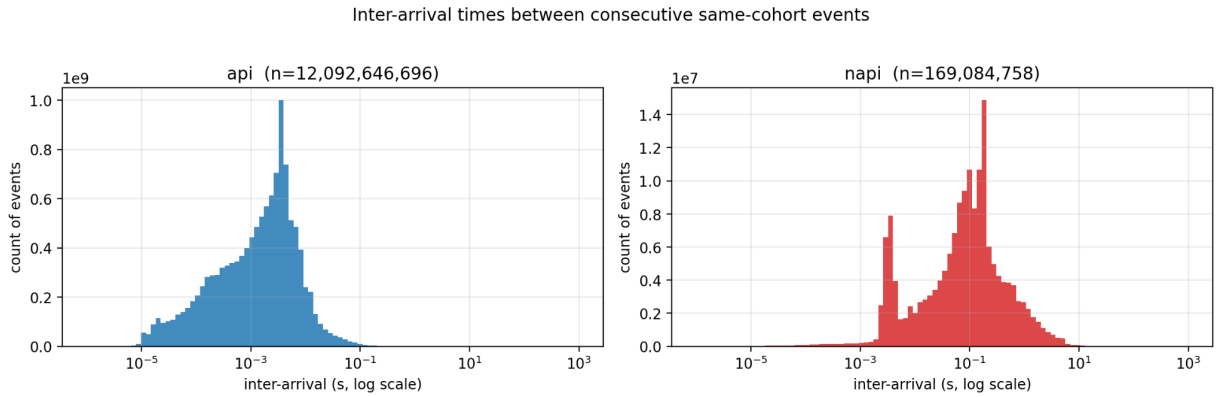


Figure IA.8. Inter-arrival times between consecutive same-route events. The horizontal axis is logarithmic and the vertical scales differ across panels. The distributions measure the clock of the aggregate route-specific event stream, not account-level decision latency.

where $N_{r,h}$ is event count and $V_{r,h}$ is matched dollar notional. Normalizing within route removes the level difference between the two populations and isolates the timing of their activity. Figure IA.9 shows that API event and notional shares are relatively elevated during the 13:00–16:00 UTC interval, while UI-associated activity is flatter and relatively stronger late in the UTC day. The differences are economically coherent with a globally traded asset and distinct participation technologies, but they are modest compared with the latency separation. We therefore use the intraday profile as corroboration and absorb hour-of-day effects in later regressions rather than treating calendar time as a cohort identifier.

Figure IA.10 reports the two label-contamination experiments summarized in the main text.

IA.B.1. Secondary Route Diagnostics: Definitions and Estimates

Three additional diagnostics provide corroborating evidence, although none is used as an identifying assumption. First, we compare the aggregate arrival processes of `received` orders after normalizing for the large difference in mean route intensity. For one-second counts N_{rs}, define excess dispersion as

$$X_r = \frac{\mathrm{Var}(N_{rs}) - \mathrm{E}(N_{rs})}{\mathrm{E}(N_{rs})^2}. \tag{IA.7}$$

This statistic equals zero under a homogeneous Poisson benchmark and is invariant to scale under a mixed-Poisson interpretation. Its daily mean is 0.43 for API-routed flow and 1.74 for UI-associated flow. The difference is 1.31, with a seven-day block interval of [1.06, 1.57]. UI-associated submissions are therefore more clustered relative to their own average arrival rate. Figure IA.8 reports the corresponding event-stream inter-arrival distributions. We do not interpret aggregate inter-arrival times as individual reaction times because the feed does not disclose a persistent client identifier suitable for that purpose.

Second, UI-associated new orders are relatively more concentrated during 01:00–03:00 and 20:00–23:00 UTC, whereas API-routed orders are relatively more concentrated during 08:00–16:00

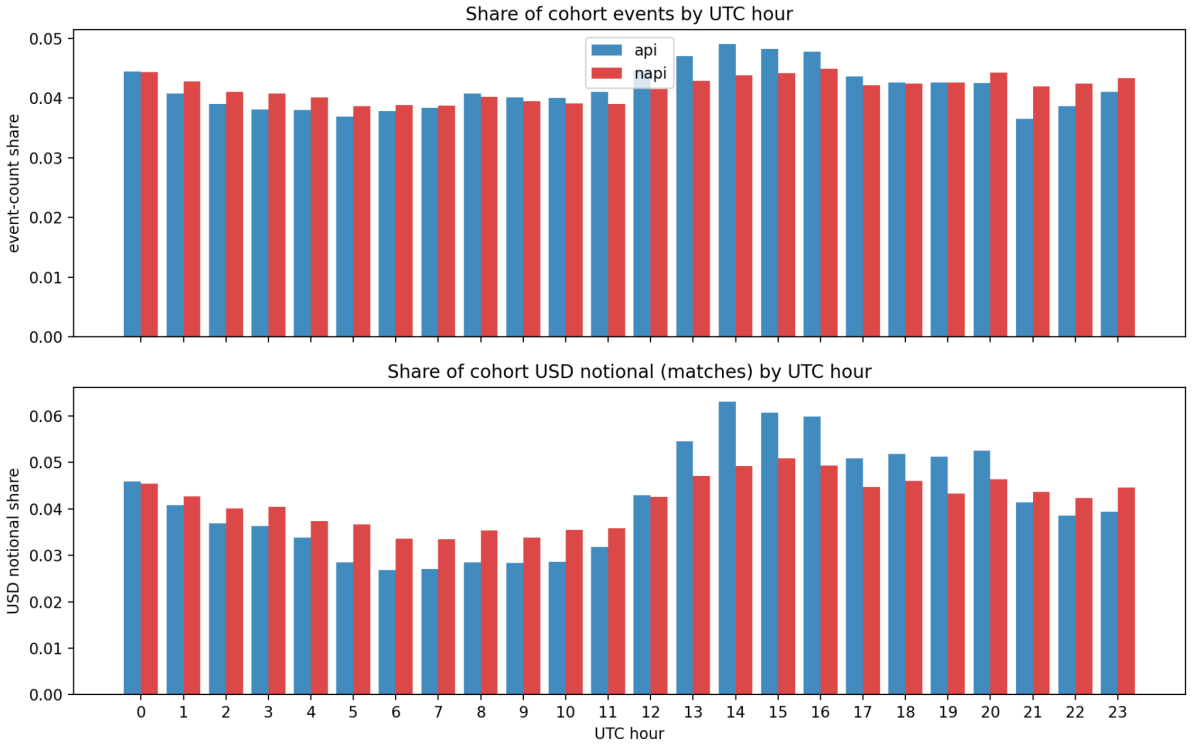


Figure IA.9. Intraday composition by order-entry route. The top panel reports each UTC hour's share of route-specific event counts. The bottom panel reports the analogous share of matched dollar notional. Each series sums to one across the 24 UTC hours, so the figure compares timing rather than activity levels.

UTC. The largest simultaneous-band UI excess is 0.59 percentage points at 21:00 UTC, and the largest API excess is 0.40 percentage points at 14:00 UTC. Figure IA.9 shows both within-cohort event-count shares and within-cohort executed-notional shares by UTC hour. These differences are systematic but modest relative to the cancellation-latency evidence.

Third, we study bunching around round dollar notionals after removing first-order differences in the size distribution. For each multiple m, we estimate roundness rates within notional bins and standardize both routes to the same pooled common-support distribution,

$$\widetilde{p}_r(m) = \sum_b \omega_b \Pr(|V - m \text{ round}(V/m)| \leq 1 \mid V \in b, r) \,, \tag{IA.8}$$

where V is order notional in dollars. The standardized UI/API point estimates are 1.50, 1.79, and 2.55 around \$100, \$500, and \$1,000 multiples. The local density plots also show visible mass at exact round values. However, the seven-day block intervals include one and the monthly estimates are unstable. We therefore treat round-number bunching as suggestive corroboration rather than evidence on which the cohort definition rests.

Taken together, these diagnostics reinforce the interpretation of the observable route field as a technological margin. API-routed activity operates on a faster cancellation and event clock,

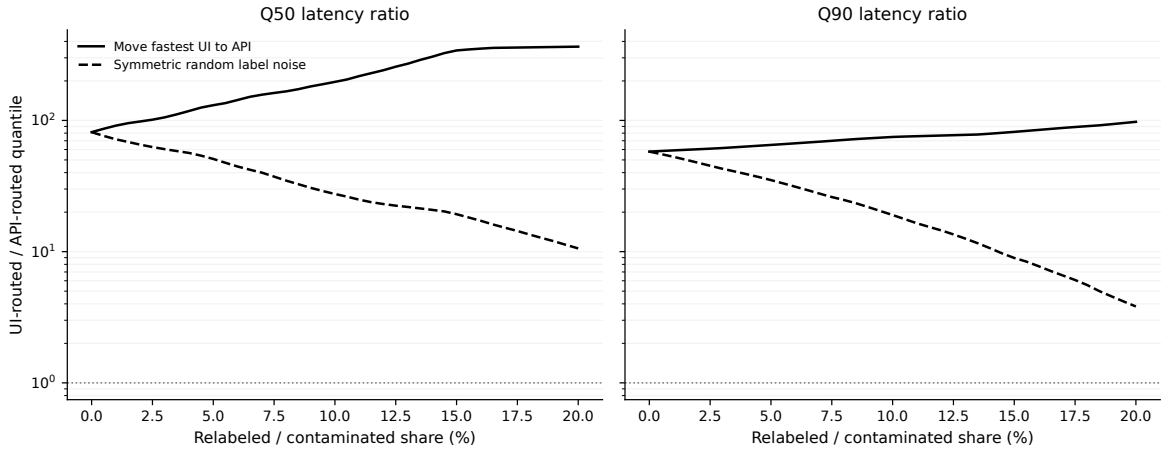


Figure IA.10. Sensitivity of latency separation to route-label contamination. The solid line reassigns the fastest UI-associated cancellations to the API cohort. The dashed line symmetrically mixes the two normalized route distributions. The panels report the UI/API ratio at the median and the 90th percentile.

produces substantially more quote-removal messages per execution message, and has a somewhat different intraday composition. The paper's empirical claims do not rely on any one of these visual differences. They rely on the distribution-wide cancellation-latency separation and its formal robustness checks, with the figures in this appendix providing a transparent view of the underlying data.

IA.C. Market Manipulation: Derivations and Additional Results

This appendix collects the mathematical derivations used in Section 4. The results are stated for the population objects; the corresponding sample estimators replace population distributions and expectations with their empirical analogues.

IA.C.1. Modeling Choices and Literature Anchors

Each modeling choice in Section 4 pairs an economic channel with a statistical tool, and this subsection records the pairings. The FPY depth-curve state is used because manipulation changes the geometry of displayed depth, so the book should be represented as a curve over distance from the touch rather than as unrelated quote levels, and RKHS curve smoothing with smoothness-ordered factors adapts the term-structure logic of Filipović et al. (2022) and Filipović et al. (2024) to liquidity and imbalance curves. Multi-level depth imbalance is used because spoofing works through visible supply away from the touch as well as at it, so the statistic must retain posting distance, in line with limit-order-book theory and spoofing models (Glosten, 1994; Cont et al., 2014; Tao et al., 2022; Cartea et al., 2023; Gu et al., 2024). The Wasserstein benchmark comparison is used because a spoof-like path is a distortion of the whole distribution of imbalance paths rather

than a single threshold crossing, and the Wasserstein distance preserves the geometry of those states with precedent in order-book classification (Tao et al., 2022; Wright et al., 2022; Villani, 2009; Peyré and Cuturi, 2019). The high-CV regime split reflects that artificial depth should matter most when message traffic is bursty and quote revision is active (Hendershott et al., 2011; Hasbrouck and Saar, 2013; Kirilenko and Lamacie, 2015; Easley et al., 2021). The Poisson GLR is used for ignition because the object is a short-window acceleration in one-sided aggressive trading, and likelihood-ratio change detection gives a transparent state-scaled test for abnormal counts (Page, 1954; Lorden, 1971; Basseville and Nikiforov, 1993). Price confirmation and human follow-through implement the price-impact and adverse-selection logic of the microstructure literature and surveillance descriptions of disruptive trading (Kyle, 1985; Glosten and Milgrom, 1985; Gatheral, 2010; Easley et al., 2016; CME Group, 2021; IOSCO, 2013; Trading Technologies, 2025). Event-time plots are used because manipulation is a local state, so the natural clock is distance from the triggering state rather than time of day (MacKinlay, 1997; Bouchaud et al., 2018). The purpose throughout is interpretability rather than maximal in-sample fit.

IA.C.2. Spoofing detector: benchmark construction, perturbations, and false positives

This subsection gives the full spoofing construction used behind Figure 9. The main text uses the simple imbalance statistic because it is easier to interpret. The full detector compares the observed book to two empirical benchmark distributions, one representing legitimate depth and one representing spoofed depth, both conditioned on the contemporaneous market state.

Let Z_t be the state vector observed just before the candidate order-book configuration,

$$Z_t = (\log s_t,\ \log D_t(K),\ |r_{t-60s,t}|,\ \log N_{t-60s,t},\ I_t(K))\,, \tag{IA.9}$$

where s_t is the quoted spread, $D_t(K) = D_t^b(K) + D_t^a(K)$ is multi-level displayed depth, $r_{t-60s,t}$ is the last-minute return, $N_{t-60s,t}$ is message count, and $I_t(K)$ is multi-level imbalance. We discretize each component of Z_t into empirical terciles inside the same calendar month. Let $\mathcal{B}(z)$ be the resulting state bucket.

The legitimate benchmark distribution is constructed from book states that do not contain a candidate transient wall and that survive a minimum lifetime filter. Formally,

$$\mathcal{L}(z) = \{I_u(K) : \mathcal{B}(Z_u) = \mathcal{B}(z),\ W_u = 0,\ \ell_u \geq \ell_{\min}\}\,, \tag{IA.10}$$

where W_u is an indicator that a single order or order cluster accounts for more than ω of same-side depth inside the monitored band, and ℓ_u is the lifetime of the largest displayed order in the band. The lifetime filter removes fleeting wall-like states from the legitimate comparison group.

The spoofed benchmark distribution is constructed by injecting a candidate wall into otherwise comparable pre-event books. If the candidate side is $d \in \{+1, -1\}$, candidate notional is v, and

posting distance is a ticks from the best quote, the perturbed imbalance is

$$I_u^{\mathrm{sp}}(K; d, v, a) = \frac{D_u^b(K) - D_u^a(K) + d\, w(a) v}{D_u^b(K) + D_u^a(K) + w(a) v + \varepsilon}, \qquad w(a) = \exp(-a/\eta), \tag{IA.11}$$

computed over states u with $\mathcal{B}(Z_u) = \mathcal{B}(z)$. The exponential distance weight reflects the fact that a near-touch wall is more informative for immediate execution than an otherwise identical wall deep in the book. The spoof-strength statistic is the one-sided Wasserstein displacement between the spoofed and legitimate conditional laws,

$$S_t = \max\left\{0,\ W_1\left(\widehat{F}_t^{\mathrm{sp}}, \widehat{F}_t^{\mathrm{legit}}\right) - q_{1-\alpha}\left[W_1\left(\widehat{F}_t^{\mathrm{pl}}, \widehat{F}_t^{\mathrm{legit}}\right)\right]\right\}, \tag{IA.12}$$

where $\widehat{F}_t^{\mathrm{pl}}$ is a placebo distribution obtained by injecting the same wall into a randomly time-shifted book from the same state bucket and UTC hour. The subtraction of the placebo quantile makes the detector conservative in naturally unbalanced states.

For each horizon h, cohort c, and target Y, the predictive projection is

$$Y_{c,h}(t) = \alpha_{c,h} + \beta_{c,h} S_t + u_{c,h}(t). \tag{IA.13}$$

For a candidate order or order cluster i, the observed target admits the decomposition

$$Y_{c,h}(t) = R_{i,c,h}(t) + Y_{c,h}^{(-i)}(t), \tag{IA.14}$$

where $R_{i,c,h}(t)$ records the candidate's own cancellation, execution, or replacement contribution and $Y_{c,h}^{(-i)}(t)$ is revision elsewhere in the book. The projection coefficient $\beta_{c,h}$ therefore combines predictability of the candidate lifecycle with the response of other displayed liquidity; a leave-candidate-out target is required to identify the second component alone.

Uncertainty is computed by resampling whole UTC days. In bootstrap draw b, we sample days with replacement, recompute the projection, and store $R_{b,c,h}^2$. Confidence intervals are percentile intervals from the empirical distribution of $R_{b,c,h}^2$. Multiple testing across outcomes, cohorts, and horizons is controlled by the Benjamini-Hochberg false-discovery-rate procedure applied to the bootstrap p-values for $\beta_{c,h} = 0$.

The robustness results support a narrow interpretation. The spoof-strength signal is not a universal predictor of all future order flow. It is most informative for the liquidity margin and for machine response in bursty states. That pattern is precisely what the mechanism predicts. A false displayed-depth state matters when fast traders are actively revising quotes and slower traders are more likely to treat visible depth as an informative signal.

Depth imbalance and spoof perturbations. The purpose of this block is sensitivity calibration. We inject synthetic walls into real books and verify that the imbalance statistic responds as designed, so that the detector's response to a wall is a property of the construction rather than an

Outcome	Tested cells	Positive cells	Survive BH 5%	Placebo false-positive rate
Buy taker flow	14	13	11	4.10%
Sell taker flow	14	12	10	4.80%
Net liquidity	14	14	13	2.20%
All cells	42	39	34	3.70%

Table IA.I Spoofing robustness, multiple testing, and placebo false positives. Positive cells have the predicted sign for the spoof-strength coefficient. The false-positive rate is the share of time-shift placebo cells that would be selected by the same threshold rule. Reported rates are below the nominal 5% design level, while the majority of economically relevant cells survive the Benjamini-Hochberg correction.

