Queue Priority as Adverse-Selection Exposure

Aryan Ayyar

Manipal Academy of Higher Education

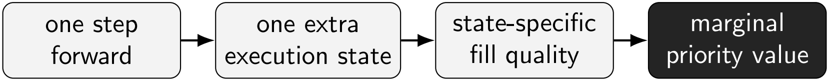
29 August 2026

We Ask a Very Simple Question

Does Moving Toward the Front Bench Imply You Are the Topper?

- Earlier rank raises the probability of execution.
- But it changes *which* executions the order receives.
- The extra fill may arrive exactly when the quote is stale.
- Queue depth ahead can therefore provide delay or protection.

The argument in one line



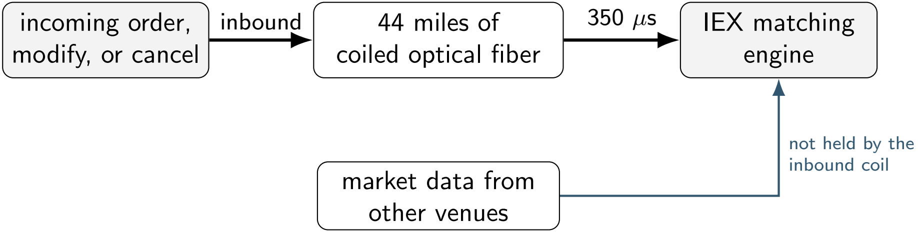
- If the marginal fill is benign, moving forward is valuable.
- If the marginal fill is toxic, standing farther back is a shield.
- Mixed flow can create an interior best position.

The IEX Exchange

- U.S. national securities exchange since 2016.
- Founded in 2012 to build a more transparent market that better protected investors from structural speed advantages.
- Continuous electronic trading in U.S. equities.
- Displayed and non-displayed liquidity.
- Displayed same-price orders expose a FIFO queue.

Source: IEX, "About Us: IEX History"; SEC, Release approving IEX as a national exchange (17 June 2016).

IEX's 350 μs speed bump



- The buffer lets IEX process external quote changes before an incoming order executes.
- It targets stale-price latency arbitrage; it is **not** batch clearing.
- FIFO priority remains meaningful inside the IEX book.

Source: IEX, "Technology" and "Breaking Down M-Peg and D-Peg."

Why IEX Is Useful Here

- DPLS records each order's lifecycle: add, modify, execute, cancel, or delete.
- Deterministic replay recovers exact displayed rank and shares ahead.
- Executions and local quote states make marginal queue exposure observable.
- The speed bump changes the race to the engine while FIFO still allocates displayed priority.

Protection from latency arbitrage does not make every unit of queue priority equally valuable.

Source: IEX DPLS specification; IEX, "Discretionary Limit" and IEX Signal documentation.

From Order Lifecycles to Sparse Event Deltas

For each partition

$$g = (\text{symbol}, \text{clear epoch}, \text{side}, \text{price}),$$

compile every order lifecycle into signed size changes:

$$\delta_{g,t} = \text{adds}_{g,t} - \text{executions}_{g,t} - \text{cancels}_{g,t} + \text{modifications}_{g,t}.$$

- Priority spells replace repeated mutation of a global FIFO book.
- Clear events start a new independent epoch.
- Sorting occurs once within each partition.

The Prefix Operator Vectorizes Replay

Let $\boldsymbol{\delta}_g = (\delta_{g,1}, \dots, \delta_{g,T_g})^\top$ and define

$$(L_g)_{ts} = \mathbf{1}\{s \le t\}.$$

Then the complete displayed-size path is recovered at once:

$$\mathbf{x}_g = L_g \boldsymbol{\delta}_g, \qquad x_{g,t} = \sum_{s \le t} \delta_{g,s}.$$

- Across partitions, $L = \operatorname{diag}(L_1, \dots, L_G)$ is block diagonal.
- A segmented cumulative sum applies L without materializing the dense matrix.
- The operation is algebraically identical to sequential replay within every clear epoch.

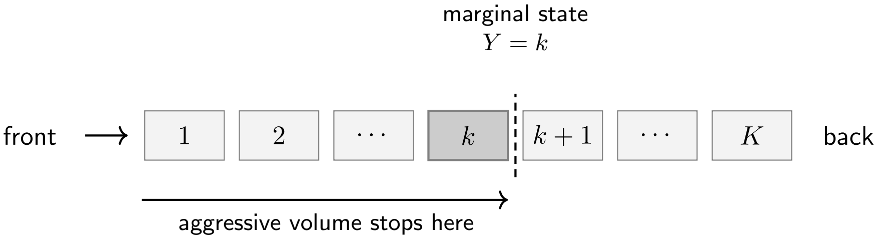
From State Paths to Queue Coordinates

For active order i with priority key π_i:

$$\text{shares ahead}_{i,t} = \sum_{j:\,\pi_j<\pi_i} s_{j,t}, \qquad \text{rank}_{i,t} = 1 + \sum_{j:\,\pi_j<\pi_i} \mathbf{1}\{s_{j,t} > 0\}.$$

- Segmented scans recover displayed size and depth.
- Priority-prefix or Fenwick queries recover shares ahead and rank.
- Only execution-touched levels need to emit rank-exposure rows.
- Certification and production-prefix gates matched canonical replay exactly.

One step of priority is one marginal execution state



- If $Y < k$, neither adjacent rank executes.
- If $Y \geq k+1$, both execute.
- Only $Y = k$ distinguishes rank k from rank $k+1$.

The marginal-priority identity

$$V_k(h) = \mathbb{E}[m(h)\mathbf{1}\{Y \geq k\}] \tag{1}$$

$$\begin{aligned} \Delta_k(h) &= V_k(h) - V_{k+1}(h), && (2) \\ &= \mathbb{E}[m(h)\mathbf{1}\{Y = k\}]\,, && (3) \\ &= q_k\, M_k(h). && (4) \end{aligned}$$

- $m(h)$ is the provider-signed post-fill markout.
- $q_k = \mathbb{P}(Y = k)$ is marginal stopping probability.
- $M_k(h) = \mathbb{E}[m(h) \mid Y = k]$ is marginal fill quality.
- The sign of priority is the sign of the marginal fill's expected surplus.

The three measured objects

Three components of the measurement

1. **Probability, $q(K)$:** How often does aggressive flow terminate in rank block K?
2. **Fill quality, $M(K, h)$:** Conditional on that stop, is the passive fill favorable or adverse?
3. **Priority value, $\Delta V(K, h) = q(K)M(K, h)$:** What is the direct expected surplus of owning that marginal exposure?

Probability and fill quality must be measured together.

Why full clears are censored

interior stop
rank k fills; $k+1$ does not
↓
identifies adjacent priority value

full level clear
both ranks fill
↓
right-censored for the adjacent-rank margin

- A very toxic sweep can have a bad absolute markout.
- Yet it contributes zero to the adjacent margin if both ranks execute.