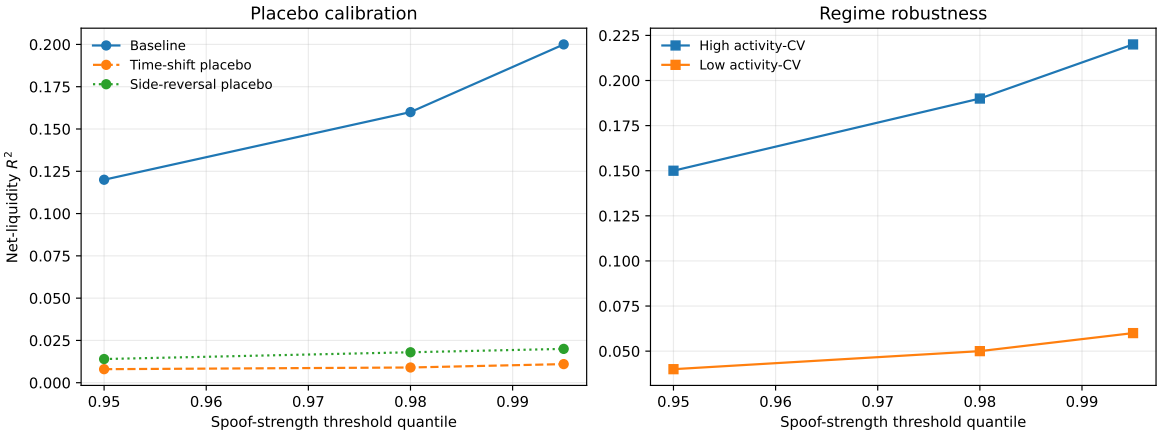


Figure IA.11. Spoofing robustness diagnostics. The left panel compares the baseline detector with time-shift and side-reversal placebos. The right panel shows that predictability is concentrated in high activity-CV states, where displayed depth is most likely to affect quote revision and human attention.

accident of the data.

Boundedness and monotonicity of multi-level imbalance. Fix a depth K and suppose $B_{K,t} \geq 0$, $A_{K,t} \geq 0$, and $\varepsilon > 0$. Define

$$I_K(t) = \frac{B_{K,t} - A_{K,t}}{B_{K,t} + A_{K,t} + \varepsilon}.$$

Then $I_K(t) \in [-1, 1]$. If an additional ask-side order of size $v \geq 0$ is added within the depth K, so that

$$I_K^a(t; v) = \frac{B_{K,t} - A_{K,t} - v}{B_{K,t} + A_{K,t} + v + \varepsilon},$$

then $I_K^a(t;v)$ is weakly decreasing in v. If an additional bid-side order of size $v \geq 0$ is added within the depth K, so that

$$I_K^b(t;v) = \frac{B_{K,t} + v - A_{K,t}}{B_{K,t} + A_{K,t} + v + \varepsilon},$$

then $I_K^b(t;v)$ is weakly increasing in v.

Liquidity-cost formula in a block-shaped book. Consider an ask book with constant depth $a > 0$ per tick and tick size $\delta > 0$. A buy market order of size $H > 0$ walks approximately $F(H) = H/a$ ticks and pays liquidity cost $\delta G(H)$ in excess of the best ask value, with

$$G(H) = \frac{H^2}{2a} + O(H).$$

In the continuous block approximation the leading term is $H^2/(2a)$.

In the continuous block approximation, the marginal unit $x \in [0, H]$ of the market order executes at depth x/a ticks above the best ask. The excess tick cost is therefore

$$G(H) = \int_0^H \frac{x}{a}\, dx = \frac{H^2}{2a}.$$

In a discrete tick book, the execution depth is $\lceil x/a \rceil$, so replacing the integral by the corresponding Riemann sum changes the result by a boundary term of order $O(H)$. Multiplication by the tick size δ gives the dollar excess cost. This is the approximation used to connect displayed depth, order size, and spoofing incentives.

Wasserstein distance and empirical spoof comparison. The Wasserstein distance is used because the full detector compares the entire distribution of the observed book state with the legitimate benchmark distribution, not just a mean, and this distance is the natural notion of how far one distribution must be moved to match another.

One-dimensional quadratic Wasserstein distance. Let F and G be probability distributions on $\mathbb{R}$ with finite second moments, and let F^{-1} and G^{-1} be their generalized quantile functions. Then

$$W_2^2(F, G) = \int_0^1 \left(F^{-1}(u) - G^{-1}(u)\right)^2 du.$$

Moreover, if $\widehat{F}_n$ is the empirical distribution of a strictly stationary ergodic sample with marginal law F and finite second moment, then $W_2(\widehat{F}_n, F) \to 0$ almost surely.

This is the classical one-dimensional optimal-transport formula (Villani, 2009).

The proposition justifies the spoof-strength comparison in equation (IA.12). The conditional implementation applies the same argument within post-trade imbalance buckets. With a fixed number of buckets and a positive limiting probability in each bucket, the empirical conditional laws converge to their population analogues by the same ergodic argument.

IA.C.3. Momentum-ignition detector: likelihood ratio, thresholds, and inference

One-sided Poisson GLR. Let $N = n$ be the number of events in a window of length w_1, and suppose under the null that $N \sim \text{Poisson}(\mu)$ with $\mu = \lambda_0 w_1 > 0$. For the one-sided alternative $\lambda \geq \lambda_0$, the generalized log-likelihood ratio is

$$\Lambda(n;\mu) = \begin{cases} n\log(n/\mu) - (n-\mu), & n \geq \mu, \\ 0, & n < \mu. \end{cases}$$

Equivalently, for $n \geq \mu$, $\Lambda(n;\mu) = \mu D(n/\mu \| 1)$, where $D(x\|1) = x\log x - (x-1)$ is the Poisson large-deviation rate function.

The log likelihood for rate λ is

$$\ell(\lambda; n) = n\log(\lambda w_1) - \lambda w_1 - \log(n!).$$

The unrestricted maximum likelihood estimator is $\widehat{\lambda} = n/w_1$, obtained from

$$\frac{\partial \ell}{\partial \lambda} = \frac{n}{\lambda} - w_1 = 0.$$

Under the constrained alternative $\lambda \geq \lambda_0$, the maximizer is $\lambda^* = \max\{n/w_1, \lambda_0\}$. If $n < \mu = \lambda_0 w_1$, then $n/w_1 < \lambda_0$ and $\lambda^* = \lambda_0$, so the log-likelihood ratio is zero. If $n \geq \mu$, then $\lambda^* = n/w_1$ and

$$\begin{aligned} \Lambda(n;\mu) &= \ell(n/w_1; n) - \ell(\lambda_0; n) \\ &= n\log(n) - n - \log(n!) - \{n\log(\mu) - \mu - \log(n!)\} \\ &= n\log(n/\mu) - (n-\mu). \end{aligned}$$

Writing $x = n/\mu$ gives $\Lambda = \mu[x\log x - (x-1)] = \mu D(x\|1)$.

Nonnegativity and local quadratic approximation. For $x \geq 0$, $D(x\|1) = x\log x - (x-1) \geq 0$, with equality if and only if $x = 1$. Moreover, as $x \to 1$,

$$D(x\|1) = \frac{(x-1)^2}{2} + O((x-1)^3).$$

The function D is convex because $D''(x) = 1/x > 0$ for $x > 0$, and $D'(x) = \log x$, so the unique minimizer is $x = 1$. Since $D(1\|1) = 0$, nonnegativity follows. A Taylor expansion around $x = 1$ gives $D(1+y\|1) = \frac{y^2}{2} + O(y^3)$.

The second proposition shows why the GLR behaves like a squared standardized surprise for small deviations but retains the correct large-deviation scaling for extreme bursts. This is the reason for using a likelihood-ratio statistic rather than a raw count threshold.

State-conditioned thresholds.

Conditional size of the MI trigger. Let T be a scalar score and $B \in \mathcal{B}$ a finite-valued state variable. Under the null hypothesis, suppose the conditional distribution of $|T|$ given $B = b$ is continuous with $(1-\gamma)$ quantile $q_{1-\gamma}^{(b)}$. Define the trigger $\mathcal{A} = \{|T| \geq q_{1-\gamma}^{(B)}\}$. Then

$$\mathbb{P}_0(\mathcal{A} \mid B = b) = \gamma \quad \text{for each } b \in \mathcal{B},$$

and hence $\mathbb{P}_0(\mathcal{A}) = \gamma$.

By definition of the conditional quantile and continuity,

$$\mathbb{P}_0(|T| < q_{1-\gamma}^{(b)} \mid B = b) = 1 - \gamma.$$

Taking complements gives $\mathbb{P}_0(|T| \geq q_{1-\gamma}^{(b)} \mid B = b) = \gamma$. Averaging over the states yields

$$\mathbb{P}_0(\mathcal{A}) = \sum_{b \in \mathcal{B}} \mathbb{P}_0(\mathcal{A} \mid B = b)\mathbb{P}_0(B = b) = \gamma.$$

If the conditional quantiles are estimated from a stationary mixing training sample and each state has positive probability, the empirical quantiles are consistent for the population quantiles. The feasible trigger therefore has asymptotic conditional size γ state by state. This is the formal justification for comparing a high-activity state to its own tail rather than to a pooled unconditional threshold.

Bootstrap uncertainty and multiple-testing convention. For statistics based on calendar-time prediction, the bootstrap resamples whole UTC days. Let $\mathcal{D}$ be the set of trading days and let $\theta = T(\{X_d : d \in \mathcal{D}\})$ be any scalar statistic such as an out-of-sample R^2, an average cost, or a pre/post contrast. Bootstrap draw b samples $|\mathcal{D}|$ days with replacement, recomputes the statistic on the sampled set, and produces $\theta^{*(b)}$. The reported 95 percent interval is

$$\left[Q_{0.025}\{\theta^{*(b)}\}_{b=1}^{B},\ Q_{0.975}\{\theta^{*(b)}\}_{b=1}^{B} \right]. \tag{IA.15}$$

For event studies with clustered events, the event bootstrap samples event clusters rather than individual event-time observations. A cluster is a maximal set of events whose windows overlap. Resampling clusters preserves within-window serial dependence and avoids treating multiple events in the same burst as independent draws.

For a family of m hypothesis tests with bootstrap p-values $p_1, \ldots, p_m$, let $p_{(1)} \leq \cdots \leq p_{(m)}$ be the ordered values. The Benjamini-Hochberg rejection index at false-discovery rate q is

$$\widehat{k} = \max \left\{ k : p_{(k)} \leq \frac{k}{m} q \right\}. \tag{IA.16}$$

The rejected set is $\{j : p_j \leq p_{(\widehat{k})}\}$. We use this correction for families defined by the same economic exercise, such as spoofing outcomes across cohorts and horizons. The correction is intentionally reported rather than hidden because a result that survives only as an uncorrected point estimate is not strong enough to support the paper's main claims.

IA.C.4. Event-study estimand and consistency

For each event e, let $Y_e(\ell)$ denote the event-time outcome in bin ℓ after applying the appropriate transformation $g_Y(\cdot, s_e)$. The event-study target is the conditional path

$$\mu_Y(\ell) = \mathbb{E}\left[Y_e(\ell) \mid e \in \mathcal{E}\right].$$

The event-study estimator of Section 4.3.2 is the sample mean

$$\widehat{\mu}_Y(\ell) = \frac{1}{E}\sum_{e=1}^{E} Y_e(\ell).$$

The interpretation of this path requires two economic assumptions. First, events must be isolated enough that the window is not mechanically mixing different shocks. Second, the conditional path must be homogeneous enough that the average is a meaningful object.

ASSUMPTION IA.1 (Event isolation): *For each event type used in an event study, the event window contains no other event of a different type whose own event-time effect loads on the same outcome at the same horizon, except within clusters that are explicitly collapsed or resampled as one unit.*

Economically, event isolation means that an ignition window is not simultaneously being used as a news window or a spoofing window when the interpretation requires an ignition-only path. It is not literally true in every second of a crypto market. The empirical compromise is to cluster overlapping windows and to treat clustered observations as the sampling unit for uncertainty. The resulting event study is therefore a conditional mean around a detected state, not a structural impulse response to a randomized shock.

ASSUMPTION IA.2 (Homogeneous conditional path): *For a fixed event type and conditioning set, the average potential event-time path is stable across calendar time and across the cross-section of events after conditioning on the variables used to define the event and the state splits.*

Economically, homogeneity means that the plotted curve is a meaningful average of similar events rather than a mixture of unrelated states. This assumption is more plausible for narrow bins of volatility, spread, and intensity than for the fully pooled sample. We therefore report placebo checks, high-activity splits, and interval estimates when the mechanism predicts heterogeneity, and we interpret pooled event studies as descriptive conditional means.

Consistency of the event-time mean. Under stationarity and ergodicity over events the binned mean is consistent, and under standard mixing conditions it is asymptotically normal.

These are textbook results, and the design content lies in the event definition and the day-block inference rather than in the limit theory.

Consistency of pre/post contrasts. Under the conditions of the preceding proposition, the event-minus-pre and post-minus-pre contrast estimators of Section 4.3.2 are consistent for their corresponding population contrasts.

IA.C.5. Parameter sensitivity of the manipulation detectors

The figures in this subsection report model-implied sensitivity rather than new parameter estimates. They are included to make clear which assumptions drive the two manipulation detectors and to show that the baseline parameters have standard microstructure interpretations.

Sensitivity of the imbalance perturbation. Let $I = (B - A)/(B + A)$ be the pre-spoof imbalance and let $x = w_k v/(B + A) \geq 0$ denote the depth-weighted spoof size as a fraction of displayed depth. A bid-side spoof maps I into $I^b(x) = (I + x)/(1 + x)$ and an ask-side spoof maps I into $I^a(x) = (I - x)/(1 + x)$. Hence

$$\frac{\partial I^b(x)}{\partial x} = \frac{1 - I}{(1 + x)^2}, \qquad \frac{\partial I^a(x)}{\partial x} = -\frac{1 + I}{(1 + x)^2}.$$

The absolute marginal effect is decreasing in x and is increasing in the depth weight w_k.

The bid-side formula follows from

$$I^b(x) = \frac{B - A + w_k v}{B + A + w_k v} = \frac{I + x}{1 + x}.$$

Differentiation gives $(1 + x - (I + x))/(1 + x)^2 = (1 - I)/(1 + x)^2$. The ask-side formula is

$$I^a(x) = \frac{B - A - w_k v}{B + A + w_k v} = \frac{I - x}{1 + x},$$

whose derivative is $[-(1 + x) - (I - x)]/(1 + x)^2 = -(1 + I)/(1 + x)^2$. Since $x = w_k v/(B + A)$, increasing the depth weight increases x for fixed raw order size. The derivative magnitudes are decreasing in x because the denominator $(1 + x)^2$ is increasing.

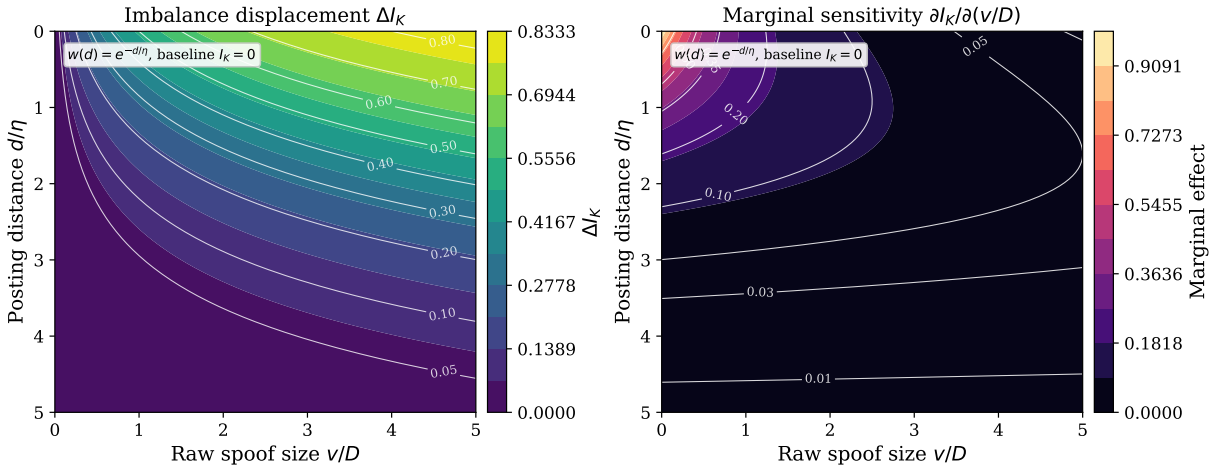


Figure IA.12. Parameter sensitivity of the spoofing imbalance model. The left panel plots the model-implied displacement in multi-level imbalance from a bid-side wall as a function of raw spoof size v/D and posting distance d/η, with depth weight $w(d) = \exp(-d/\eta)$. The right panel plots the marginal derivative. The sensitivity is strongest for large, near-touch walls and exhibits diminishing marginal effects in size, which is why the empirical detector weights both order size and posting distance. The ask-side case is symmetric.

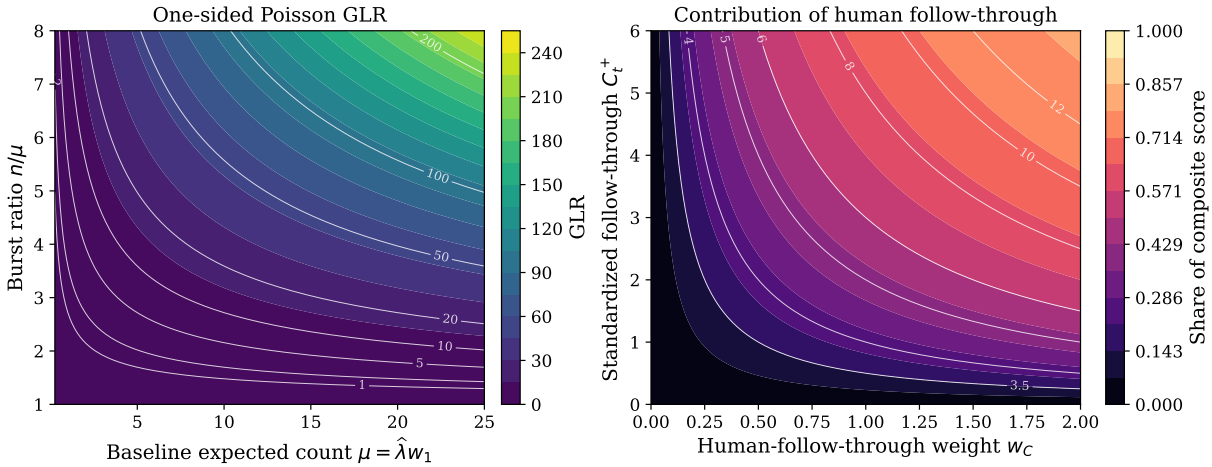


Figure IA.13. Parameter sensitivity of the momentum-ignition detector. The left panel plots the one-sided Poisson GLR as a function of the baseline expected count μ and the realized burst ratio n/μ. The right panel plots the share of the composite score attributable to human follow-through as the weight w_C and standardized follow-through C_t^+ vary, holding the acceleration and price-confirmation components fixed (default $w_C = 0$)

IA.C.6. Additional Momentum-Ignition Results

The predictive-content validation of the signed MI score is shown in the main text (Figure 13 in Section 4.3.3). Table IA.II records the baseline detector specification and its rationale. Table IA.III reports the post-window and event-window take-ratio decomposition underlying the estimates discussed in Section 4.3.3. Figure IA.14 shows the smooth response of displayed liquidity, and Figure IA.15 separates the volatility burst from directional drift.

Table IA.II Momentum-Ignition Detector Baseline Specification

Object	**Baseline value**	**Economic rationale**
Detection window w_1	2 seconds	Captures abrupt aggressive bursts before they diffuse into ordinary volume.
Baseline window w_0	20 seconds	Estimates local intensity while adapting to short-lived regime changes.
Follow-through sensitivity Δ	3 seconds	Used only when $w_C > 0$ to assess how including the post-burst human response would change event selection; the baseline sets $w_C = 0$.
State buckets	Volatility $\times$ intensity $\times$ imbalance	Prevents high-volume or high-volatility states from mechanically generating triggers.
Trigger quantile	$1 - \gamma = 0.995$	Selects extreme state-conditioned acceleration rather than ordinary active trading.
Core weights	$w_A, w_Z, w_C \geq 0$	Separates order-flow acceleration, price confirmation, and human follow-through; baseline $w_C = 0$.

Notes: The detector is estimated on event-time order-flow objects and evaluated in clock-time event studies. The design uses interpretable components rather than unrestricted high-dimensional classifiers.

Table IA.III Take Ratios around Momentum-Ignition Events ($h = 30$s)

	Δ Post–Pre	Δ%	Δ Event–Pre	Δ% Event–Pre
Human	0.00061	10.51	**0.02451**	**424.54**
Machine	0.00007	2.47	0.00134	48.90
Human − Machine	**0.00054**	**8.04**	**0.02318**	**375.64**

Notes: Take ratio equals taker volume divided by taker plus limit-order volume. Entries are trimmed means of event-bin ratios, not ratios of aggregate event-bin sums. The bold entries mark the event-time human liquidity-demand spike and the human-machine differential.

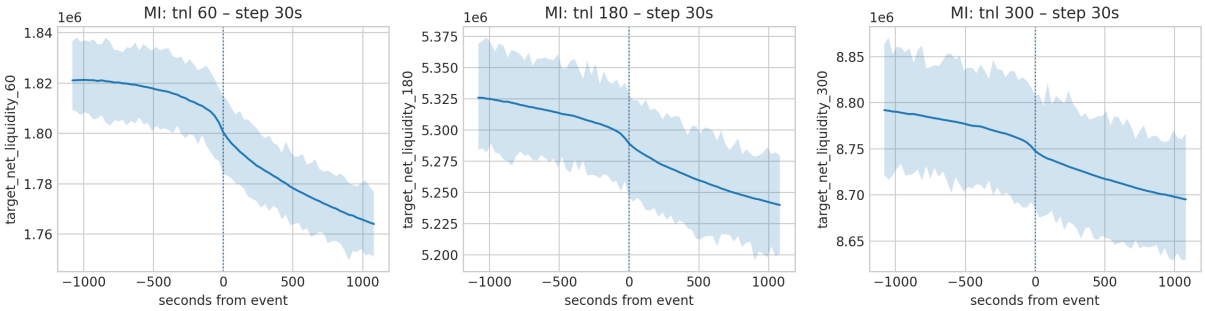


Figure IA.14. Target net liquidity around momentum ignition. The panels plot total displayed liquidity within the target band for $h \in \{60, 180, 300\}$ seconds. Liquidity declines around ignition but does so smoothly, indicating that supply-side adjustment operates mainly through quote protection rather than wholesale withdrawal.

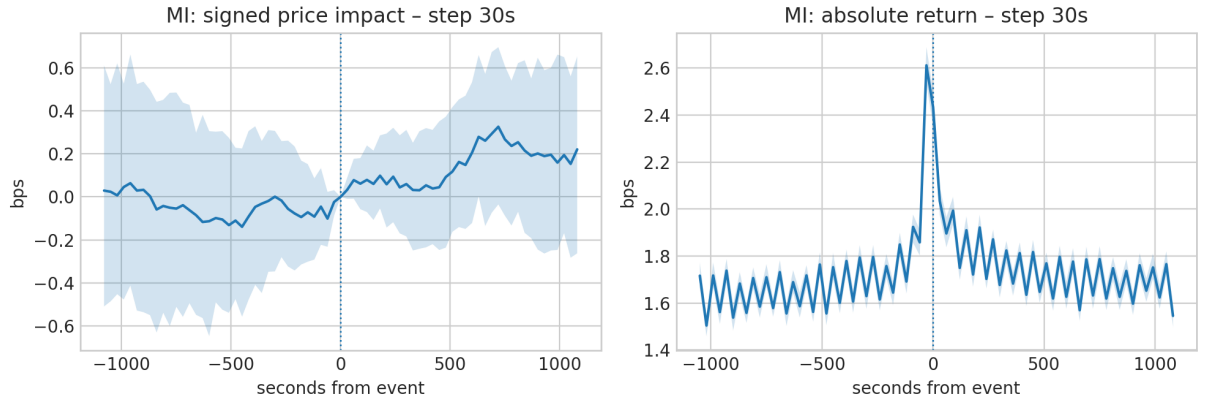


Figure IA.15. Price dynamics around momentum ignition. The left panel reports sign-aligned price impact and the right panel reports absolute returns. Ignition is associated with a sharp volatility burst, while average directional drift is financially smaller and statistically less precise.

IA.D. FPY Depth-Curve Construction and Diagnostics

This appendix gives the mathematical details behind the depth-curve representation used in Section 4.1. The construction follows the logic of FPY curve estimation, with maturity replaced by log distance from the best quote. The object of interest is not a single best bid or ask. It is the entire displayed supply curve, transformed into a smooth low-dimensional state.

IA.D.1. Grid, Transforms, Kernel-Ridge Estimator, and FPY Norm

Let $0 = d_0 < d_1 < \cdots < d_J$ denote depth-bin endpoints measured in ticks from the relevant best quote, and let $u_j = \log(1 + d_j)$ be the log-depth coordinate. If $B_t(d_j)$ and $A_t(d_j)$ are cumulative

bid and ask dollar depth within d_j ticks, the bin increments are

$$\Delta B_t(j) = B_t(d_j) - B_t(d_{j-1}), \qquad \Delta A_t(j) = A_t(d_j) - A_t(d_{j-1}). \tag{IA.17}$$

Throughout this appendix, FPY refers to the framework of Filipović, Pelger, and Ye (Filipović et al., 2022, 2024), whose estimator and theoretical results are applied here to order-book depth curves. We convert increments into log-depth densities,

$$b_t(u_j) = \frac{\Delta B_t(j)}{\Delta u_j}, \qquad a_t(u_j) = \frac{\Delta A_t(j)}{\Delta u_j}, \qquad \Delta u_j = u_j - u_{j-1}, \tag{IA.18}$$

and then work with liquidity and imbalance density curves

$$\ell_t(u_j) = b_t(u_j) + a_t(u_j), \qquad \iota_t(u_j) = b_t(u_j) - a_t(u_j). \tag{IA.19}$$

The transformations

$$x_t^L(u_j) = \log\{1 + \ell_t(u_j)\}, \qquad x_t^I(u_j) = \operatorname{asinh}\{\iota_t(u_j)\} \tag{IA.20}$$

stabilize heavy-tailed depth while preserving the sign and ordering of imbalance. For each curve type we estimate a latent smooth function g_t by the weighted kernel-ridge problem

$$\widehat{g}_t = \underset{g \in \mathcal{H}_{\alpha,\delta}}{\arg\min} \sum_{j=1}^{J} w_j(\alpha)\{x_t(u_j) - g(u_j)\}^2 + \lambda \|g\|^2_{\mathcal{H}_{\alpha,\delta}}, \qquad w_j(\alpha) \propto \exp(-\alpha u_j), \tag{IA.21}$$

where the FPY norm is

$$\|f\|^2_{\mathcal{H}_{\alpha,\delta}} = f(0)^2 + \int_0^{u_J} e^{\alpha u} \left[(1-\delta)\{f''(u)\}^2 + \delta\{f'(u)\}^2\right] du. \tag{IA.22}$$

The sign of the exponential tilt gives more weight to the financially informative region near the touch. The derivative penalty rules out curve fits that chase isolated stale quotes in the deep book.

IA.D.2. Exact recovery of bid and ask depth from liquidity and imbalance

Fix a time t and depth grid $0 = d_0 < d_1 < \cdots < d_J$. Let $b_t(u_j)$ and $a_t(u_j)$ denote bid and ask density in the log-depth coordinate. Define

$$\ell_t(u_j) = b_t(u_j) + a_t(u_j), \qquad \iota_t(u_j) = b_t(u_j) - a_t(u_j). \tag{IA.23}$$

The mapping between (b_t, a_t) and (ℓ_t, ι_t) is pointwise linear.

PROPOSITION IA.1 (Exact bid-ask recovery): *For each depth point u_j, the bid and ask densities*

are uniquely determined by

$$b_t(u_j) = \frac{\ell_t(u_j) + \iota_t(u_j)}{2}, \qquad a_t(u_j) = \frac{\ell_t(u_j) - \iota_t(u_j)}{2}. \tag{IA.24}$$

Consequently, cumulative bid and ask depth on the original grid are recovered by

$$B_t(d_j) = \sum_{r=1}^{j} b_t(u_r)\Delta u_r, \qquad A_t(d_j) = \sum_{r=1}^{j} a_t(u_r)\Delta u_r. \tag{IA.25}$$

Proof. At a fixed grid point, write the linear system

$$\begin{pmatrix} \ell_t(u_j) \\ \iota_t(u_j) \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} b_t(u_j) \\ a_t(u_j) \end{pmatrix}. \tag{IA.26}$$

The determinant is -2, so the matrix is invertible. Multiplication by the inverse gives Equation (IA.24). Cumulative depth follows by summing density times bin width over the log-depth grid. □

In the empirical implementation we smooth transformed curves $x_t^L = \log(1 + \ell_t)$ and $x_t^I = \operatorname{asinh}(\iota_t)$. The inverse transformations are

$$\ell_t = \exp(x_t^L) - 1, \qquad \iota_t = \sinh(x_t^I). \tag{IA.27}$$

The exact recovery theorem therefore applies after inverse transformation. When a low-rank approximation produces a small negative bid or ask density because of approximation error, the reconstruction is truncated at zero only for visualization. The factor scores and regressions are always computed in transformed curve space.

Error propagation from low-rank factors to depth reconstruction. Let $\widehat{x}_t^L(k)$ and $\widehat{x}_t^I(k)$ denote k-factor approximations to the transformed liquidity and imbalance curves. Suppose the true transformed curves lie in the bounded set $\|x_t^L\|_\infty \le M_L$ and $\|x_t^I\|_\infty \le M_I$, and the approximations also lie in the same bounded set. Define inverse-transformed approximations

$$\widehat{\ell}_t(k) = \exp\{\widehat{x}_t^L(k)\} - 1, \qquad \widehat{\iota}_t(k) = \sinh\{\widehat{x}_t^I(k)\}. \tag{IA.28}$$

Depth reconstruction error. For any vector norm dominated by the Euclidean norm up to a finite grid constant, the reconstructed bid-density error satisfies

$$\|\widehat{b}_t(k) - b_t\|_2 \le \frac{1}{2}\left[e^{M_L}\|\widehat{x}_t^L(k) - x_t^L\|_2 + \cosh(M_I)\|\widehat{x}_t^I(k) - x_t^I\|_2\right], \tag{IA.29}$$

and the same bound holds for ask-density error. If C is the cumulative summation operator that maps bin densities to cumulative depth, then

$$\|\widehat{B}_t(k) - B_t\|_2 \le \|C\|_{op}\|\widehat{b}_t(k) - b_t\|_2, \qquad \|\widehat{A}_t(k) - A_t\|_2 \le \|C\|_{op}\|\widehat{a}_t(k) - a_t\|_2. \tag{IA.30}$$

The mean-value theorem gives

$$|\exp(x) - \exp(y)| \le e^{M_L}|x - y| \tag{IA.31}$$

whenever x and y lie in $[-M_L, M_L]$. It also gives

$$|\sinh(x) - \sinh(y)| \le \cosh(M_I)|x - y| \tag{IA.32}$$

whenever x and y lie in $[-M_I, M_I]$. Applying these inequalities componentwise and using the exact-recovery result above yields Equation (IA.29). The cumulative bounds follow from linearity of cumulative summation and the definition of the operator norm.

The proposition explains why a low-dimensional factor representation can be used for state classification. The factor approximation error translates directly into bounded errors in reconstructed bid and ask depth. A spoof-like residual is therefore financially interpretable as the part of displayed depth that cannot be explained by the smooth low-rank book state.

Weighted RKHS estimator and representer form. The estimator and the results in this and the next paragraph are taken from Filipović et al. (2022) and are invoked, not new. Let $\mathcal{H}_{\alpha,\delta}$ be the Hilbert space with norm in Equation (IA.22). For fixed $(\alpha, \delta, \lambda)$ and positive weights w_j, define

$$\widehat{g}_t = \operatorname*{arg\,min}_{g \in \mathcal{H}_{\alpha,\delta}} \sum_{j=1}^{J} w_j \{x_t(u_j) - g(u_j)\}^2 + \lambda \|g\|^2_{\mathcal{H}_{\alpha,\delta}}. \tag{IA.33}$$

Existence, uniqueness, and finite representation. If $\lambda > 0$, the weighted RKHS problem has a unique solution. If K is the reproducing kernel of $\mathcal{H}_{\alpha,\delta}$, then the solution has the finite representation

$$\widehat{g}_t(u) = \sum_{j=1}^{J} c_{t,j} K(u, u_j), \tag{IA.34}$$

where the coefficient vector solves a finite linear system.

This is the standard representer-theorem result for kernel ridge regression as used in Filipović et al. (2022). The strict ridge penalty gives uniqueness, and the representer theorem gives the finite expansion because the loss depends on g only through point-evaluation functionals at the observed depth grid.

The derivative penalty in the norm makes the first factors interpretable. Low-order kernel eigenfunctions are smooth directions in depth. High-order eigenfunctions are rough directions. A

localized wall or hole in displayed depth necessarily loads more on rough directions than a parallel shift or slope change. This is the mathematical reason the FPY representation separates baseline depth variation from spoof-like distortions.

PCA benchmark versus shrinkage representation. By the classical Eckart-Young theorem, PCA gives the best rank-k linear reconstruction of the smoothed curve panel, so curve PCA shows how much of the book can be reconstructed, while the kernel-eigenbasis shrinkage representation gives a more conservative state vector by penalizing rough directions. The paper uses this distinction in the event-time summaries.

IA.D.3. Additional FPY reconstruction diagnostics

This subsection reports the diagnostics that discipline the FPY representation used in Section 4.1. The purpose is not only to show that the order book can be compressed. A high-dimensional order book can almost always be compressed mechanically if enough factors are allowed. The harder requirement is that the compression respect the financial object of interest. The estimator should preserve financially meaningful depth movements, suppress fleeting local contamination, and deliver factors that can be interpreted as shifts in liquidity supply rather than as arbitrary rotations of the event-level state vector.

Table IA.IV Variance concentration in the leading depth-curve factors

Curve	Factor 1	Factor 2 incremental	First two cumulative
Liquidity	94.49%	5.47%	99.96%
Imbalance	90.78%	8.92%	99.70%

Notes. Entries are shares of total cross-sectional variation in the RKHS-smoothed curve panel explained by PCA. The cumulative first-two-factor share is the sum of the first two incremental shares up to rounding. The figures establish that level and slope capture essentially all smooth curve variation, while localized high-order residuals remain available to identify wall-like and other non-smooth states.

A two-factor representation is sufficiently rich to recover the persistent level and slope of displayed supply, yet sufficiently sparse that localized deviations are not absorbed into a large collection of weak factors. The residual therefore retains a clear interpretation as depth variation outside the normal smooth state. This is the object used to distinguish broad liquidity adjustment from localized manipulation-like distortions.

Let $x_t(u_j)$ denote either the transformed liquidity curve or the transformed imbalance curve on the log-depth grid. For each candidate smoothing parameter vector $\theta = (\lambda, \alpha, \delta)$, the FPY curve estimate solves

$$\widehat{g}_t(\theta) = \underset{g \in \mathcal{H}_{\alpha,\delta}}{\arg\min} \sum_{j=1}^{J} w_j \{x_t(u_j) - g(u_j)\}^2 + \lambda \|g\|^2_{\mathcal{H}_{\alpha,\delta}}. \tag{IA.35}$$

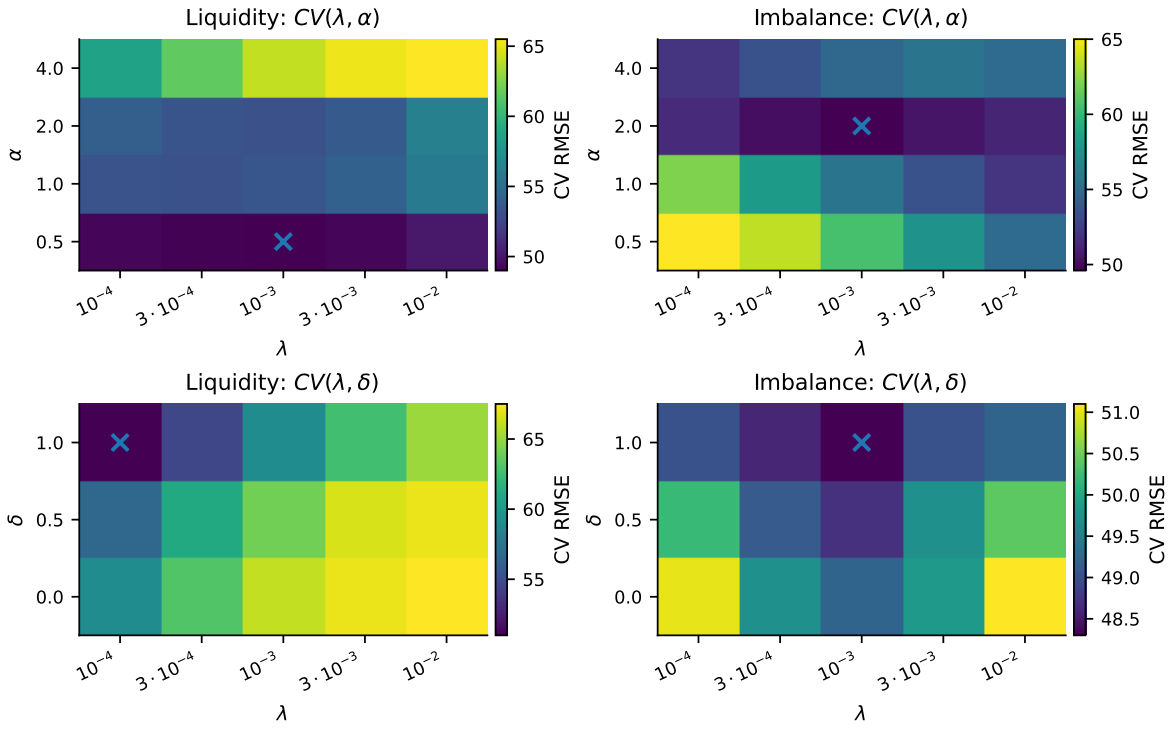


Figure IA.16. Cross-validation sensitivity for FPY curve estimation. The top row varies the ridge penalty and exponential depth weight. The bottom row varies the ridge penalty and the slope-curvature mixing parameter. The cross-validation minima lie in stable troughs rather than isolated points, which means the estimated factor space is not an artifact of a fragile smoothing choice. Economically, the selected region gives enough flexibility to fit near-touch depth while still penalizing localized deep-book spikes.