The April–May Study Window

Element	Design
Continuous study window	1 April–29 May 2026; 41 trading sessions
Presentation sample	All 41 sessions; no date is deleted
Quote-quality gate	Retain future IEX BBO when $2(a-b)/(a+b) \leq 1$, equivalently $a \leq 3b$
Queue coordinates	Front, 1–99, 100–499, 500–999, 1,000–4,999, 5,000+ shares ahead
Horizons	100 ms, 1 s, 10 s, 60 s
Uncertainty	Date-cluster bootstrap; simultaneous family across 48 broad cells

- One event-blind rule removes 2,691 of 117,066,077 valid 100-ms rows.
- The rule retains 99.9977% of support and is applied identically to every symbol and date.

Descriptive Statistics: Daily Coverage

Date	Risk rows	Stops	Date	Risk rows	Stops
01 Apr	2,793,657	674,843	01 May	2,631,767	581,535
02 Apr	2,397,587	595,291	04 May	2,856,210	649,315
06 Apr	1,982,817	501,739	05 May	2,848,915	633,864
07 Apr	2,367,584	580,644	06 May	3,234,744	712,279
08 Apr	2,737,504	652,949	07 May	3,289,445	732,660
09 Apr	2,533,148	603,928	08 May	3,018,574	671,446
10 Apr	2,549,584	608,165	11 May	3,219,396	714,807
13 Apr	2,521,693	597,571	12 May	3,240,749	735,614
14 Apr	2,917,663	687,884	13 May	3,224,583	687,999
15 Apr	2,959,343	687,084	14 May	2,991,556	653,332
16 Apr	3,003,129	683,906	15 May	2,912,682	644,269
17 Apr	3,152,349	717,644	18 May	3,056,722	673,537
20 Apr	2,637,037	621,701	19 May	3,194,630	691,435
21 Apr	2,778,773	649,951	20 May	3,109,335	675,931
22 Apr	2,736,045	627,986	21 May	3,006,244	667,160
23 Apr	2,743,800	619,561	22 May	2,736,882	618,481
24 Apr	2,587,205	602,776	26 May	3,088,243	680,070
27 Apr	2,733,828	608,248	27 May	2,915,215	640,721
28 Apr	2,630,467	596,170	28 May	3,226,212	698,673
29 Apr	2,749,597	604,730	29 May	2,931,831	667,879
30 Apr	2,816,641	629,282			
April total	**56,329,451**	**13,152,053**	**May total**	**60,733,935**	**13,431,007**

Corrected 41-session sample at 100 ms. Risk rows and terminal stops are shown after the uniform quote-quality gate.

Descriptive Statistics: Bid-Side Queue Support

Shares ahead	Risk rows	Stops	q (%)	M	Δ
Front	27,008,285	11,098,567	41.09	2.199	0.9037
1–99	3,387,614	688,982	20.34	0.533	0.1084
100–499	11,981,866	990,331	8.27	-0.365	-0.0302
500–999	4,513,966	170,342	3.77	0.190	0.0072
1,000–4,999	7,494,916	157,411	2.10	0.263	0.0055
5,000+	4,385,852	40,508	0.92	0.357	0.0033

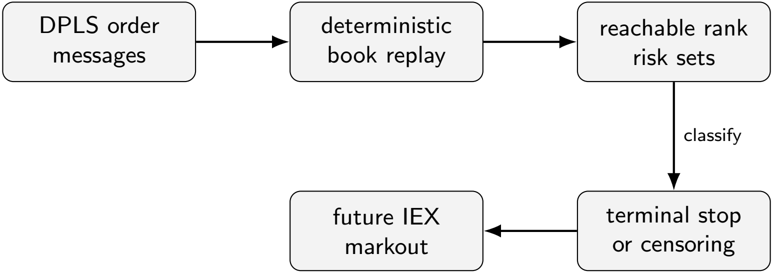
Corrected 41-session sample at 100 ms. M and Δ are provider-signed ticks.

Descriptive Statistics: Ask-Side Queue Support

Shares ahead	Risk rows	Stops	q (%)	M	Δ
Front	26,691,708	11,434,798	42.84	1.177	0.5042
1–99	3,114,704	645,173	20.71	-1.017	-0.2106
100–499	12,058,035	985,618	8.17	-0.424	-0.0347
500–999	4,553,846	168,653	3.70	0.193	0.0072
1,000–4,999	7,554,564	162,481	2.15	0.256	0.0055
5,000+	4,318,030	40,196	0.93	0.295	0.0027

Corrected 41-session sample at 100 ms. M and Δ are provider-signed ticks.

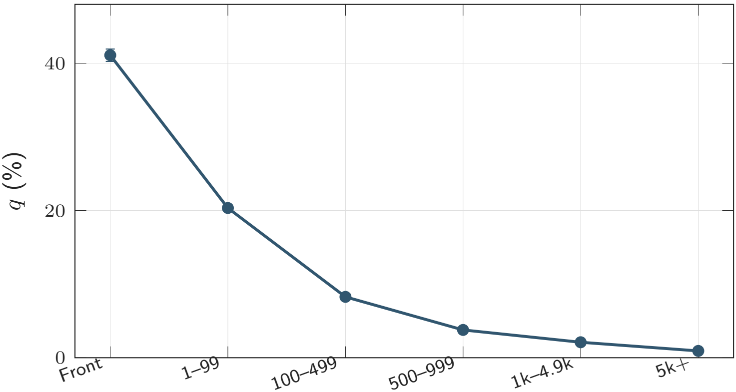
From DPLS messages to the economic object



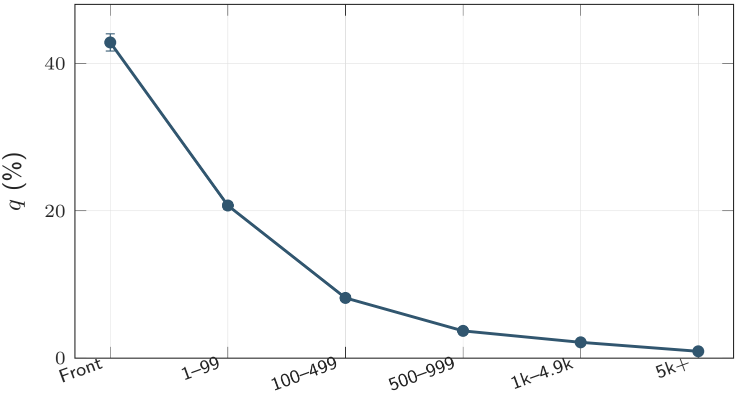
How the Measurement Is Constructed

- Adds, modifies, executions, deletes, and clear-book events reconstruct displayed FIFO.
- Aggressive episodes identify the volume path through the queue.
- Markouts are signed from the passive provider's perspective.
- The 350 μs delay is already embedded before the recorded system event; it is not added again.