The ridge parameter λ controls the usual bias-variance tradeoff. The exponential depth parameter α controls how quickly the estimator downweights the deep book, where displayed quantity is less informative about immediate execution. The slope-curvature mixing parameter δ controls whether the norm penalizes gradual tilts or localized bends more aggressively. The tuning criterion is a leave-one-depth-out cross-validation loss,

$$\mathrm{CV}(\theta) = \frac{1}{J} \sum_{j=1}^{J} \left\{ x_t(u_j) - \widehat{g}_t^{(-j)}(u_j; \theta) \right\}^2, \tag{IA.36}$$

averaged over the estimation panel. This criterion is appropriate because the object being reconstructed is a cross-sectional supply curve. The relevant out-of-sample experiment is whether a fitted curve can interpolate omitted depth points without chasing depth-specific noise.

Figure IA.16 is important for identification. Spoofing and momentum ignition are rare states, and the FPY factors are used partly to distinguish abnormal book shapes from the ordinary breathing of depth. A smoothing parameter chosen only to maximize in-sample reconstruction could absorb the very localized distortions we want to detect. A parameter chosen too conservatively could

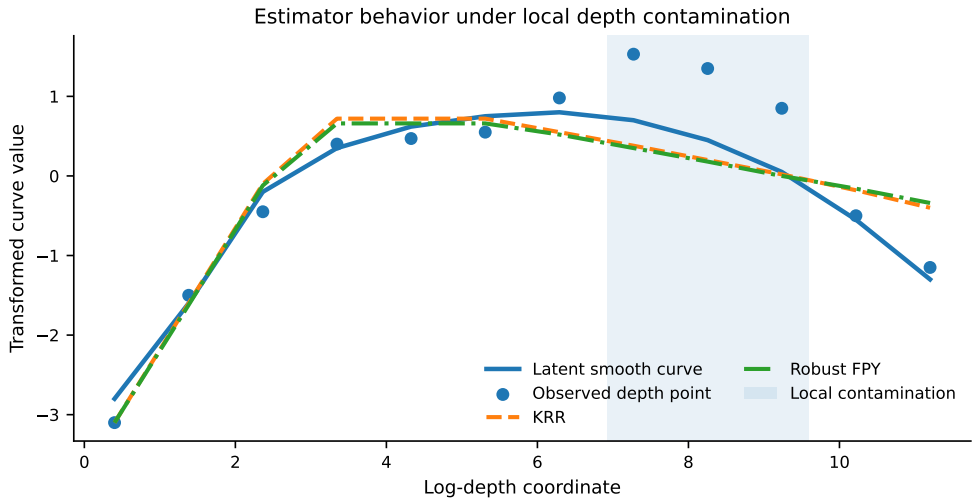


Figure IA.17. FPY smoothing under local depth contamination. The figure shows a stylized local outlier in a depth curve and compares the truth, the contaminated observations, and smoothed estimates. The robust FPY fit remains close to the smooth curve outside the contaminated neighborhood, which is the desired behavior when fleeting displayed depth is not persistent enough to represent true executable supply.

wash out financially meaningful near-touch liquidity changes. The cross-validation surface shows that the selected region is in neither extreme. Reconstruction quality is stable over neighboring values, so the empirical results are not driven by a knife-edge choice of λ, α, or δ.

Figure IA.17 gives the diagnostic that is closest to the spoofing application. A spoof-like message is not a parallel shift in liquidity. It is a local deformation in displayed depth that is often cancelled quickly. A useful curve estimator should therefore not interpret every isolated depth point as a permanent movement in latent liquidity supply. The robust FPY estimate has exactly this property. It absorbs smooth information in the neighborhood of the book but does not allow a single contaminated point to dominate the factor state. High-order residuals are the part of displayed depth that cannot be rationalized by the low-dimensional smooth book.

The loading shapes in Figure 7 make the model financially interpretable. A level factor corresponds to broad replenishment or withdrawal of displayed liquidity. A slope factor distinguishes near-touch depth from deeper posted quantity. A curvature factor captures local convexity of the supply curve. These are the directions that matter for execution costs, adverse selection, and the strategic placement of displayed orders. By contrast, arbitrary principal components of a raw order-book vector can mix the bid side, the ask side, near-depth noise, and deep-book noise in ways that are difficult to map into financial economics. The FPY transformation imposes the right geometry before factor extraction.

Figure IA.18 provides the benchmark for the curve approach. Full-state PCA has excellent asymptotic reconstruction once enough factors are included, but this is not the criterion used in the paper. The empirical question is whether a small number of factors can summarize the financially relevant state while leaving localized distortions visible. The curve FPY representation

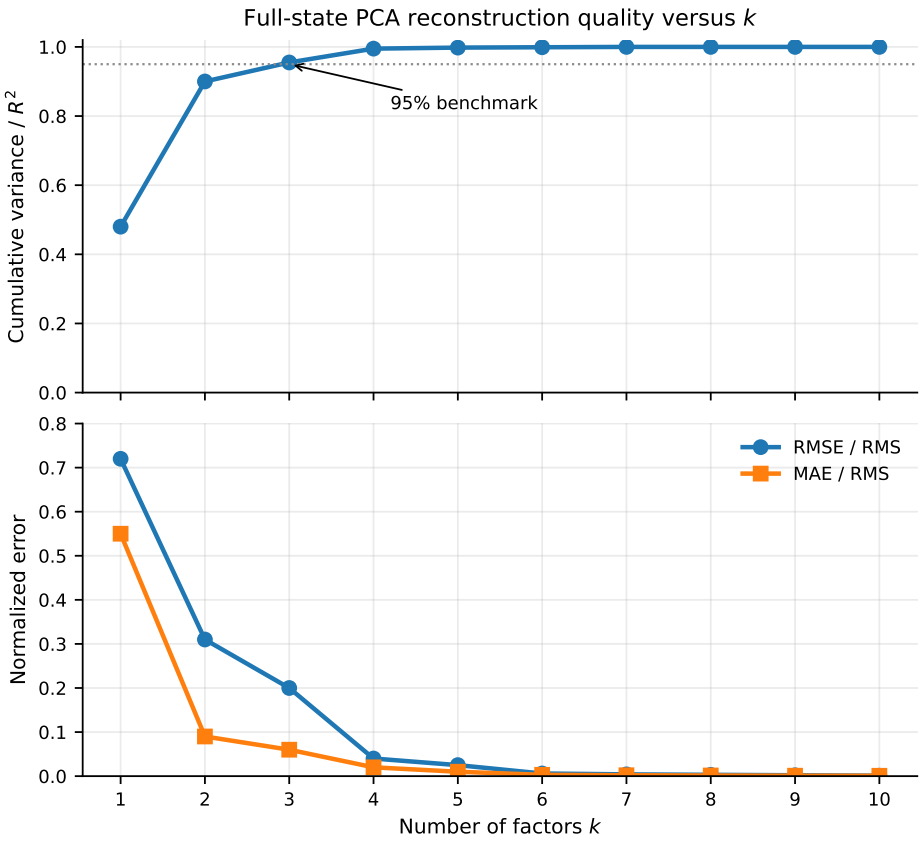


Figure IA.18. Full-state PCA reconstruction quality as the number of factors increases. The top panel plots cumulative variance explained, and the bottom panel reports normalized reconstruction errors. Full-state PCA eventually reconstructs the order book well, but the error reduction is gradual over the first few factors and the factors need not be smooth depth directions. This benchmark is useful because it separates pure reconstruction power from financially interpretable reconstruction.

is biased toward smooth depth directions. That bias is valuable because it produces factors that can be interpreted as liquidity level, liquidity slope, imbalance level, and imbalance slope, while rough residuals remain available as diagnostics for spoof-like and ignition-like states.

The joint conclusion from these diagnostics is that FPY is not used as an arbitrary dimension reducer. It is a structured model of the displayed supply curve. Cross-validation chooses a stable smoothing region, the robust estimator avoids chasing isolated contaminated depth, the loadings have low-order economic shapes, and the full-state benchmark shows why reconstruction alone is not enough. These properties justify the use of FPY factors in Section 4.1 and the interpretation of manipulation states as specific movements in liquidity, imbalance, and high-order residual directions.

IA.D.4. Depth-Factor Interpretation

The factor directions of the transformed liquidity and imbalance curves map into baseline and manipulation states as follows, and the same algebra maps factor approximations back into bid and ask depth through Equation (IA.19).

- **Liquidity level F_1^L.** A parallel movement in displayed depth across both sides of the book. It is high when the book is broadly deep and immediacy is cheap, and it falls when liquidity suppliers protect quotes during ignition or withdraw after distorted depth.
- **Liquidity slope F_2^L.** Reallocates displayed quantity between near-touch and deep depth. It is stable when depth is replenished close to the touch, and it moves when near-touch liquidity thins while deeper depth remains visible.
- **Imbalance level F_1^I.** The broad bid-minus-ask asymmetry across depth. It stays close to its local mean during symmetric supply and demand, and it shifts when one side of the book is made to appear unusually strong.
- **Imbalance slope F_2^I.** Contrasts near-touch and deep asymmetry in the signed book. It is small in quiet periods because imbalance is not concentrated at one depth, and it is large when a wall appears at a selected distance or when aggressive flow exhausts one side near the touch.
- **High-order residual share.** The fraction of depth variation outside the first smooth factors. It is low when the book is a smooth depth curve, and it is high when the state contains localized humps, gaps, or fleeting walls.

Additional Depth-Curve Event-Time Summaries. The quiet-period baseline is shown in the main text (Figure 14 in Section 4.4). There the curve factor and full-RKHS benchmarks are nearly parallel in their first imbalance loading over the informative portion of the book, and the event-time scores remain close to their local means. In a normal state the book is not literally constant. Depth breathes, orders arrive, and cancellations occur, but most variation is carried by smooth level and slope movements rather than by localized depth defects.

Figure 16 gives the corresponding state around momentum-ignition candidates. The response is different from spoofing. The ignition state is not primarily a false wall in the cross-section of displayed depth. It is an aggressive-flow state that changes the near-touch execution environment. The curve factors move around the event, but the interpretation is supply protection following pressure from marketable orders. The depth perspective is important because it explains why the event-study evidence below shows a large increase in effective spreads and human taker intensity even though displayed depth does not collapse. Liquidity suppliers can leave orders in the book while moving the relevant part of the depth curve away from the touch, which raises the marginal cost faced by traders who demand immediacy.

The corresponding event-time summary is shown in the main text (Figure 16 in Section 4.4).

IA.E. News Measurement, Evaluation, and Additional Results

This appendix gives the formal details for Section 5. The goal is to construct features that are measurable at the prediction time and to evaluate them without overlap leakage.

IA.E.1. Measurement: alignment, decay features, and smooth-ridge estimator

Let an article have timestamp τ_i in continuous time, measured in minutes. Let the prediction grid be the integer minute grid. Define

$$m(i) = \min\{m \in \mathbb{Z} : m \geq \tau_i\}. \tag{IA.37}$$

For any prediction minute m, the article is included in the information set only if $m(i) \leq m$.

No-lookahead property. If article timestamps are correctly recorded, the feature vector X_m constructed from articles with $m(i) \leq m$ is measurable with respect to the news filtration at minute m and excludes all articles with publication time after m.

The stricter delay-one-minute robustness design replaces $m(i)$ by $m(i) + 1$. The no-lookahead property continues to hold, and the design becomes conservative when timestamps are rounded or recorded with small latency.

Exponential decay features. The decay features exist because news is digested at different speeds. Some headlines move the market within minutes and others diffuse over hours, and rather than imposing one horizon we carry the same news signal at four fixed half-lives and let the forecasting model weight them. Each feature is simple exponential smoothing of the minute-level impulse. For a given topic impulse $x_{b,m}$ and half-life τ, define $\rho_\tau = 2^{-1/\tau}$. The recursive feature

$$z^{(\tau)}_{b,m} = \rho_\tau z^{(\tau)}_{b,m-1} + x_{b,m} \tag{IA.38}$$

has the distributed-lag representation

$$z^{(\tau)}_{b,m} = \sum_{\ell=0}^{\infty} \rho_\tau^\ell x_{b,m-\ell}, \tag{IA.39}$$

provided the initialization is taken in the distant past or the effect of initialization is discarded after burn-in.

Iterating the recursion gives

$$z^{(\tau)}_{b,m} = x_{b,m} + \rho_\tau x_{b,m-1} + \cdots + \rho_\tau^K x_{b,m-K} + \rho_\tau^{K+1} z^{(\tau)}_{b,m-K-1}. \tag{IA.40}$$

Because $0 < \rho_\tau < 1$, the final term vanishes as $K \to \infty$ when the process is bounded in mean square or after a finite burn-in. This gives Equation (IA.39).

The half-life interpretation follows from $\rho_\tau^\tau = 1/2$. An article impulse therefore receives half its original weight after τ minutes.

Smooth-ridge estimator. With $Y \in \mathbb{R}^{n\times H}$ the matrix of targets, $X \in \mathbb{R}^{n\times p}$ the feature matrix, and $D \in \mathbb{R}^{(H-1)\times H}$ the first-difference matrix across adjacent horizons, the smooth-ridge objective of footnote 16 is $Q(B) = \|Y - XB\|_F^2 + \kappa\|B\|_F^2 + \rho\|BD'\|_F^2$, and for $\kappa > 0$ its unique minimizer satisfies the linear system $(X'X + \kappa I_p)\widehat{B} + \rho\widehat{B}D'D = X'Y$, a standard ridge-type result. This estimator is a plain ridge regression across stacked horizons and does not use the reproducing-kernel machinery of the depth-curve construction.

This estimator and its properties are taken from Filipović et al. (2022), applied here with forecast horizons in place of maturities; we refer to that paper for the analysis.

The linear system is a Sylvester equation, and the positive ridge term makes the solution well determined even when news features are correlated or sparse.

IA.E.2. Evaluation: purged walk-forward design and event-study consistency

For a test block $\mathcal{I}_j$ and horizon h, the training set is

$$\mathcal{T}_j(h) = \{m : m \notin \mathcal{I}_j,\ [m, m+h] \cap \mathcal{I}_j = \varnothing\}. \tag{IA.41}$$

This definition removes every training target whose realized outcome window overlaps the test block.

No label-overlap leakage. For any $m \in \mathcal{T}_j(h)$ and any $m' \in \mathcal{I}_j$, the outcome interval used to construct $Y_{m,h}$ does not contain m'.

This purge is needed because high-frequency forward targets overlap mechanically. Without it, the training set could contain observations whose labels share realized order flow or liquidity outcomes with the test set.

Event-study consistency. The event studies are read as descriptive reaction profiles, with the identifying content in the no-lookahead alignment, the purge rule, and the day-block bootstrap. What they do not control for, and do not claim to, are confounders arriving in the same minute as a headline, which is why the interpretation rests on the trend-split and cross-exchange designs.

IA.E.3. Additional News Panels, Examples, and Feature Comparisons

Figure IA.19 describes the news panel. Figure IA.20 compares the FinBERT and Gemini score distributions on the matched sample. The worked headline examples, the feature decomposition, and the model choices are discussed in the main text, and Figure IA.21 shows the FPY depth-curve state around high-news-intensity periods. Table IA.V reports the timing and novelty-weight robustness of the news predictability results.

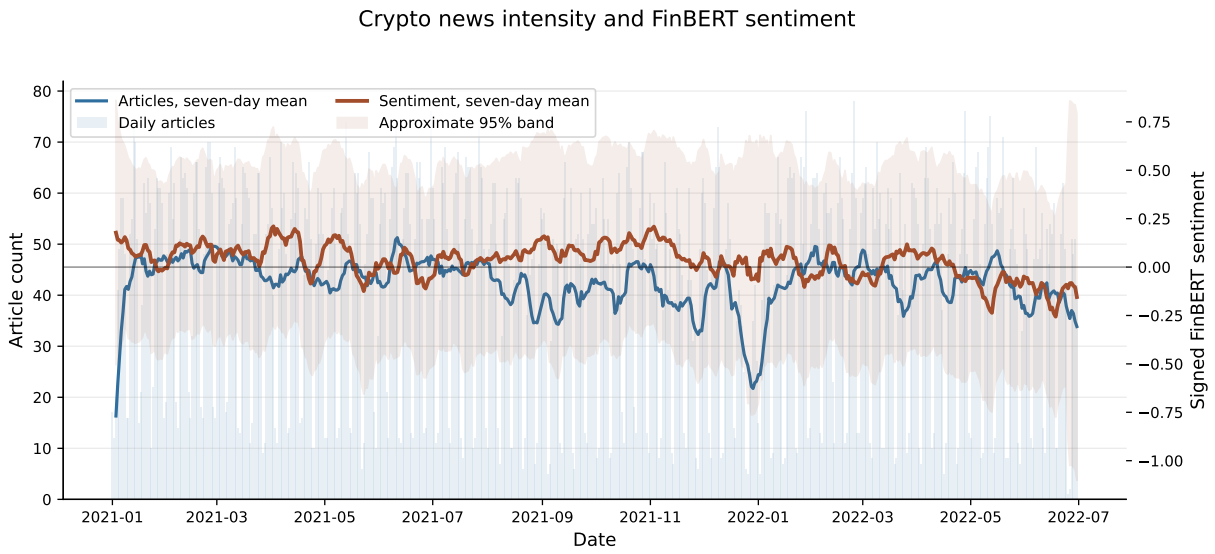


Figure IA.19. Crypto news intensity and FinBERT sentiment. Bars report daily article counts from the two-source news panel. The blue line is the seven-day moving average of article counts. The brown line is the seven-day moving average of signed FinBERT sentiment, with an approximate confidence band for bounded scores. The figure shows that sentiment variation is not reducible to publication intensity.

The worked headline examples are discussed in Section 5.3 of the main text.

Each empirical choice in the news design pairs the economics of public information arrival with the statistical problem created by high-frequency overlapping outcomes. Next-minute alignment is used because a headline cannot enter the information set before its timestamp and crypto venues trade continuously without a market open that would absorb overnight news, which eliminates mechanical look-ahead in the tradition of high-frequency event studies (MacKinlay, 1997; Foucault et al., 2016). Finance-domain sentiment is used because crypto headlines contain financial language, market jargon, and rhetorical price-action phrases that generic language models can misclassify, and FinBERT is a reproducible, auditable benchmark (Devlin et al., 2019; Araci, 2019; Loughran and McDonald, 2011). Novelty weights are used because multiple outlets and follow-up stories repeat the same information, and duplicate counting would confuse information intensity with syndication intensity, a conservative analogue of relevance filtering in text-based asset pricing (Tetlock et al., 2008; Boudoukh et al., 2013). Topic buckets and the decay features are used because different types of headlines have different response horizons, and exponential kernels approximate a distributed lag with few parameters. Ridge and smooth ridge are used because public news signals are many and correlated while adjacent forecast horizons should have related responses. Purged walk-forward validation is used because the relevant question is whether news known at time m predicts future book states in held-out time blocks, and purging removes overlapping-label leakage so the reported R^2 is a genuine out-of-sample statistic.

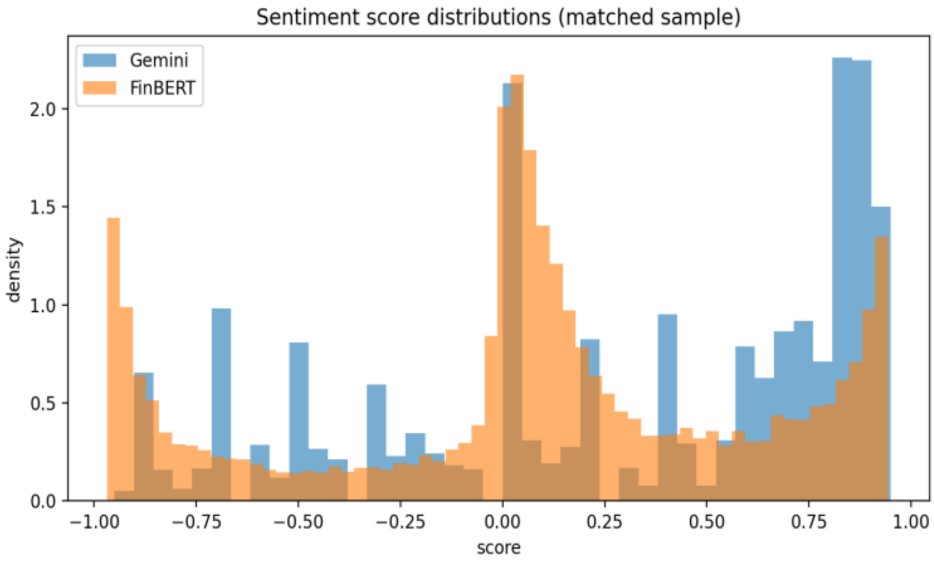


Figure IA.20. Distribution of headline sentiment scores. The matched-sample distribution shows that the Gemini score assigns more extreme polarity, while FinBERT is more conservative and more concentrated near zero. This comparison motivates using strong-score event studies for return interpretation and FinBERT for the main predictive microstructure design.

Table IA.V News Predictability under Alignment and Novelty Variants

Target	Baseline	Delay one minute	No novelty weights
Non-API net liquidity	0.410	0.410	0.409
Non-API taker buys	0.311	0.311	0.311
Non-API taker net	0.132	0.131	0.130

Notes: Entries are horizon averages of R^2_{OS} for the FinBERT-Topics smooth-ridge specification. Delay one minute maps each article to the next full minute after release. No novelty weights sets duplicate downweighting to one.

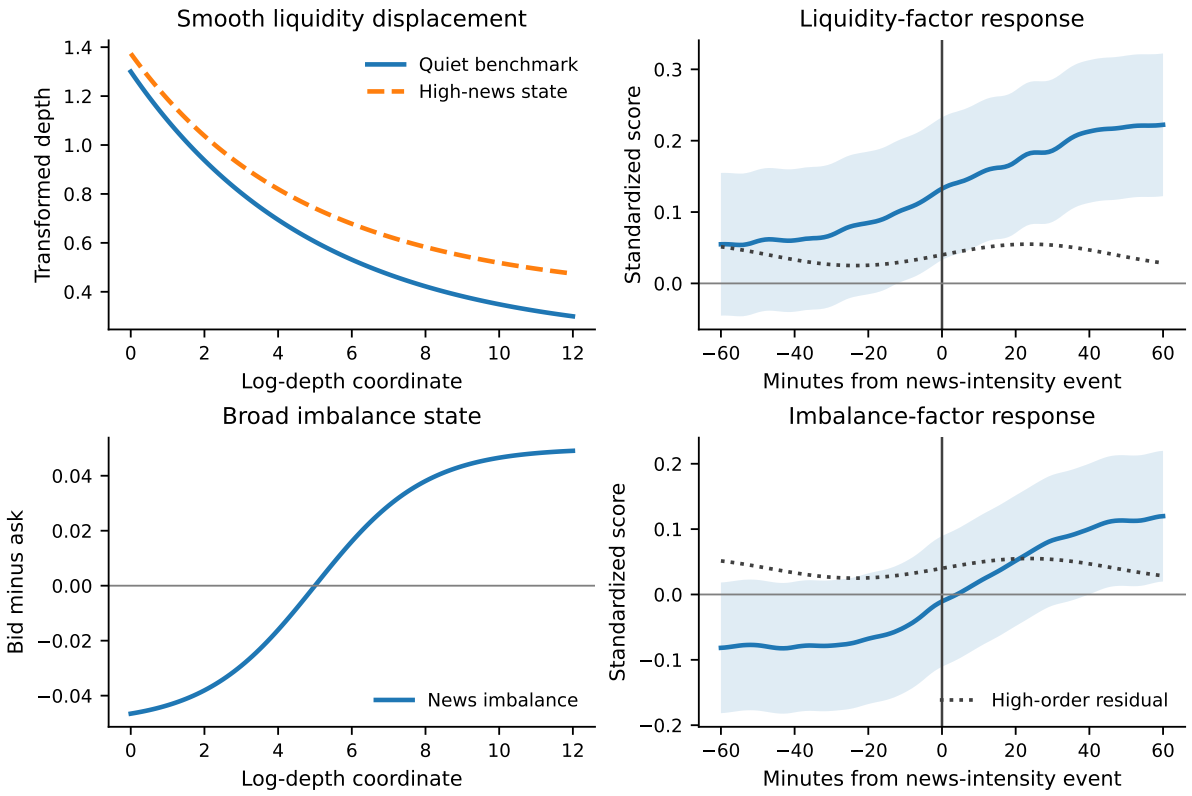


Figure IA.21. FPY depth-curve state around high-news-intensity periods. The news state produces a smoother factor response than spoofing, which is consistent with broad public information rather than localized displayed-depth distortion.

IA.F. Cross-Exchange Shocks: Additional Results

Figure IA.22 plots the one-second cross-exchange lead-lag function estimated in the main text. Figure IA.23 reports the per-event distribution of the standardized cohort difference behind every cell of the cross-exchange shock analysis. The clipped means and the unclipped medians agree in sign and approximate magnitude on every cell. In the event-time paths, API maker limit buys in top-tail events rise from a pre-event mean of \$0.94M to \$1.50M at $t = 0$ (+59%), and non-API maker limit buys rise from \$2.6k to \$3.9k (+47%); the bottom-tail sell-side measures behave symmetrically.

Table IA.VI reports the full with-the-move sensitivity analysis across clipping thresholds $c \in \{10, 25, 50, 100\}$. Cancellation estimates move only from about 1.18 to 1.13 in the bottom tail and from 1.16 to 1.12 in the top tail, and maker-side match estimates stay between 0.12 and 0.14. Taker-side match estimates remain large and positive at every threshold. The maker-limit magnitudes are more sensitive, declining from about 1.11 at $c = 10$ to 0.77 at $c = 100$ in the bottom tail and from 1.14 to 0.81 in the top tail. The conclusion supported by the table is stability of sign and economic ordering, not literal invariance of every coefficient. At the baseline $c = 50$, no more than 0.60% of observations are clipped in any cell, and against-the-move placebo estimates remain close to zero at all four thresholds.

Table IA.VI Cross-exchange $\hat{D}(Y)$ sensitivity to clip threshold, with-the-move cells.

Measure	Tail	$c = 10$	$c = 25$	$c = 50$	$c = 100$
Cancellations	Bottom	1.180	1.164	1.150	1.132
Cancellations	Top	1.164	1.149	1.137	1.124
Maker limit orders	Bottom	1.110	0.942	0.839	0.769
Maker limit orders	Top	1.136	0.978	0.879	0.806
Maker-side matches	Bottom	0.122	0.134	0.136	0.134
Maker-side matches	Top	0.118	0.127	0.127	0.125
Taker-side matches	Bottom	1.246	1.422	1.446	1.407
Taker-side matches	Top	1.264	1.399	1.395	1.340

Notes. Point estimates for $\hat{D}(Y)$ in equation (36) across four clip thresholds $c \in \{10, 25, 50, 100\}$. $N \approx 470{,}000$ events per tail. The paper's headline uses $c = 50$. Against-the-move (placebo) cells are near zero in absolute magnitude at every threshold and are omitted from the table. The maximum fraction of events clipped across all 16 cells is 5.09% at $c = 10$, 1.48% at $c = 25$, 0.60% at $c = 50$, and 0.25% at $c = 100$. Confidence intervals are day-block bootstrap 95% percentile intervals over 2,000 resamples; day-block CI half-widths at $c = 50$ are under 0.14 on every cell.

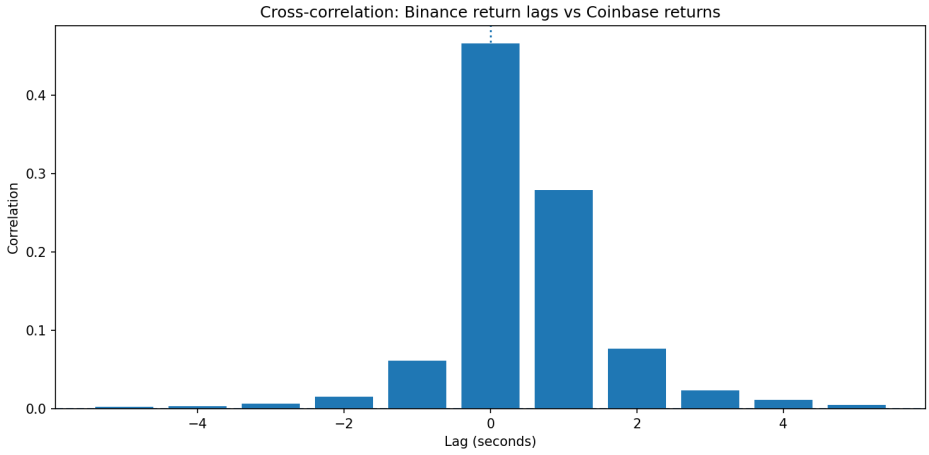


Figure IA.22. Cross-exchange lead-lag at one-second resolution. Sample cross-correlation $\Gamma_{B,C}(\ell)$ between one-second log returns on Binance BTCUSDT and Coinbase BTC–USD, January 2021 through June 2022. Positive ℓ compares the current Binance return to a future Coinbase return.

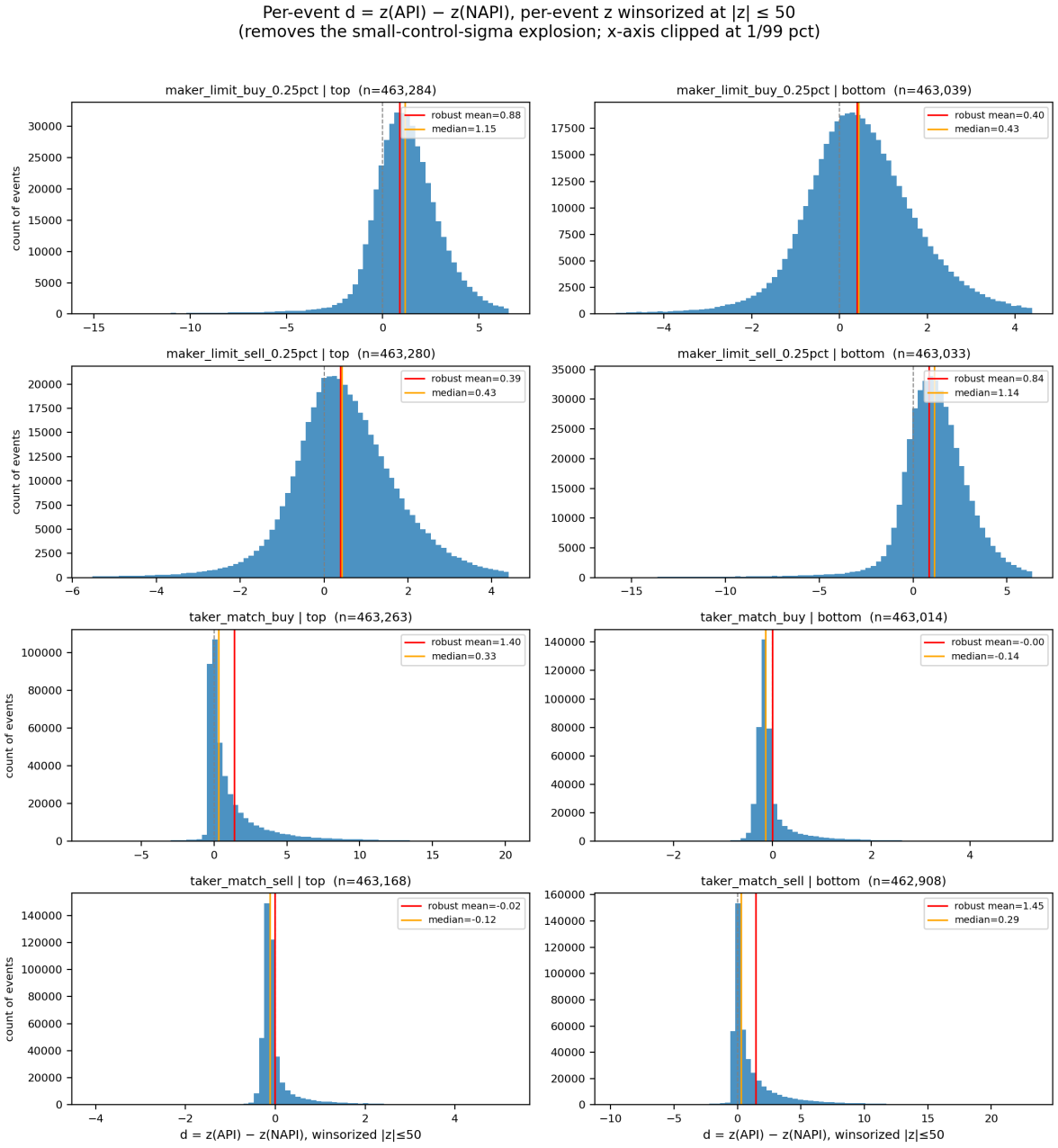


Figure IA.23. Per-event distribution of $d_i = z_i^{\text{API}} - z_i^{\text{NAPI}}$ for the four near-book measures, on each tail, after the per-event z-score is clipped at $|z| \leq 50$. Vertical lines mark the event-level mean (red, $\hat{D}(Y)$ in equation (36)) and the unclipped event-level median (orange). The clipped mean and the unclipped median agree in sign and approximate magnitude on every cell.

IA.G. Investor Cost Accounting: Derivations and Additional Results

This appendix gives the formal details behind Section 6. The objects are event-level functionals of observed trades, quotes, and midquote returns. The purpose is to make explicit which parts of the cost accounting are identities, which parts are estimators, and which parts require counterfactual assumptions.

IA.G.1. Cost identities, basis-point aggregation, and impact diagnostic

Let $m_i = m(t_i)$ and define the relative signed half-spread $s_i = q_i(p_i - m_i)/p_i$. Because a taker buy has $q_i = +1$ and a taker sell has $q_i = -1$, s_i is positive exactly when the taker executes on the costly side of the midquote.

Taker cost as realized implementation shortfall. For any trade set I with $D(I) > 0$, let $\omega_i = p_i v_i / D(I)$. Then

$$TC(I) = 10^4 \sum_{i \in I} \omega_i s_i \tag{IA.42}$$

is the notional-weighted average signed half-spread paid by takers, measured in basis points. If executions occur at the best quotes, $TC(I)$ equals one half of the notional-weighted effective spread.

The relative shortfall for trade i is $q_i(p_i - m_i)/p_i$. Weighting by $p_i v_i / D(I)$ gives

$$\sum_{i \in I} \frac{p_i v_i}{D(I)} q_i \frac{p_i - m_i}{p_i} = \frac{\sum_{i \in I} q_i (p_i - m_i) v_i}{D(I)}. \tag{IA.43}$$

Multiplication by 10^4 expresses the rate in basis points. At the best quotes, $2q_i(p_i - m_i)/p_i$ is the relative effective spread, which proves the half-spread identity.

Maker markout as adverse-selection loss. For a passive maker fill i, $r_i^M(\tau) > 0$ if and only if the midquote at horizon τ moves against the passive side of the trade. Hence $MC(I) = 10^4 \sum_{i \in I} \omega_i r_i^M(\tau)$ is the notional-weighted adverse-selection markout in basis points.