Result 1: Marginal Stopping Probability — Bid



Result 1: Marginal Stopping Probability — Ask

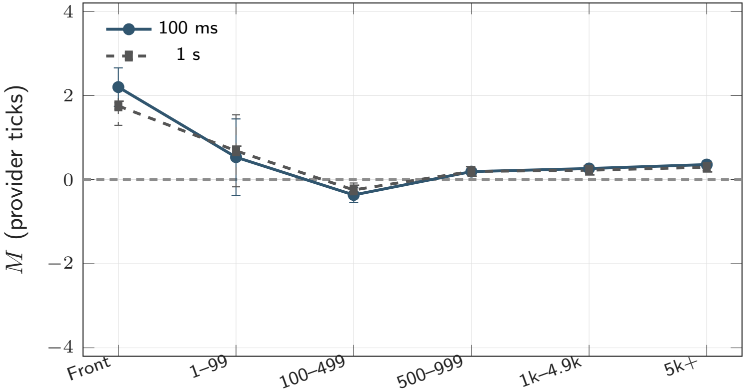


How to Read Marginal Fill Quality

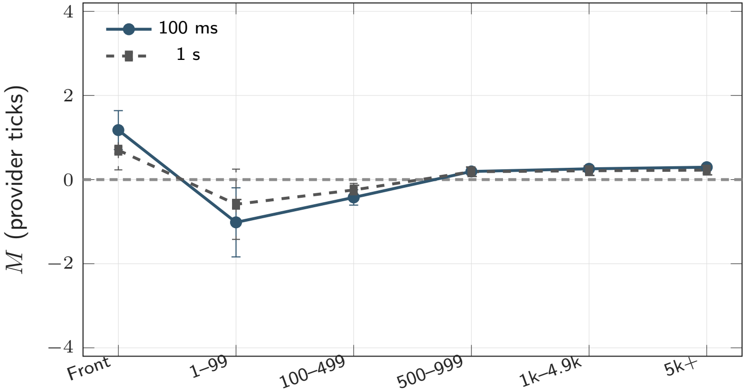
$$M(K, h) = \mathbb{E}[m(h) \mid Y \in K]$$

- A label such as 1,000–4,999 is a **shares-ahead bin**.
- $M(K, h)$ averages provider-signed markouts conditional on a marginal stop in bin K.
- The plotted value is a conditional mean, not a sum or cumulative markout.
- Positive values favor the passive provider; negative values indicate adverse selection.

Result 2: Marginal Fill Quality — Bid



Result 2: Marginal Fill Quality — Ask



From Fill Quality to Priority Value

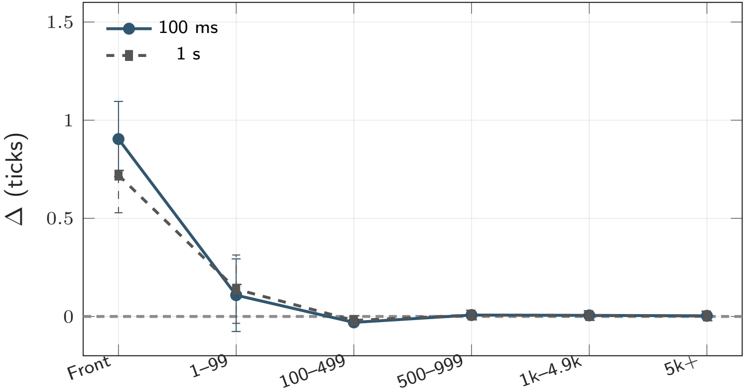
$$\Delta(K,h) = q(K)\,M(K,h)$$

- $M(K,h)$ asks how good the fill is *when* the marginal state occurs.
- $q(K)$ measures how often aggressive flow stops in that state.
- $\Delta(K,h)$ is the direct expected contribution of owning that marginal queue exposure.

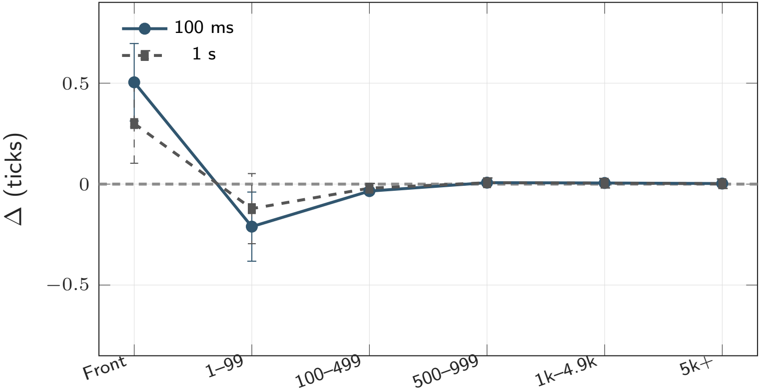
Bid, 100–499 shares ahead, 100 ms

$$q = 8.27\%, \qquad M = -0.365 \text{ ticks}, \qquad \Delta = -0.0302 \text{ ticks.}$$

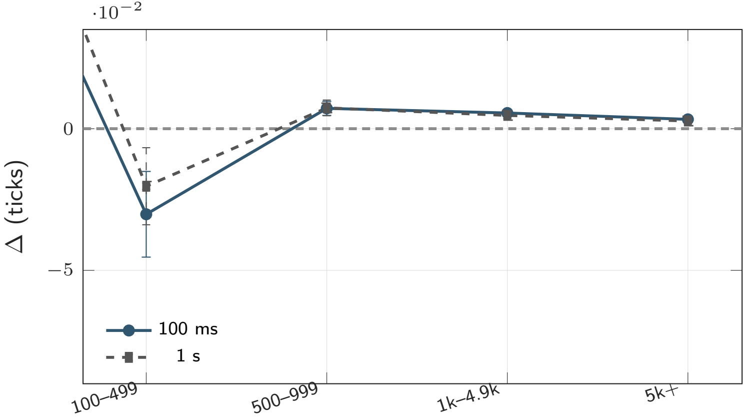
Result 3: Marginal FIFO Priority Value — Bid



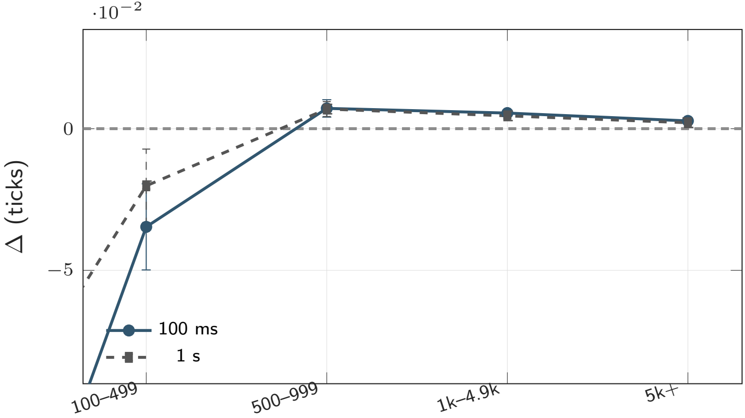
Result 3: Marginal FIFO Priority Value — Ask



Result 3: Non-Front Queue Value — Bid



Result 3: Non-Front Queue Value — Ask



Load-Bearing Point Estimates

Bid

	Front	100–499	1k–4.9k	5k+
q (%)	41.09	8.27	2.10	0.92
M	2.199	-0.365	0.263	0.357
Δ	0.9037	-0.0302	0.0055	0.0033

Ask

	Front	100–499	1k–4.9k	5k+
q (%)	42.84	8.17	2.15	0.93
M	1.177	-0.424	0.256	0.295
Δ	0.5042	-0.0347	0.0055	0.0027

Values are 100-ms provider ticks from the corrected 41-session sample.

Load-Bearing Date-Cluster Uncertainty

Bid

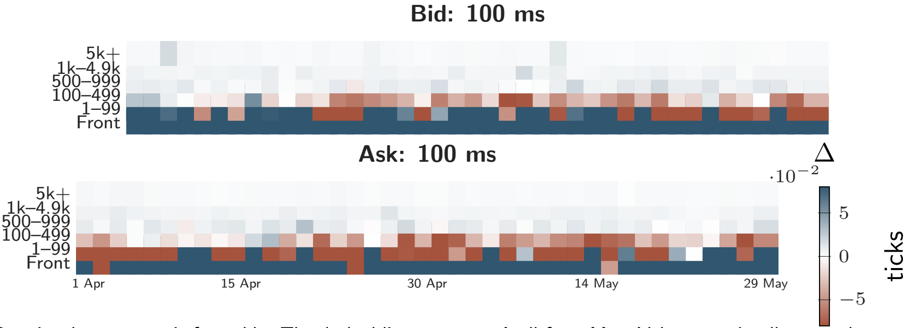
Queue bin	Simultaneous 95% band	Sign
Front	[0.7119, 1.0955]	strong +
100–499	[-0.0453, -0.0151]	−
1k–4.9k	[0.0044, 0.0067]	strong +
5k+	[0.0023, 0.0043]	strong +

Ask

Queue bin	Simultaneous 95% band	Sign
Front	[0.3121, 0.6963]	strong +
100–499	[-0.0498, -0.0195]	−
1k–4.9k	[0.0048, 0.0063]	strong +
5k+	[0.0025, 0.0030]	strong +

9,999 trading-date cluster bootstrap replications; simultaneous bands across all 48 side-by-horizon-by-bin cells. These quantify day-to-day stability, not BOLO causality.

Persistence across 41 trading sessions



Rust is adverse; navy is favorable. The dashed line separates April from May. Values are visually capped at ± 0.08 ticks so the interior cells remain visible.

Quote-Quality Gate: Screen the Far Quote, Not the Day

Exact raw-feed event

- 15 April, 11:36:43 ET: a HUBC ask was modified from $5 to $3,000.
- It became the IEX best ask only when lower asks disappeared: 311.5 seconds in total.
- Future IEX midpoints then jumped near $1,500. This is a real DPLS order, not a parser error.

Uniform screen, not a deleted day

- Keep all 41 sessions.
- Require

$$\frac{2(a-b)}{a+b} \leq 1 \quad \Longleftrightarrow \quad a \leq 3b.$$

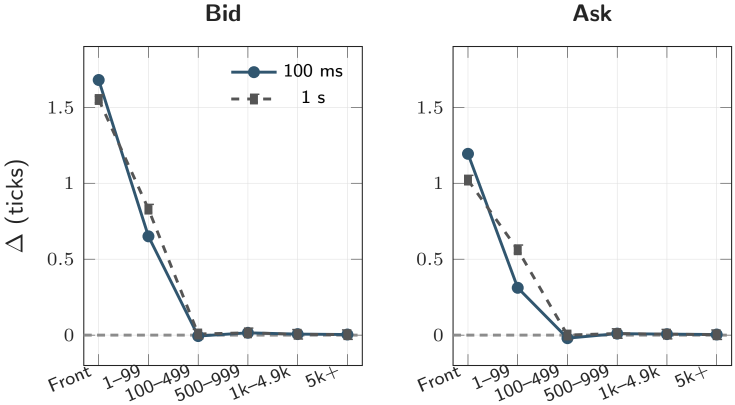
- Remove 2,691 of 117,066,077 valid 100-ms rows; retain 99.9977%.

Corrected 41-session sample

117,063,386 valid 100-ms risk rows and **26,583,060** valid terminal stops remain.

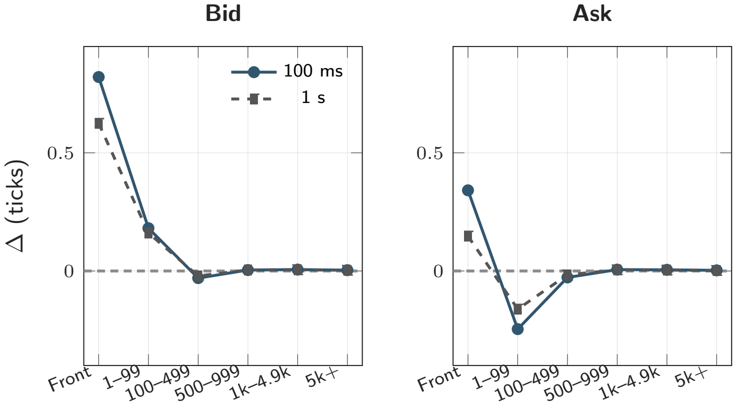
Source: Raw IEX DPLS sequence 141900297; screen applied identically to every symbol and date. DPLS-only, IEX-local evidence.

Opening Queue-Priority Curve: 09:45–10:45 ET



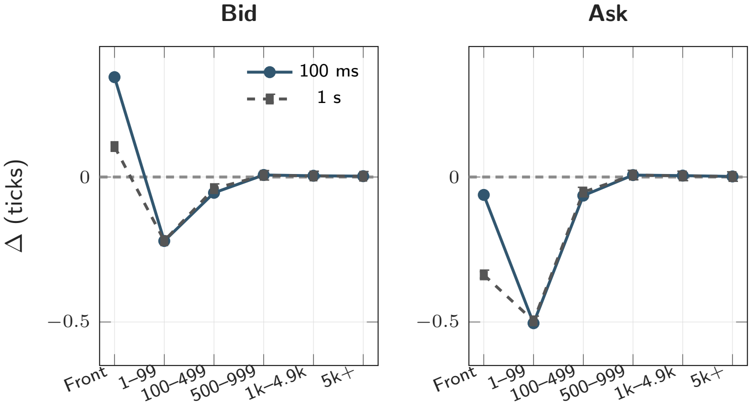
Source: Corrected 41-session sample; uniform future-IEX-BBO quote-quality gate.

Midday Queue-Priority Curve: 12:00–14:00 ET



Source: Corrected 41-session sample; uniform future-IEX-BBO quote-quality gate.

Closing Queue-Priority Curve: 14:45–15:45 ET



Source: Corrected 41-session sample; uniform future-IEX-BBO quote-quality gate.

Constructing Side-Specific Expected Drift

1. Define provider-signed drift

For a terminal stop at time t:

$$d_t^B = 10^4 \frac{m_{t+1s} - m_t}{p_t},$$

$$d_t^S = 10^4 \frac{m_t - m_{t+1s}}{p_t}.$$

B: passive bid hit; S: passive ask lifted. On either side, $d < 0$ means adverse selection for the passive provider.