If the maker buys, a subsequent decline gives $p_i - m(t_i + \tau) > 0$. If the maker sells, a subsequent increase gives $m(t_i + \tau) - p_i > 0$. Dividing by p_i converts each outcome into a relative loss. Notional weighting and multiplication by 10^4 produce the basis-point rate.

Notional-weighted basis-point aggregation. Let $Y_e^T = \Delta TC_e^{\text{DiD}}$ and $Y_e^M = \Delta MC_e^{\text{DiD}}$. For an event family $\mathcal{E}_j$, define

$$H_j^T = \frac{\sum_{e\in\mathcal{E}_j} D_e^+ (Y_e^T)^+}{\sum_{e\in\mathcal{E}_j} D_e^+}, \tag{IA.44}$$

$$H_j^M = \frac{\sum_{e\in\mathcal{E}_j} D_e^+ (Y_e^M)^+}{\sum_{e\in\mathcal{E}_j} D_e^+}, \tag{IA.45}$$

$$H_j^{\text{gross}} = H_j^T + H_j^M, \qquad H_j^{\text{net}} = \frac{\sum_{e\in\mathcal{E}_j} D_e^+ (Y_e^T + Y_e^M)}{\sum_{e\in\mathcal{E}_j} D_e^+}. \tag{IA.46}$$

All four quantities are measured in basis points.

Component-level gross harm is monotone in harmful rate changes. For fixed nonnegative weights D_e^+, H_j^{gross} is nonnegative and weakly increasing in each harmful rate change. A favorable movement in one component does not erase a harmful movement in the other component.

The maps $x \mapsto x^+$ and $y \mapsto y^+$ are nonnegative and weakly increasing. Multiplication by $D_e^+ \geq 0$, summation, and division by positive total notional preserve these properties.

This is the formal reason for applying the positive part before aggregation. The result is an incidence rate for exposed accounts rather than a social-surplus measure.

IA.G.2. Matched-Control DiD Identification

Let $Y_e(1)$ denote the post-minus-pre cost-rate change that occurs for event e under treatment, and let $Y_e(0)$ denote the post-minus-pre change that would have occurred in the same state absent the event. Let X_e be the vector of pre-event state variables used for matching. For controls $c(e,k)$, let $Y_{c(e,k)}(0)$ be their realized non-event post-minus-pre cost-rate changes.

ASSUMPTION IA.3 (Conditional counterfactual comparability): *For each event e,*

$$\mathbb{E}\{Y_e(0) \mid X_e\} = \mathbb{E}\{Y_c(0) \mid X_c = X_e\}, \tag{IA.47}$$

and the matched controls satisfy $\max_k \|X_{c(e,k)} - X_e\| \to 0$ *as the control pool grows.*

Consistency of the matched-control drift correction. Under Assumption IA.3 and standard moment conditions, the matched-control average

$$\widehat{\mu}_0(X_e) = \frac{1}{K}\sum_{k=1}^{K} Y_{c(e,k)}(0) \tag{IA.48}$$

converges in probability to $\mathbb{E}\{Y_e(0) \mid X_e\}$. Consequently, $Y_e(1) - \widehat{\mu}_0(X_e)$ consistently estimates the event-level abnormal rate change relative to the matched counterfactual mean.

By conditional comparability, the conditional mean of matched controls at X_e equals the event counterfactual mean. Because the matching discrepancy converges to zero and the conditional mean is continuous in the state variables, the bias from matching on nearby rather than identical states vanishes. The law of large numbers for the matched control draws gives convergence of the control average to its conditional expectation. Combining the two steps gives the result.

The assumption is weaker than a fully specified structural model of execution costs. It requires only that same-day non-event windows with similar spread, depth, volatility, notional, imbalance, and time-of-day states provide the local drift in trading costs absent the event.

Kyle-λ impact-rate diagnostic. In the Kyle diagnostic, the one-second midquote change is projected on signed net flow,

$$\Delta m_s = a + \lambda Q_s + u_s. \tag{IA.49}$$

For a reference parent order of signed size Q^{ref} and initial price p_0, the linear price-impact schedule is $p(q) = p_0 + \lambda q$ along cumulative quantity $q \in [0, Q^{\text{ref}}]$.

Linear-impact rate for a reference order. The average impact paid relative to initial notional is

$$IC^{\text{bps}}(\lambda, Q^{\text{ref}}, p_0) = 10^4 \frac{\int_0^{Q^{\text{ref}}} \lambda q \, dq}{p_0 Q^{\text{ref}}} = 5{,}000 \frac{\lambda Q^{\text{ref}}}{p_0}. \tag{IA.50}$$

The event-level change is therefore

$$\Delta IC_e^{\text{bps}} = 5{,}000 \frac{(\widehat{\lambda}_e^{+} - \widehat{\lambda}_e^{-}) Q^{\text{ref}}}{p_0}. \tag{IA.51}$$

Integrating the marginal price displacement gives $\frac{1}{2}\lambda (Q^{\text{ref}})^2$, and dividing by initial notional $p_0 Q^{\text{ref}}$ and multiplying by 10^4 yields the stated basis-point rate.

This diagnostic is not designed to equal the realized-cost estimator. It abstracts from realized exposure, spread changes, passive markouts, and the identity of exposed traders.

IA.G.3. Volatility channel: expected shortfall, capacity, and liquidation risk

For a 60-second window, RV estimates local quadratic variation. If returns are continuous semimartingales with sufficiently regular volatility, realized variation converges in probability to integrated variance under infill sampling. Bipower variation estimates the continuous component of variation, and $JV = (RV - BV)^+$ isolates jump-like residual variation in the finite sample. The paper uses these objects as transparent nonparametric inputs to the risk-capital mapping rather than as a full model of volatility dynamics.

Gaussian expected shortfall in basis points. Let $R \sim N(0, \sigma^2)$ and let relative loss be $L = -R$. Then

$$ES_\alpha^{\text{bps}}(L) = 10^4 \mathbb{E}[L \mid L \geq VaR_\alpha(L)] = 10^4 \sigma \frac{\varphi(z_\alpha)}{1-\alpha}, \qquad z_\alpha = \Phi^{-1}(\alpha). \tag{IA.52}$$

Write $L = \sigma Z$ for $Z \sim N(0,1)$. The standard truncated-normal identity gives $\mathbb{E}[Z \mid Z \geq z_\alpha] = \varphi(z_\alpha)/(1-\alpha)$. Multiplying by σ and then by 10^4 gives the result.

The basis-point formulation is scale free. It reports the tail-loss rate of the position rather than introducing a reference notional amount or aggregating hypothetical exposure across events.

Participation capacity under a fixed risk budget.

Capacity loss. If an investor faces the constraint $ES_\alpha(N, \sigma_h) \leq B$, then maximum admissible notional is $N^{\max}(\sigma_h) = B/(\kappa_\alpha \sigma_h)$, where $\kappa_\alpha = \varphi(z_\alpha)/(1-\alpha)$. The positive proportional capacity loss from a counterfactual volatility σ^{cf} to post-event volatility σ^+ is

$$L^{\text{cap}} = \left[1 - \frac{\sigma^{\text{cf}}}{\sigma^+}\right]^+. \tag{IA.53}$$

By the expected-shortfall expression above, the constraint is $N\kappa_\alpha\sigma_h \leq B$. Solving for the largest feasible N gives $N^{\max}(\sigma_h) = B/(\kappa_\alpha\sigma_h)$. The proportional decline in capacity is

$$1 - \frac{N^{\max}(\sigma^+)}{N^{\max}(\sigma^{\text{cf}})} = 1 - \frac{\sigma^{\text{cf}}}{\sigma^+}. \tag{IA.54}$$

Taking the positive part reports only capacity losses rather than capacity gains.

The result shows why volatility shocks are also liquidity shocks. A risk-constrained liquidity supplier must reduce displayed size one-for-one with the increase in horizon volatility, all else equal.

Liquidation probability under a Brownian barrier proxy. Let $X_t = \sigma W_t$ be a zero-drift Brownian return process over horizon h, and let $\tau_{-a} = \inf\{t \leq h : X_t \leq -a\}$ be the first time the return breaches a downside margin barrier $a > 0$.

Reflection-principle liquidation proxy. Under the Brownian proxy,

$$\mathbb{P}(\tau_{-a} < h) = 2\Phi\left(-\frac{a}{\sigma\sqrt{h}}\right). \tag{IA.55}$$

This is the classical reflection-principle bound for Brownian motion with drift.

For liquidation penalty rate π, the expected penalty rate is

$$LP^{\text{bps}} = 10^4 \pi \mathbb{P}(\tau_{-a} < h). \tag{IA.56}$$

The event-attributable proxy is the positive difference between post-event and counterfactual penalty rates. The approximation abstracts from exchange-specific liquidation rules, maintenance margin schedules, funding, and discrete liquidation cascades. Its role is to quantify the convexity of leveraged exposure to the same volatility shocks measured above.

IA.G.4. Additional Cost Tables and Figures

The estimator choices trade off as follows. Signed slippage and implementation shortfall measure the realized cost paid by aggressive traders relative to the local midquote, observed trade by trade, which makes them the primary taker component, although they do not value unexecuted orders or latent private values. Effective and realized spreads are standard market-quality rates comparable across markets, used here as diagnostics and aggregated on the event-window exposure denominator, because unweighted rates ignore how much exposure is concentrated in stress states. Short-horizon markouts capture the adverse selection faced by passive liquidity suppliers and form the primary maker component with $\tau = 10$ seconds, with the caveat that markouts depend on the horizon and can mix inventory with information effects. The matched-control DiD supplies the local counterfactual drift and is the primary estimator for realized investor incidence, though it cannot fully separate overlapping stressors if several occur at once. Kyle-λ serves as a mechanism diagnostic rather than a headline cost estimator, since a short-window impact slope can be weakly identified. Gross accounting measures adverse incidence for exposed accounts and net accounting allows favorable outcomes to offset losses, and the two are reported side by side because gross is not a social-surplus measure while net can conceal incidence. Expected shortfall supplements the execution rates with a tail-loss rate for inventory exposure. Figure IA.24 accumulates the pooled abnormal rate over calendar time. Figure IA.25 reports gross harm by event type, Figure IA.26 compares estimator rankings, Figure IA.27 places execution and volatility-risk incidence in common basis-point units, Figure IA.28 reports the per-event expected-shortfall distribution, Figure IA.29 displays tail concentration, Figure IA.30 reports participation-capacity losses and the leveraged liquidation-rate proxy, and Figure IA.31 plots liquidation-like stress windows on known liquidation dates.

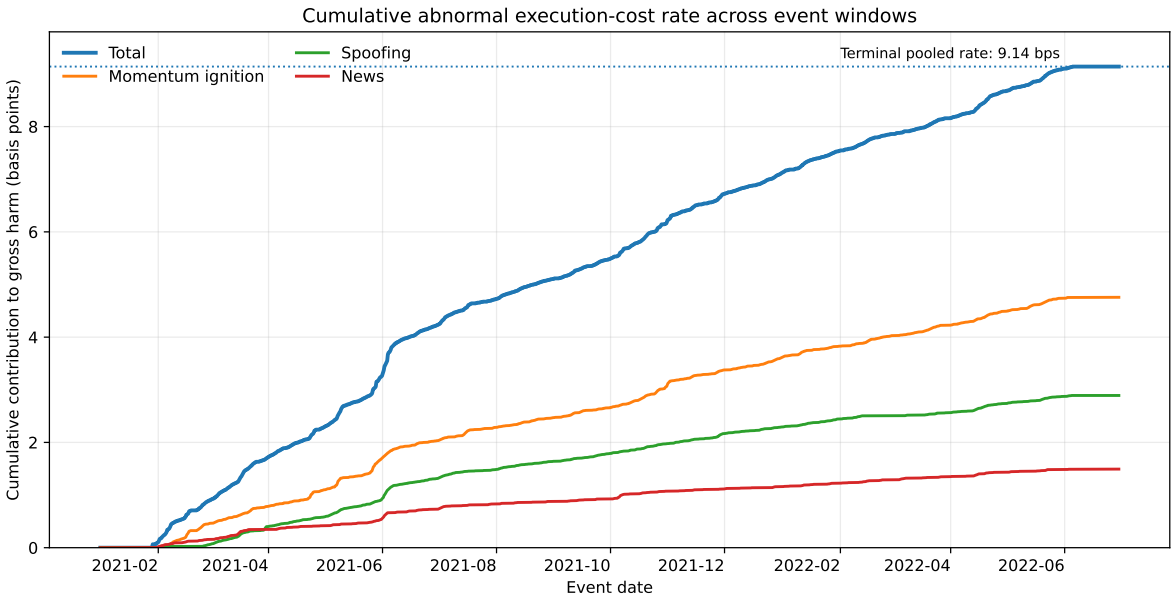


Figure IA.24. Cumulative abnormal execution-cost rate. The figure plots each event family's contribution to the pooled matched-control gross rate. All paths use pooled post-event human taker notional as the fixed denominator, so the terminal cross-type value is 9.14 basis points. The step-like path shows that incidence is persistent but tail driven.

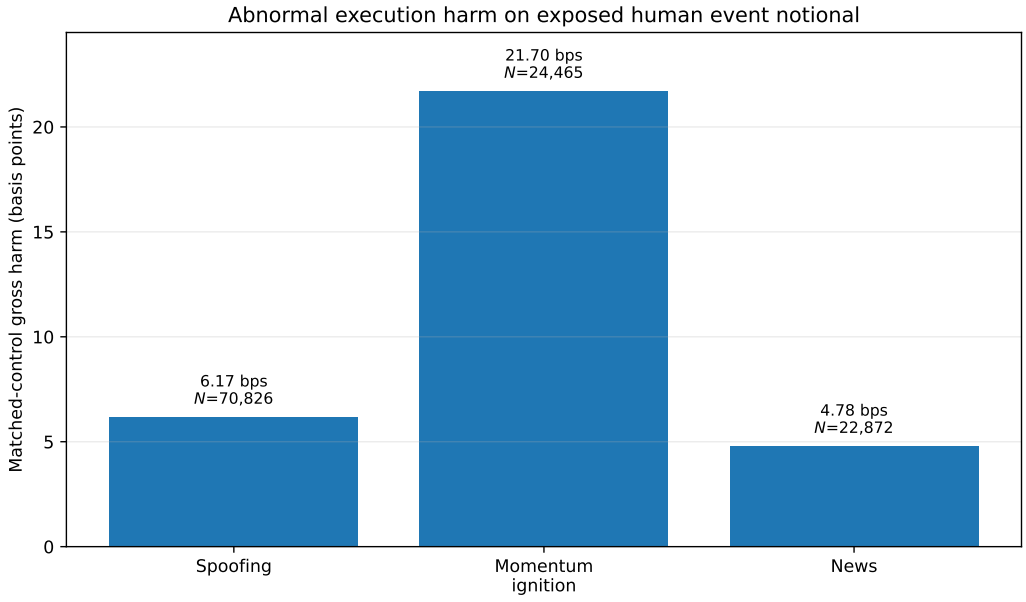


Figure IA.25. Aggregate realized execution harm in basis points. The bars report matched-control gross rates on event-family human exposure. Momentum ignition has the largest abnormal rate, while spoofing is broader in frequency.

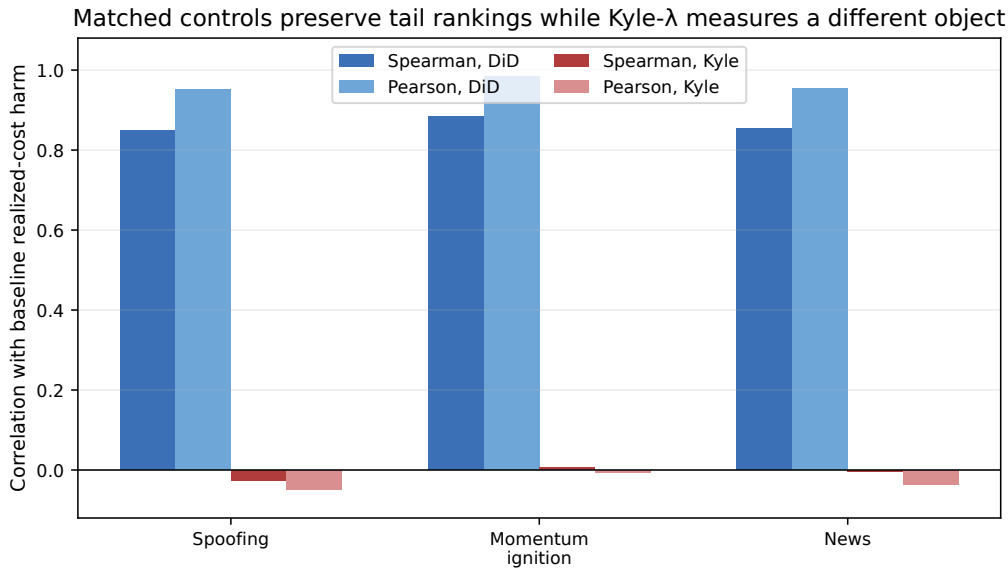


Figure IA.26. Estimator agreement. The figure reports correlations with the local baseline realized-cost estimator. Matched-control DiD preserves the ranking of harmful episodes, while Kyle-λ is nearly orthogonal because it measures a different economic object.



Figure IA.27. Execution and volatility-risk incidence in common units. Gross execution harm is normalized by event-family human exposure. Expected-shortfall harm is the mean event-level tail-loss rate. The rankings differ because the first object prices realized execution, while the second prices inventory risk.