2. Form expected drift on each side

$$\bar{d}^B = \frac{\sum_{B \text{ stops}} d_t^B}{N^B_{\text{valid opps}}},$$

$$\bar{d}^S = \frac{\sum_{S \text{ stops}} d_t^S}{N^S_{\text{valid opps}}}.$$

Separate denominators incorporate reach probability and keep unequal bid/ask counts from weighting the score.

Source: m_t and m_{t+1s}: pre-event and one-second IEX midpoints; p_t: execution price. Five-minute bins; at least five valid opportunities and one terminal stop per side.

Constructing the Directional-Toxicity Score

3. Take one continuous imbalance

$$T^{\text{raw}} = \bar{d}^{B} - \bar{d}^{S}.$$

$T^{\text{raw}} > 0$: ask-side provider drift is relatively more adverse $\Rightarrow$ upward pressure dominates.
$T^{\text{raw}} < 0$: bid-side drift is relatively more adverse $\Rightarrow$ downward pressure dominates.

4. Rank and standardize

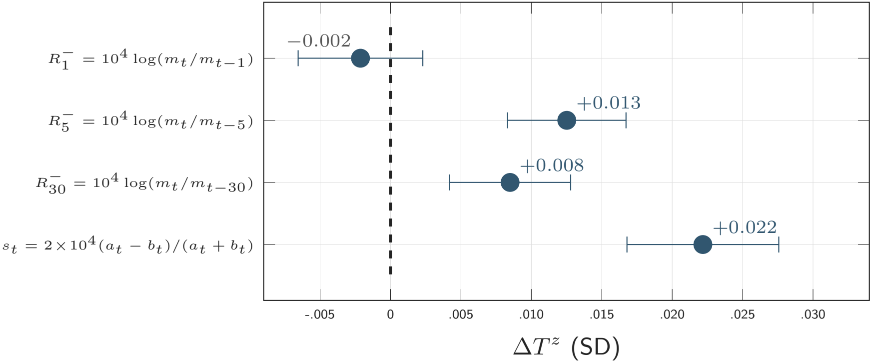
Within each date and five-minute clock bin, percentile-rank T^{raw} across eligible stocks and rescale the rank to one SD:

$$T^{z} = \text{rank\_z}(T^{\text{raw}})\,.$$

T^{z} is the stock's continuous relative position in that contemporaneous cross-section.

Source: Within every date $\times$ five-minute clock bin, the score compares each stock only with contemporaneous eligible stocks.

Step 1: What Predicts the Directional-Toxicity Score?

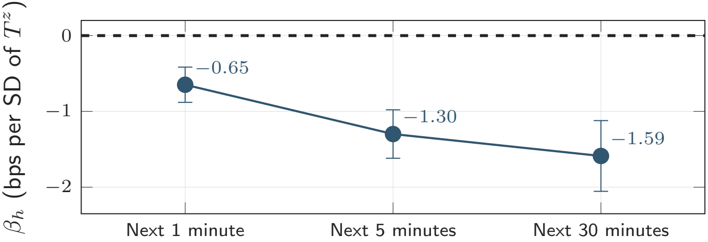


Source: t is the midpoint immediately before formation; a_t, b_t are the IEX ask and bid. 2,329 date–five-minute cross-sectional regressions; pre-states rank-standardized; predictors jointly entered; activity controlled; 41 sessions; Holm-adjusted 95% intervals.

Step 2: Directional Toxicity Forecasts Reversal

1. Recent pressure $\longrightarrow$ **2. Directional toxicity** $\longrightarrow$ 3. Correction and reversal

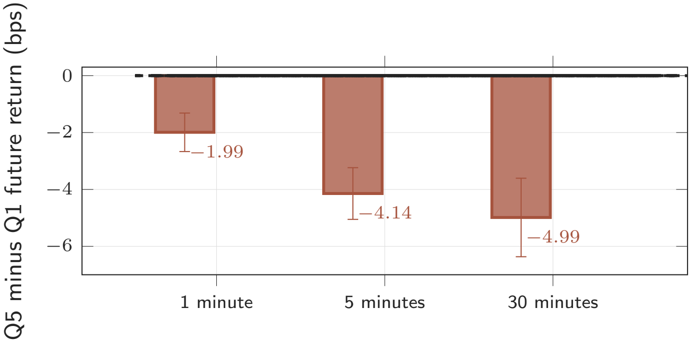
What each point means. After the full five-minute state window, $R^{(h)} = 10^4 \log(m_{\mathrm{end}+h}/m_{\mathrm{end}})$. Each β_h is controlled future midpoint basis points for a one-cross-sectional-SD higher toxicity score; $\beta_h < 0$ means reversal after ask/upward pressure.



Source: 403,558 symbol–five-minute cells, 2,614 ordinary shares, 41 sessions; pre-1/5/30-minute returns, spread, and activity controlled; Holm-adjusted 95% intervals.

Step 3: The Correction Widens from 2 to 5 Basis Points

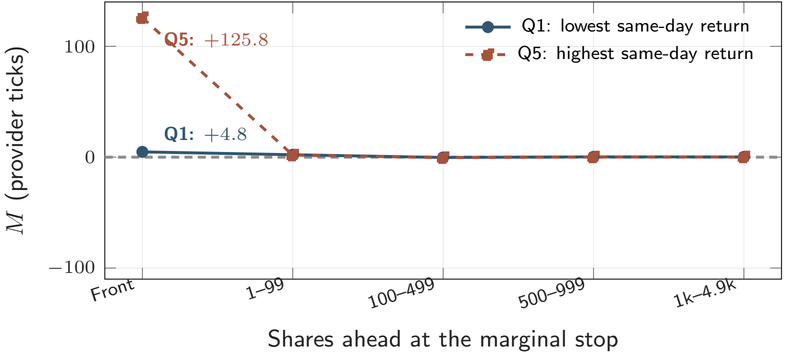
1. Recent pressure ⟶ 2. Directional toxicity ⟶ **3. Correction and reversal**



The most toxic quintile underperforms the least toxic by about 2, 4, and 5 bps over 1, 5, and 30 minutes.

Source: Date-level mean Q5–Q1 spreads; 41 sessions; 95% intervals.

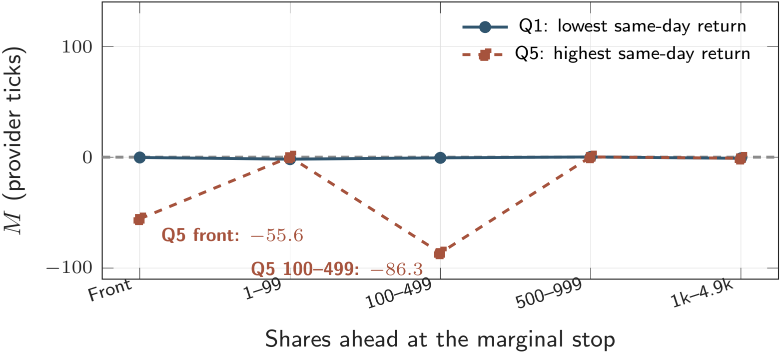
Low- and High-Return Stocks Trace Different Buy-Side Curves



At the front, the pooled 100-ms buy-side markout is sharply larger in Q5; the whole-curve contrast remains statistically uncertain.

Source: 41 IEX sessions; date-specific same-day total-return quintiles; ordinary shares; pooled conditional means.

Low- and High-Return Stocks Trace Different Sell-Side Curves



Q5 is adverse at the front and at 100–499 shares ahead; the whole-curve contrast remains statistically uncertain.

Source: 41 IEX sessions; date-specific same-day total-return quintiles; ordinary shares; pooled conditional means.

Large Point Estimates, Wide Simultaneous Bands

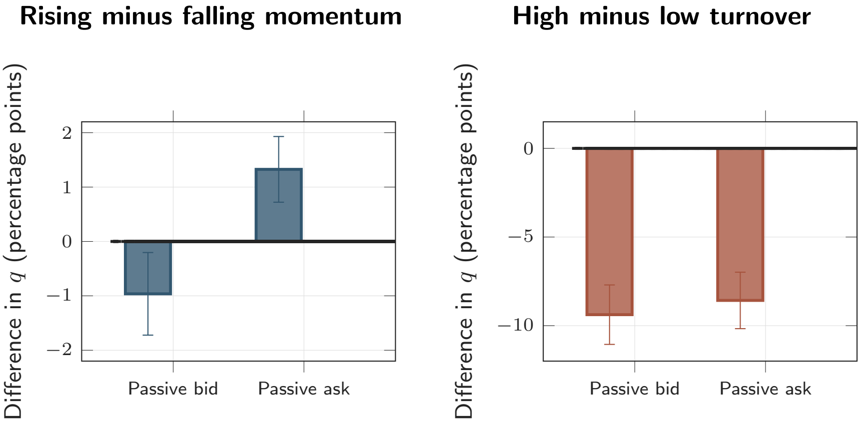
Side	Horizon	M (ticks)	q (pp)	Δ (ticks/opportunity)
Passive buy	100 ms	$+121.0$ $[-260.1, +502.1]$	-1.47 $[-3.11, +0.17]$	$+45.6$ $[-98.9, +190.2]$
Passive buy	1 s	$+121.2$ $[-260.5, +502.9]$	-1.47 $[-3.11, +0.17]$	$+45.7$ $[-99.1, +190.6]$
Passive sell	100 ms	-55.3 $[-231.8, +121.2]$	$+1.00$ $[-0.88, +2.88]$	-22.3 $[-94.3, +49.7]$
Passive sell	1 s	-95.8 $[-398.3, +206.7]$	$+1.00$ $[-0.88, +2.89]$	-38.7 $[-162.1, +84.8]$

- Brackets report 95% simultaneous intervals from 9,999 whole-date bootstrap draws.
- Every front-rank M, q, and $\Delta = qM$ interval includes zero.
- The side alignment is economically visible, but same-day return overlaps the markout window.

This is a descriptive state-dependent curve shape, not a momentum, predictive, or causal effect.

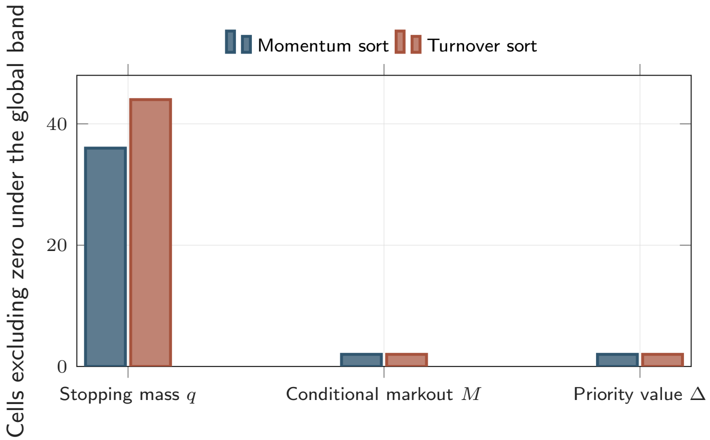
Source: 166,124 eligible symbol-dates across all 41 sessions; Q5 minus Q1; front rank.

Momentum and Turnover Shift Front Stopping Mass



Source: 41 sessions, front/100 ms; LSEG total return and turnover joined to IEX DPLS. Bars show 9,999-draw simultaneous 95% intervals.

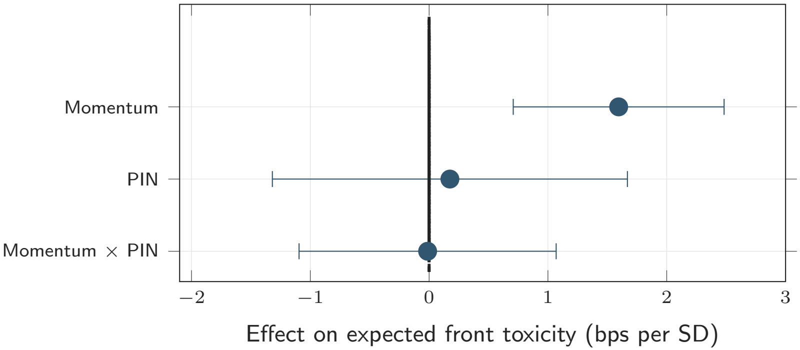
The Cross-Sectional Signal Is Concentrated in q



Stock characteristics reorganize who is reached much more consistently than the value of the marginal fill.

Source: Broad shares-ahead analysis; 48 side × rank × horizon cells per metric and sort; simultaneous 95% family.

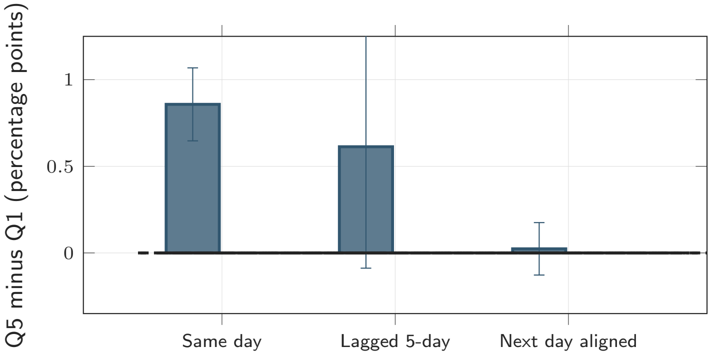
PIN Does Not Explain Overall Queue Toxicity



- Symmetric EKOP PIN is estimated stock by stock from 41 daily IEX buy/sell arrival counts.
- The positive unconditional PIN relation disappears after turnover and spread controls; conditional PIN and the interaction are null.
- Momentum remains positive after turnover and spread controls.

Source: $N = 3{,}063$ ordinary shares; HC3 95% intervals; same-window, model-implied IEX PIN.