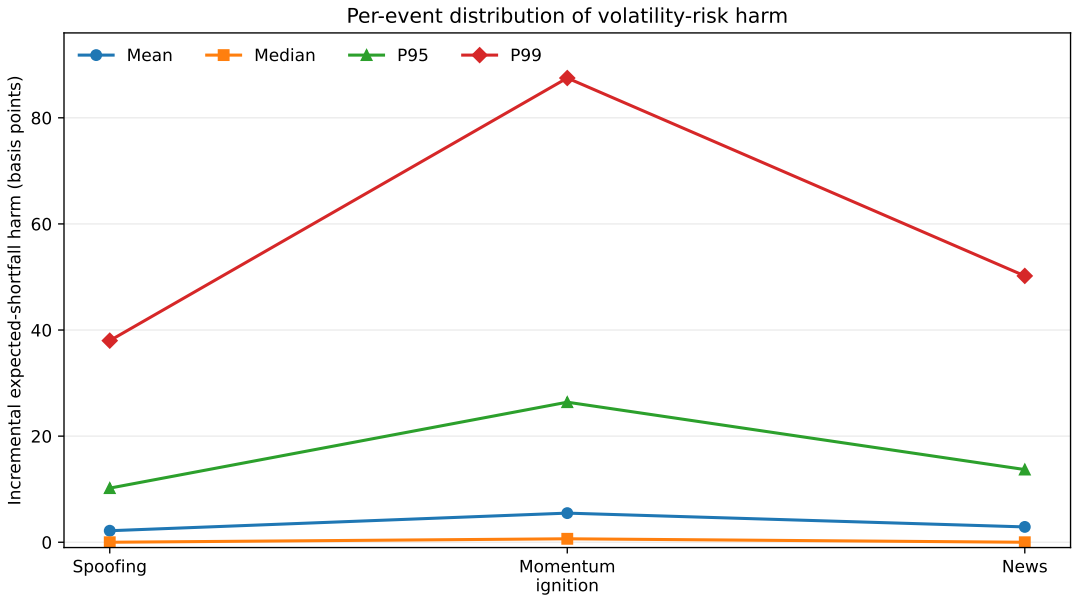


Figure IA.28. Distribution of expected-shortfall harm in basis points. The figure reports event-level means, medians, P95s, and P99s. The gap between the center and upper tail is largest for momentum ignition.

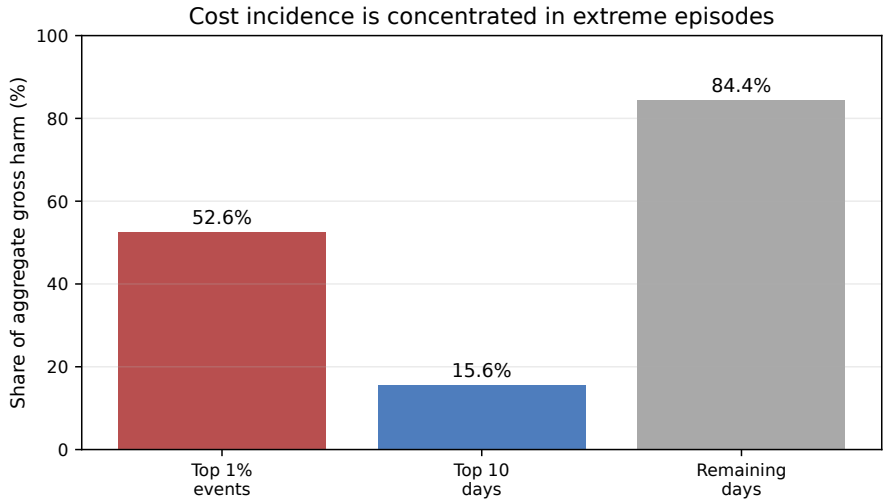


Figure IA.29. Concentration of investor harm. The figure reports the share of aggregate notional-weighted gross harm attributable to the top 1% of events and the top ten calendar days. Extreme episodes dominate incidence, so unconditional average market quality is incomplete.

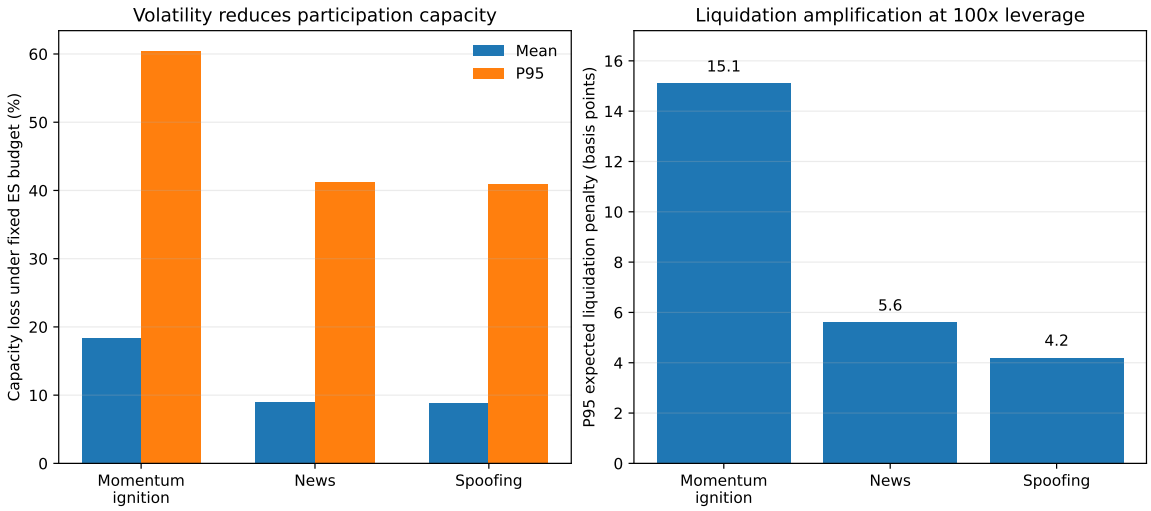


Figure IA.30. Participation capacity and liquidation amplification. The left panel reports capacity loss under a fixed expected-shortfall budget. The right panel reports the P95 liquidation-penalty rate at 100× leverage. Volatility shocks matter because they force smaller positions and can generate nonlinear losses under leverage.

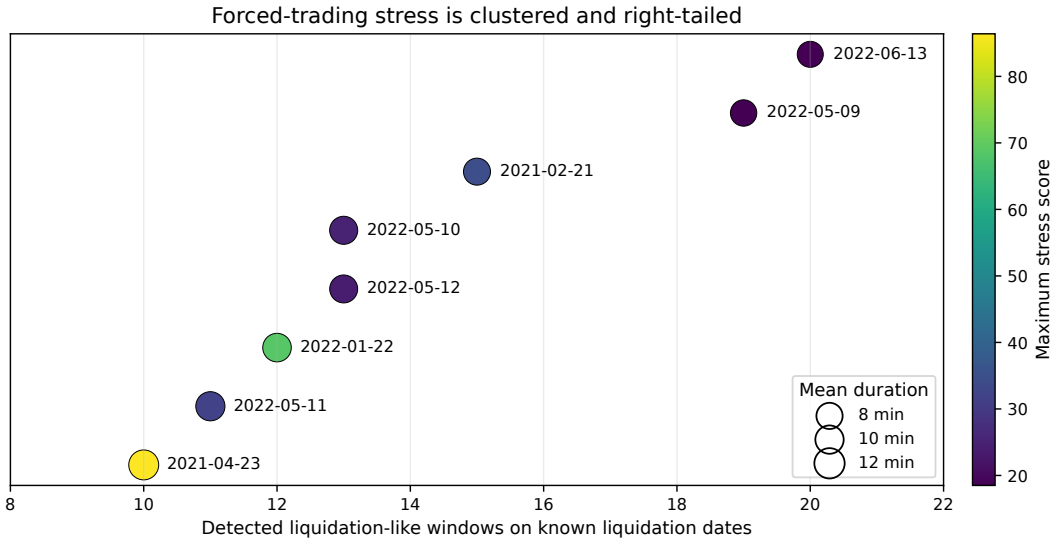


Figure IA.31. Liquidation-like stress windows on known liquidation dates. Points report the number of detected windows on major liquidation dates. Color is the maximum directional stress score and marker size is mean duration. Forced-trading stress is both clustered and right tailed.

IA.G.5. Market-Phase Classification and Regime Robustness

Emission model and persistence. For each candidate state count $K \in \{2, 3, 4\}$, the robustly standardized feature vector in Equation (46) is modeled as

$$Z_t \mid S_t = s \sim \mathcal{N}(\mu_s, \Sigma_s), \qquad \Sigma_s \text{ diagonal.} \tag{IA.57}$$

The diagonal restriction is appropriate for the 546-day sample because an unrestricted state-specific covariance matrix would add weakly identified parameters and could fit transient correlations rather than persistent phases. Dependence across the four features is retained through their common latent state. To prevent one-day reversals from producing economically meaningless phase switches, transition estimates use a sticky diagonal pseudocount $\kappa = 30$,

$$\widehat{A}_{ij} = \frac{\sum_t \xi_t(i,j) + \kappa \mathbf{1}\{i = j\}}{\sum_{k=1}^{K} \left[\sum_t \xi_t(i,k) + \kappa \mathbf{1}\{i = k\}\right]}, \tag{IA.58}$$

where $\xi_t(i,j)$ is the smoothed probability of transition $i \to j$. Each model is estimated from 15 initializations and the highest-likelihood solution is retained.

State count and economic labels. Table IA.VII reports in-sample BIC and expanding blocked validation. The validation exercise uses four nonoverlapping 50-day test blocks, with robust scaling and model parameters estimated only on the preceding training sample. Both criteria select four states.

Table IA.VIII characterizes the selected states. The labels are assigned after estimation from the ordering of 20-day return and 60-day drawdown and therefore do not affect likelihood maximization. Strong advance and steady advance are combined into the bull phase. Correction and stress are combined into the bear phase.

Figure IA.32 compares smoothed and filtered bull probabilities. The filtered probability uses observations only through date t. The 97.8% agreement in binary phase labels indicates that the main boundaries are not generated by future observations, although the smoothed state remains the appropriate object for the descriptive partition reported in the paper.

Cost incidence and conditional intensity. Figure IA.33 reports both the notional-weighted phase decomposition and the active-day intensity comparison. Panel (a) is an accounting partition in basis points. Panel (b) is the economically relevant severity comparison. Similar conditional intensity across phases shows that the larger bear-period share is not evidence of a multiplicative bear-market treatment effect.

Confidence intervals use a 14-day circular moving-block bootstrap with 5,000 replications, following Künsch (1989). Each block resamples the paired daily vector containing the phase label, event-family harm, and active-day indicator, preserving short-run clustering in market phases and execution costs. For each draw, the bear-to-bull active-day intensity ratio is recomputed.

Table IA.VII Selection of the Market-Phase Hidden Markov Model

States	Log likelihood	BIC	Blocked validation log score	Selected
2	−2,576.30	5,272.34	−5.407	No
3	−2,208.82	4,619.32	−5.082	No
4	−2,084.09	4,464.41	−4.879	Yes

Notes. Lower BIC and higher blocked validation log score are preferred. Each fit uses diagonal Gaussian emissions, persistent transitions, and multiple initializations.

Table IA.VIII Economic Characterization of the Four Latent States

Latent state	Mean 5-day return	Mean 20-day return	Mean drawdown	20-day volatility	Persistence
Strong advance	10.03%	19.82%	−2.75%	3.31%	0.921
Steady advance	−0.76%	10.59%	−6.45%	2.62%	0.941
Correction	−2.59%	−8.01%	−24.60%	3.24%	0.959
Stress	−5.52%	−23.13%	−53.81%	5.14%	0.976

Notes. Drawdown is measured from the trailing 60-day high. Persistence is the fitted probability of remaining in the same state on the next day.

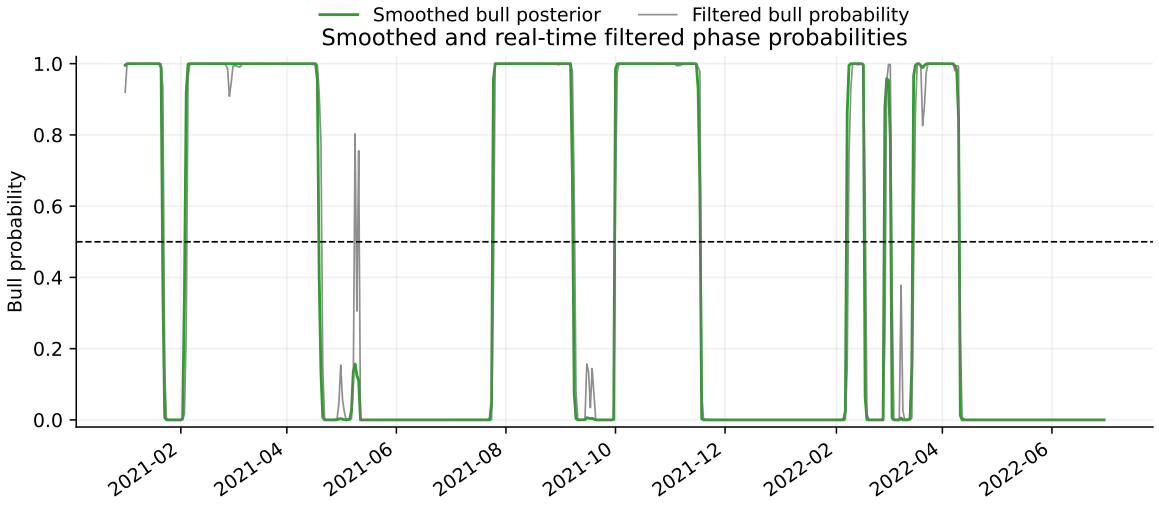


Figure IA.32. Smoothed and filtered bull probabilities. The smoothed posterior uses the full sample. The filtered posterior uses information only through date t. The corresponding bull-bear labels agree on 97.8% of sample days.

No return-jump news proxy enters the headline cost total or the market-phase partition. News uses the observed-news event definition in Section 5.

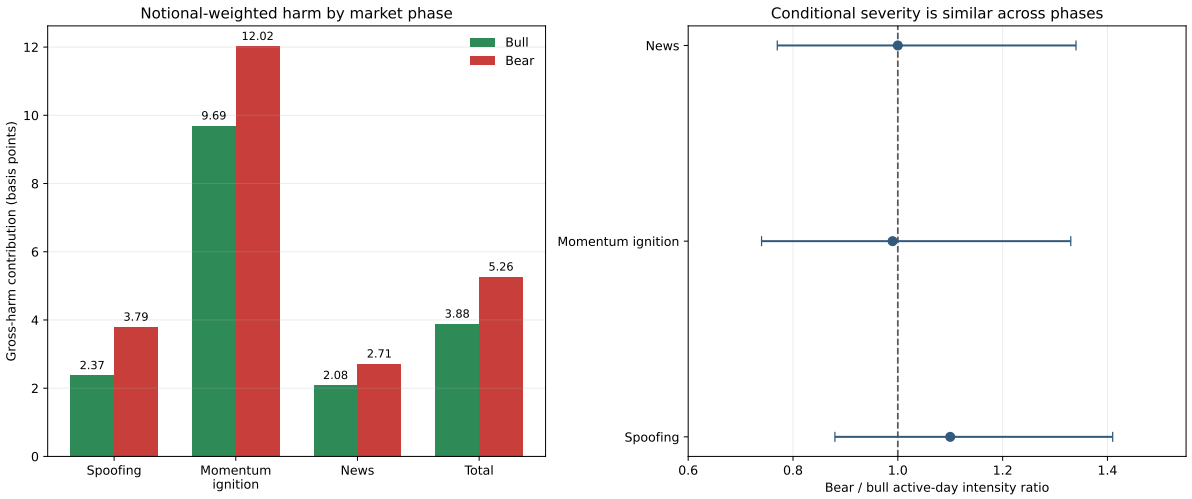


Figure IA.33. Investor-cost rates across market phases. The left panel partitions gross harm into bull and bear contributions. Event-family bars use the corresponding event-family exposure denominator, while the total bars use pooled human event-window notional. The right panel reports bear-to-bull active-day intensity ratios with 95% moving-block-bootstrap intervals.

Table IA.IX Momentum-Ignition Frequency by Market Phase

Phase	Detected events	Phase days	Events per day
Bull	7,641	228	33.51
Bear	16,824	318	52.91
Total	24,465	546	44.81

Notes. The bear-to-bull event-rate ratio is 1.58.

IA.H. Fee-Schedule Change Impact

IA.H.1. Setting, data, and empirical design

The main empirical designs identify adverse market states from order-book behavior, aggressive-flow acceleration, and public news. A natural concern is that the human–machine asymmetry could be particular to those event definitions. This appendix removes the event-detection machinery and studies an externally dated change in the direct price of trading. The exercise uses no spoofing score, momentum-ignition filter, news classifier, FPY factor, or event-cost mapping. It asks a narrower question. When Coinbase changed the relative price of supplying and demanding liquidity, did the human proxy and machine proxy adjust in the same way?

Coinbase announced on February 28, 2022 that a revised maker–taker fee schedule would take effect on March 1. The revision raised taker fees and lowered maker fees for several volume tiers, with the stated objective of increasing displayed depth and liquidity.[20] The economic effect of such

[20]Coinbase, "Coinbase Exchange fee updates—March 2022," February 28, 2022, `https://www.coinbase.com/blo`

a change is not mechanical. A higher taker fee raises the all-in cost of immediacy, while a lower maker fee improves the private return to passive supply. In equilibrium, however, the response also depends on fee pass-through, execution probability, tier assignment, order-routing technology, and the ability to revise orders when market conditions change. Theoretical and empirical work on maker–taker pricing therefore emphasizes both the total fee and its allocation between liquidity demanders and suppliers (Colliard and Foucault, 2012; Malinova and Park, 2015).

The cohort split provides an economically useful, but imperfect, exposure margin. Non-API activity is the paper's human or manual-trading proxy, while API activity is the machine proxy. Lower-volume and manually routed accounts are plausibly more sensitive to a change in the marginal cost of taking, whereas API participants can alter order type and timing at high frequency and are more likely to qualify for different volume tiers. We do not observe the fee tier of each account, so the cohort indicator is not a treatment assignment. The exercise is consequently an institutional ablation rather than a causal natural experiment. It tests whether the cohort decomposition produces economically coherent responses to a known change in transaction-cost incentives.