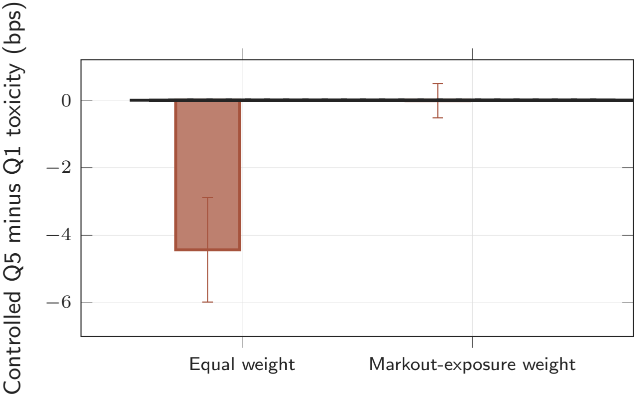
High-Information Days Move Contemporaneously, Not Predictively



High posterior information states coincide with larger returns, but the inferred news direction does not forecast the next session.

Source: Daily EKOP information-state quintiles; date fixed effects with turnover and spread controls; two-way-clustered 95% intervals.

Daily Information-State Toxicity Is Weight-Sensitive



- Equal weighting assigns the same influence to every stock-day.
- Exposure weighting gives more influence to stock-days with more valid one-second markouts.

Source: Quote-screened panel: 131,785 stock-days, 3,217 stocks, 41 sessions; date fixed effects, turnover/spread controls, two-way clustering.

Official NSE Samples Contain Substantial Hidden Reserve

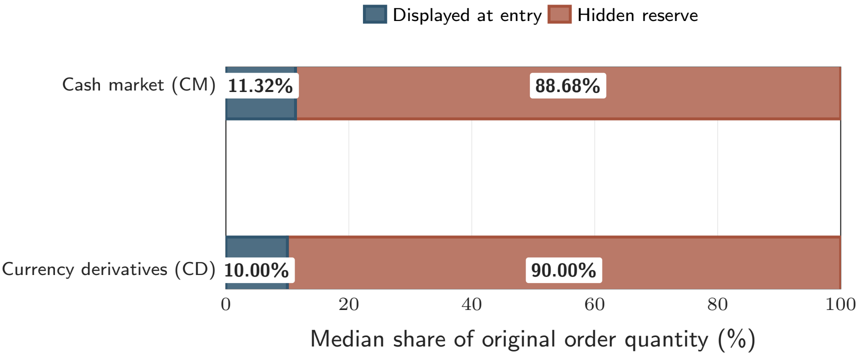
Segment	Entry records	Strict DQ	Median hidden
Capital market (CM)	213,707	19,796	88.68%
Currency derivatives (CD)	58,412	1,870	90.00%
Futures and options (FO)	98,997	0	—
Total	**371,116**	**21,666**	—

- Strict DQ requires $0 <$ disclosed quantity $<$ original quantity.
- Across all record types, the parser validates 1,514,020 order, trade, and index rows.

Official samples establish that effective priority can differ materially from visible displayed rank.

Source: Official NSE 10 April 2024 CM/FO/CD samples and product specifications.

Strict-DQ Orders Hide About 90% of Their Size



Only 10–11% of original size is visible at the median strict-DQ entry, creating a large wedge between displayed and executable priority.

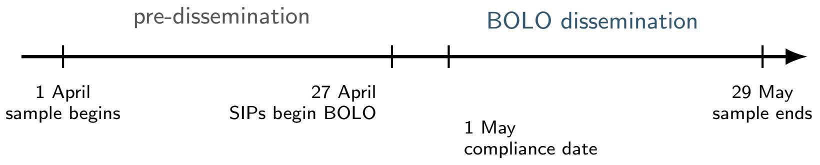
Source: Official NSE 10 April 2024 CM/CD samples; strict-DQ entries only. FO has zero strict-DQ entries. One-day feasibility sample, not a market-frequency estimate.

Why odd lots matter

- An odd lot is smaller than the security's applicable round-lot size.
- Odd-lot interest can improve the effective price frontier without defining the traditional NBBO.
- Expensive securities and retail-sized orders make odd-lot information economically important.
- Queue rank at the same nominal front can therefore represent different information states.

Source: SEC Market Data Infrastructure materials and Regulation NMS amendments.

The BOLO Transition Inside the April–May Study Window



Source: SEC Releases 34-105259 and 34-105655 (2026).

The BOLO Transition: Institutional Details

- BOLO buy: highest odd-lot bid priced above the NBB.
- BOLO sell: lowest odd-lot offer priced below the NBO.
- The formal first-business-day-of-May compliance date followed the 27 April feed start.
- DPLS alone cannot label official consolidated BOLO status.

Source: SEC Releases 34-105259 and 34-105655 (2026).

Our outcome-blind IEX-local frontier states

Three outcome-blind local frontier states

1. **Potential BOLO:** Odd-lot interest alone improves IEX's local price frontier.
2. **Mixed:** Round-lot and odd-lot interest share the local front.
3. **Round-lot:** The IEX local front is determined by round-lot interest.

The classifier uses the IEX order book and an outcome-blind local round-lot frontier. It is an economic state proxy, not a national SIP classification.

What the curves establish, and what they cannot

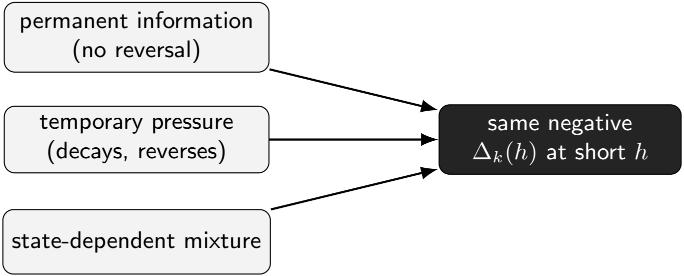
Established on 41 sessions

- $q(K)$, $M(K, h)$ and $\Delta(K, h) = qM$ are measured, censored correctly, and stable across dates.
- The interior band is adverse; the sign pattern changes between the open, midday and the close.

Not established

- Whether an adverse marginal fill is *permanent information* or *temporary pressure*.
- Whether the intraday change is a change in mechanism or a change in the mix of states.

Three mechanisms, one observed curve



- At a single horizon the three are observationally equivalent.
- They separate only through the *horizon profile* of the response and its dependence on the book state.
- Separating them requires a model with a permanent and a transitory component and a latent state that switches.

Why the standard models do not settle it

Candidate	Why it is insufficient here
Kalman filter on the midpoint alone	Signal extraction without order flow: the permanent/transitory split is driven by an assumed noise ratio, not by identified information.
Hasbrouck VAR	Linear, single-state, fixed clock, and trade-level. It measures average information content; it cannot say *which* marginal queue states carry it.
Glosten–Harris / MRR alone	Correct decomposition, but one regime and no queue coordinate. It cannot explain why the split changes across ranks and hours.
Queue-reactive / Hawkes alone	Models message arrival well, but contains no efficient-price state, so it cannot price adverse selection.
Sorts and clock dummies	Renaming the U-shaped trading day is not a mechanism; it cannot distinguish state mixing from parameter change.