Data, outcomes, and empirical design. The raw panel contains one-minute BTC–USD observations from January 1, 2021 through June 30, 2022. The estimation sample is a symmetric 244-day window around the revision. The pre period runs from October 30, 2021 through February 28, 2022, and the post period runs from March 1 through June 30, 2022. For cohort $g \in \{H, M\}$ and horizon $h \in \{5, 60\}$ minutes, let

$$V_{g,t}^{T}(h) = B_{g,t}^{T}(h) + S_{g,t}^{T}(h) \tag{IA.59}$$

denote taker notional, where B^T and S^T are buy- and sell-initiated taker flow. Let $NF_{g,t}(h) = B_{g,t}^{T}(h) - S_{g,t}^{T}(h)$ denote signed taker flow, and let $L_{g,t}^{1\%}(h)$ denote cohort-attributed displayed liquidity within one percent of the midquote. We normalize the latter to a per-minute quantity,

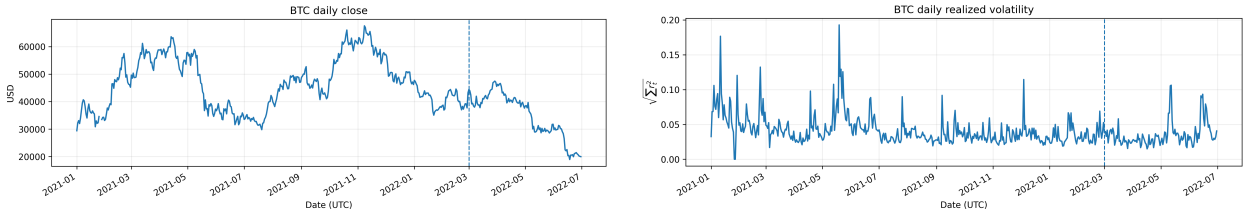
$$\overline{L}_{g,t}^{1\%}(h) = \frac{L_{g,t}^{1\%}(h)}{h}, \tag{IA.60}$$

and define the passive-intensity proxy

$$\mathcal{M}_{g,t}(h) = \frac{\overline{L}_{g,t}^{1\%}(h)}{V_{g,t}^{T}(h)}. \tag{IA.61}$$

This proxy measures displayed passive presence per dollar of aggressive activity. It is deliberately not labeled a maker-order rate because observed depth reflects additions, cancellations, executions, and order duration. The distinction matters for interpretation. A rise in $\mathcal{M}$ would indicate a greater equilibrium presence of passive liquidity relative to taking, but it would not by itself identify an

g/march-2022-coinbase-exchange-fee-updates.



(a) BTC daily close **(b)** Daily realized volatility

Figure IA.34. Market conditions around the Coinbase fee revision. The dashed line marks March 1, 2022. Realized volatility is the square root of the sum of squared one-minute log returns within each UTC day. The post window coincides with a materially different price and volatility regime, which rules out an unconditioned before–after interpretation.

increase in gross maker submissions.

The principal outcomes are within-venue cohort differences,

$$Y_{t,h}^{V} = \log V_{H,t}^{T}(h) - \log V_{M,t}^{T}(h), \tag{IA.62}$$

$$Y_{t,h}^{L} = \log L_{H,t}^{1\%}(h) - \log L_{M,t}^{1\%}(h), \tag{IA.63}$$

$$Y_{t,h}^{\mathcal{M}} = \log \mathcal{M}_{H,t}(h) - \log \mathcal{M}_{M,t}(h). \tag{IA.64}$$

The log ratios absorb multiplicative shocks that affect both cohorts proportionally, including broad changes in market-wide volume. They do not eliminate cohort-specific trends, which is why the weekly event study below is as important as the average pre/post coefficient.

The baseline regression is

$$Y_{t,h} = \alpha_h + \beta_h \,\mathrm{Post}_t + \Gamma_h' R_{t,h} + \delta_{\mathrm{dow}(t)} + \delta_{\mathrm{hour}(t)} + \mathbf{1}\{h = 5\}\delta_{\mathrm{minute}(t)} + \varepsilon_{t,h}, \tag{IA.65}$$

where Post_t equals one on and after March 1, 2022. The regime vector $R_{t,h}$ contains the log BTC price at the block endpoint, the cumulative log return $\mathrm{Ret}_t(h) = \log P_t - \log P_{t-h+1}$, and realized volatility $\mathrm{RV}_t(h) = \{\sum_{j=t-h+2}^{t} r_j^2\}^{1/2}$, where $r_j = \log P_j - \log P_{j-1}$. Standard errors are clustered by UTC date. The outcome variables are rolling sums, so using every minute would repeatedly reuse the same underlying messages. We therefore retain non-overlapping block endpoints, with minutes congruent to four modulo five for the five-minute specification and minute 59 for the hourly specification. This sampling rule removes the most direct source of mechanical pseudo-replication while preserving both a short execution horizon and a slower participation horizon.

Figure IA.34 shows why the regime controls are necessary. The post period overlaps a large BTC drawdown and several volatility episodes. These controls condition on observable market state, but they cannot reconstruct the counterfactual path of Coinbase participation in the absence of the fee revision. The estimates should therefore be read as conditional post-period differences, not as an identified treatment effect.

IA.H.2. Average post-period differences

Table IA.X reports the relative-outcome estimates. The main result is a joint contraction in human taking and human-attributed liquidity relative to the machine cohort. At the hourly horizon, the human-to-machine taker-volume ratio falls by $100\{\exp(-0.377) - 1\} = -31.4\%$, and the corresponding liquidity ratio falls by 33.2%. The five-minute estimates are nearly identical at -33.2% and -32.7%. The agreement across aggregation horizons is important because the five-minute specification is closer to order-choice adjustment, while the hourly specification averages over short-lived bursts and quote revisions.

The passive-intensity result separates participation exit from order-type substitution. The post coefficient for $\log(\mathcal{M}_H/\mathcal{M}_M)$ is economically small and statistically indistinguishable from zero at both horizons. The fee revision therefore does not coincide with a detectable rotation of the human proxy toward a greater amount of displayed liquidity per dollar of taking. Instead, the relative amount of both human taking and human-attributed depth contracts by approximately the same proportion. The directional decomposition is secondary but consistent with the same pattern. At the hourly horizon, the human-to-machine sell-taking ratio falls by about 39.9%, while the buy-taking ratio falls by about 29.6%.

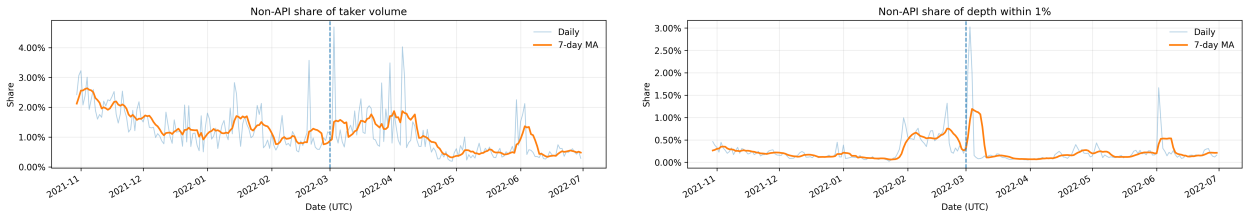
Table IA.XI shows that the relative results are driven primarily by the human proxy. At the hourly horizon, human taker volume falls by 36.1% and human-attributed liquidity falls by 36.7%. The corresponding machine changes are -6.8% for taker volume and -5.3% for liquidity, with the liquidity estimate statistically indistinguishable from zero. The within-cohort passive-intensity coefficients are also close to zero. Taken together, these estimates describe a decline in human participation on both sides of the market rather than a simple relabeling of aggressive activity as passive activity.

Figure IA.35 reports daily participation shares. The human share of taker volume trends downward before March 1 and remains low through much of the post window. The liquidity-share series is noisier because a small number of deep states can move a daily average, but its smoothed path also spends most of the later post period below the early-pre-period level. These plots are useful precisely because they are less favorable to a sharp causal interpretation than the average coefficient. The fee revision occurs during an ongoing change in participant composition rather than at the beginning of a stable pre-period plateau.

IA.H.3. Dynamic specification and identification limits

To make the identifying limitation explicit, we replace the single post indicator with week effects. Let W_t be the number of seven-day intervals between block t and March 1, with week zero beginning on the revision date. The dynamic specification is

$$Y_{t,60} = \alpha + \sum_{k=-18,\, k\neq -1}^{18} \beta_k \mathbf{1}\{W_t = k\} + \Gamma' R_{t,60} + \delta_{\mathrm{dow}(t)} + \delta_{\mathrm{hour}(t)} + \varepsilon_t, \tag{IA.66}$$



(a) Human share of taker volume

(b) Human share of liquidity within one percent

Figure IA.35. Daily cohort participation around the fee revision. The thin line is the daily statistic and the thick line is a seven-day moving average. Human denotes non-API activity. The dashed line marks March 1, 2022. The visible pre-period drift motivates the event-study and limits causal interpretation of the average post coefficient.

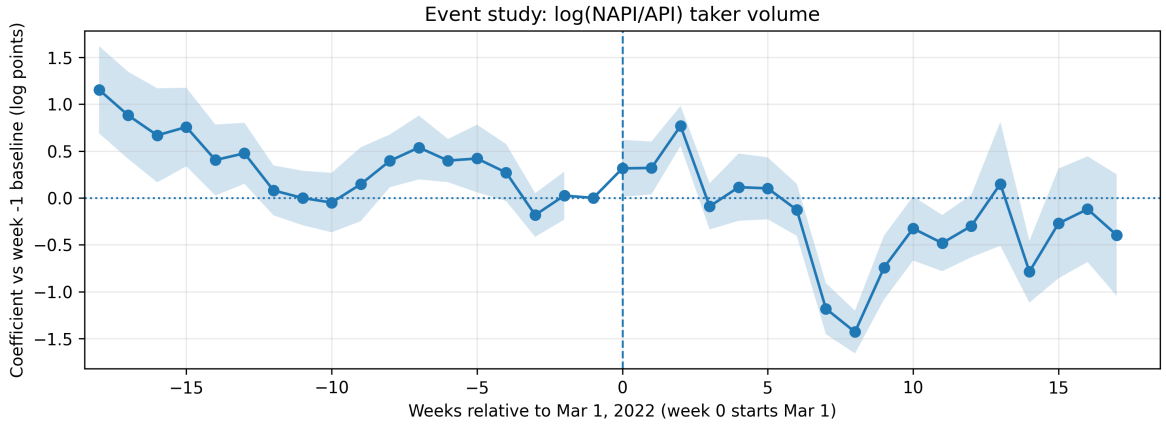
where week -1 is omitted. Figure IA.36 reports $\widehat{\beta}_k$ with 95 percent intervals based on date-clustered standard errors.

The dynamic paths are informative in two respects. First, the human-to-machine taker-volume ratio is already declining before the revision and becomes most negative several weeks afterward rather than at week zero. A January 1, 2022 cutoff placebo produces a coefficient of -0.392 for the same outcome and remains statistically significant, which confirms that the decline was not created at the March boundary. Second, the liquidity ratio falls after the cutoff but later recovers, while the passive-intensity proxy has no stable post-revision displacement. The data therefore reject a simple discontinuous-and-permanent treatment narrative. What survives is a more modest result. Conditional on observable BTC regime variables, the March–June window contains substantially less human activity relative to machine activity, and the contraction occurs in both taking and displayed liquidity rather than through a clean switch from one order type to the other.

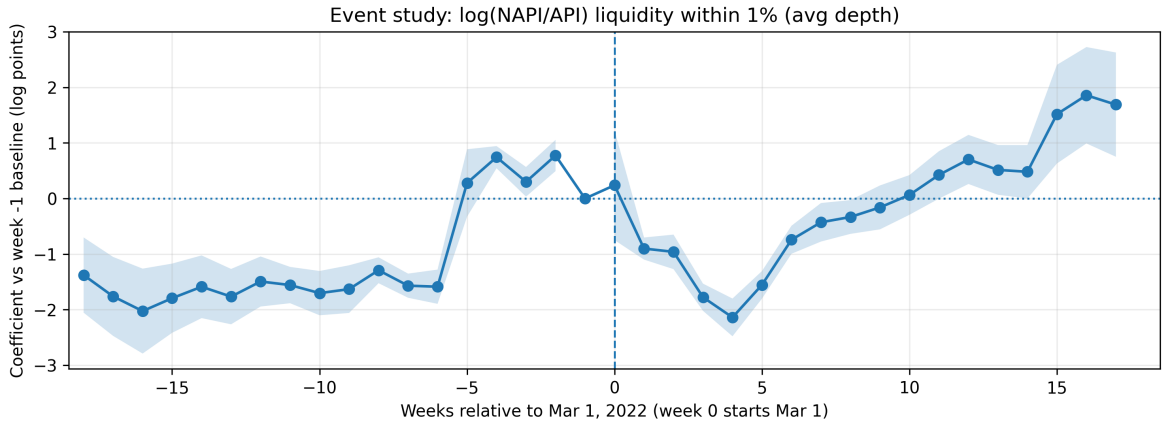
IA.H.4. Implications for the main analysis

This ablation supports two conclusions and does not support a third. It supports the view that the API label captures an economically meaningful technological margin, because the two cohorts respond differently when the venue changes the direct price of liquidity demand and supply. It also supports a participation channel. The human proxy contracts in both taker activity and displayed liquidity, with no detectable increase in displayed liquidity per dollar of taking. The evidence does not establish that the fee schedule caused the entire contraction. Fee tiers are unobserved, the post period coincides with a market-wide drawdown, and the weekly paths show differential pre-trends.

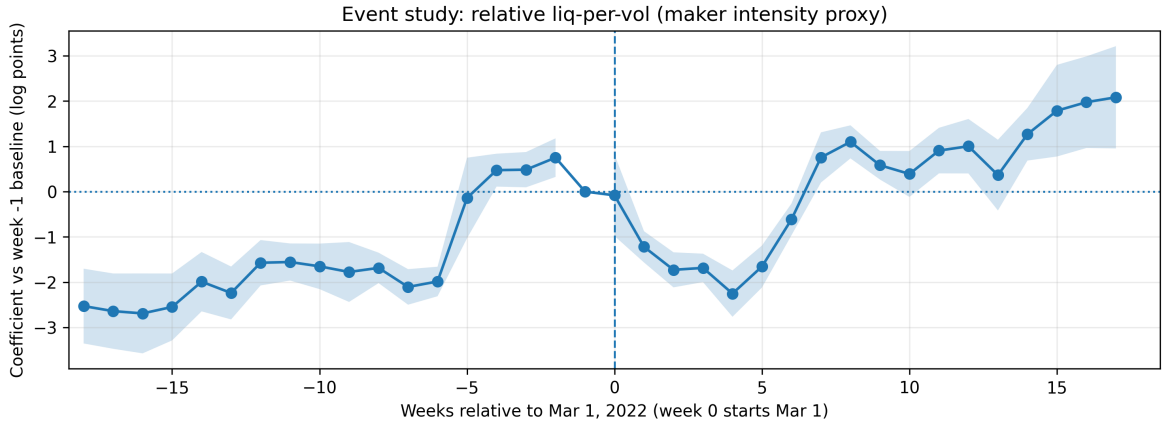
The result is therefore an external-validity check on the paper's mechanism rather than an additional component of the investor-cost total. The manipulation and news sections show that fast traders and slow traders bear different execution conditions within short adverse states. The fee-schedule ablation shows that the same cohort split also responds differently to a transparent change in direct trading costs. Together, the findings indicate that technological asymmetry is not confined to a particular detector or event label. It also appears in the way participation adjusts when the venue changes the price of immediacy.



(a) Human-to-machine taker-volume ratio



(b) Human-to-machine liquidity ratio



(c) Relative passive-intensity proxy

Figure IA.36. Weekly dynamics around the Coinbase fee revision. Points are coefficients from Equation (IA.66), with week −1 as the omitted category. Shaded regions are 95 percent confidence intervals based on standard errors clustered by UTC date. The vertical line marks the beginning of week zero. The non-flat pre-period paths and the gradual post response show why the exercise is used as an institutional ablation rather than a causal regression discontinuity.

Table IA.X Fee-schedule ablation: relative cohort outcomes

	5-minute blocks	**60-minute blocks**
Panel A. Human-to-machine relative outcomes		
$\log\left(V_H^T/V_M^T\right)$	-0.404^{***}	-0.377^{***}
	(-3.38)	(-4.16)
$\log\left(L_H^{1\%}/L_M^{1\%}\right)$	-0.396^{***}	-0.404^{***}
	(-2.71)	(-2.81)
$\log(\mathcal{M}_H/\mathcal{M}_M)$	0.009	-0.027
	(0.05)	(-0.15)
$\log\left(B_H^T/B_M^T\right)$	-0.118	-0.351^{***}
	(-0.73)	(-3.07)
$\log\left(S_H^T/S_M^T\right)$	-0.511^{***}	-0.509^{***}
	(-4.91)	(-5.41)
$NF_H/V_H^T - NF_M/V_M^T$	0.054^{*}	0.045
	(1.68)	(1.35)
Panel B. Signed taker flow, millions of dollars per block		
NF_H	0.002	0.024
	(0.65)	(0.70)
NF_M	0.061^{**}	0.667^{**}
	(2.46)	(2.24)
$NF_H - NF_M$	-0.059^{**}	-0.643^{**}
	(-2.34)	(-2.12)
Observations (maximum)	70,240	5,849
BTC regime controls	Yes	Yes
Day-of-week and hour fixed effects	Yes	Yes
Minute-of-hour fixed effects	Yes	No

Notes. The matched sample runs from October 30, 2021 through June 30, 2022. The post period begins March 1, 2022. Human denotes non-API activity and machine denotes API activity. The dependent variables in Panel A are log ratios except for the signed-flow imbalance difference. Every regression includes log price, block return, realized volatility, day-of-week fixed effects, and hour-of-day fixed effects. Five-minute regressions additionally include minute-of-hour fixed effects. Standard errors are clustered by UTC date, and t-statistics appear in parentheses. Observation counts vary slightly across log outcomes when a cohort has zero activity in a block, so the table reports the maximum count. *, **, and *** denote significance at the 10, 5, and 1 percent levels.

Table IA.XI Fee-schedule ablation: within-cohort decomposition

	5-minute blocks	**60-minute blocks**
$\log V_H^T$	-0.511^{***}	-0.448^{***}
	(-3.92)	(-4.65)
$\log V_M^T$	-0.107^{**}	-0.071^{*}
	(-2.44)	(-1.78)
$\log \overline{L}_H^{1\%}$	-0.469^{***}	-0.458^{***}
	(-3.80)	(-3.76)
$\log \overline{L}_M^{1\%}$	-0.074	-0.054
	(-1.63)	(-1.23)
$\log \mathcal{M}_H$	0.044	-0.011
	(0.23)	(-0.06)
$\log \mathcal{M}_M$	0.034	0.017
	(0.71)	(0.37)
Observations (maximum)	70,033	5,849
BTC regime controls	Yes	Yes
Calendar fixed effects	Yes	Yes

Notes. This table estimates Equation (IA.65) separately for each cohort. Liquidity is divided by the horizon length before taking logs so that it is expressed as average displayed liquidity per minute. The remaining controls, fixed effects, sampling rules, and clustered standard errors are identical to Table IA.X.