What the data appears to be saying

1. **Non-monotonicity.** The adverse band sits in the interior, not at the front, which is inconsistent with a single monotone adverse-selection gradient.
2. **Reversal.** One-second directional toxicity forecasts *correction* at 1–30 minutes; a permanent-information component is not supposed to revert.
3. **Clock dependence.** The curve rotates between the open, midday and the close while the market structure is unchanged.
4. **Concentration in q.** Momentum and turnover reorganise stopping mass far more than they move fill quality.

The canonical decomposition: Glosten and Harris (1988)

Let $D_t \in \{-1, +1\}$ be the trade direction and V_t the traded volume:

$$m_t = m_{t-1} + (z_0 + z_1 V_t)\, D_t + \varepsilon_t, \tag{5}$$

$$p_t = m_t + (c_0 + c_1 V_t)\, D_t. \tag{6}$$

- $(z_0 + z_1 V_t)$ is the adverse-selection component. It enters m_t, so it is carried into every future price and never reverses.
- $(c_0 + c_1 V_t)$ is the transitory component: order processing, inventory and the discrete grid. It affects p_t only and is undone at the next trade.
- Both are allowed to grow with trade size, which is why the split is estimated rather than assumed.
- Identification comes from the differing time-series behaviour of the two pieces, not from any auxiliary spread data.

Glosten and Harris adopt $z_0 = c_1 = 0$ as their preferred specification on a 20-stock sample.

Sadka (2006): four components of price impact

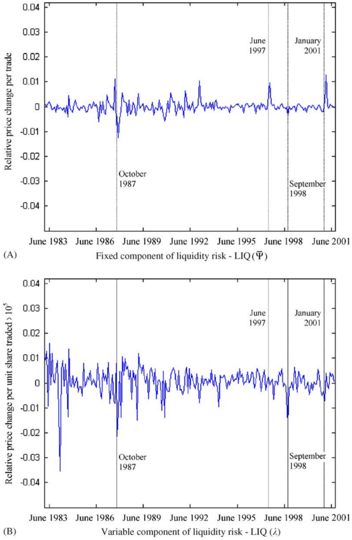
$$\Delta p_t = \Psi\,\epsilon_t^{\Psi} + \lambda\,\epsilon_t^{\lambda} + \overline{\Psi}\,\Delta D_t + \overline{\lambda}\,\Delta(D_t V_t) + y_t, \tag{7}$$

with $\epsilon_t^{\Psi} = D_t - \mathbb{E}_{t-1}[D_t]$ and $\epsilon_t^{\lambda} = D_t V_t - \mathbb{E}_{t-1}[D_t V_t]$.

	Fixed	**Variable (per share)**
Permanent (information)	Ψ	λ
Transitory (market making)	$\overline{\Psi}$	$\overline{\lambda}$

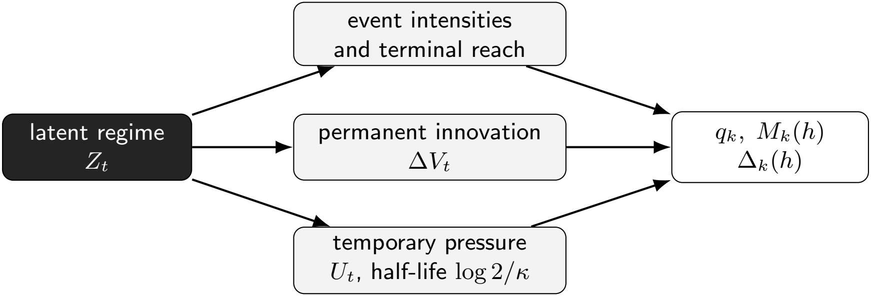
- The permanent terms respond only to *unanticipated* sign and signed volume; the transitory terms take the whole order flow.
- Sadka estimates all four, monthly and stock by stock, on NYSE 1983 to 2001, and shows they carry separate priced risk.

The four components have their own dynamics



Source: Sadka (2006), Fig. 2, *Journal of Financial Economics* 80, 309–349. Monthly innovations in the aggregate fixed ($\overline{\Psi}$) and variable (λ) components of price impact, NYSE, 1983–2001.

The model we chose: QR-PDSS



- A joint state-space *and* marked-point-process model: the same latent state generates message arrivals, queue reach, and price movement.
- The measured curves become model outputs, not the primitive.

How the model was chosen: a nested ladder

	Adds	Question it answers
M0	random walk + noise	Is any decomposition needed?
M1	$V + U$, one regime (GH/MRR)	Permanent versus temporary impact
M2	+ switching regimes	Is price discovery state-dependent?
M3	+ censored queue-reach emission	Do ranks reveal the latent state?
M4	+ queue-reactive intensities	Does full L3 flow improve inference?
M5	+ Hawkes excitation	Is clustering unexplained by the book?

- Selection is on **held-out** likelihood over an untouched final block, not in-sample fit.
- A layer is retained only if it also improves posterior reproduction of q, M and Δ.

The state and the measurement

Latent state per aggressive episode j, conditional on regime $Z_j = r$:

$$V_j = V_{j-1} + \Theta_r^\top \widetilde{g}_j + \sigma_{V,r}\varepsilon_j^V, \tag{8}$$

$$U_j = e^{-\kappa_r \Delta t_j} U_{j-1} + \Psi_r^\top g_j + \Omega_r^\top r_j + \sigma_{U,r}\varepsilon_j^U, \tag{9}$$

$$m_j = V_j + U_j + \epsilon_j^m. \tag{10}$$

- V is the efficient price: it moves only with the *unexpected* flow $\widetilde{g}_j$ and never decays.
- U is transitory displacement: it responds to total flow g_j and passive book flow r_j, and decays with half-life $\log 2/\kappa_r$.
- The observed midpoint is their sum plus tick-grid noise. This is the whole permanent/transitory split.

Only surprise is information

$$\widehat{g}_j = \mathbb{E}(g_j \mid \mathcal{F}_{j-}), \qquad \widetilde{g}_j = g_j - \widehat{g}_j, \qquad \mathbb{E}(V_j - V_{j-1} \mid \mathcal{F}_{j-}) = 0. \tag{11}$$

- g_j is not a scalar volume. It is a signed vector: size, relative reach $Y_j/D_{j-}^{\mathrm{opp}}$, level-clearing, levels crossed.
- $\widehat{g}_j$ is forecast from the exact prior book and message history, cross-fitted chronologically, never using anything after t_j.
- Order flow is strongly autocorrelated; without this step, predictable flow would mechanically create a predictable efficient price.

Queue reach enters the likelihood as censored survival

Terminal-stop hazard at rank block k:

$$\begin{aligned}\pi_{jk} &= \mathbb{P}\big(K_j = k \mid K_j \geq k,\ X_{t_j-}, Z_j, B_{j-}\big) \\ &= \operatorname{logit}^{-1}\big(\alpha_{k,Z_j} + \beta_k^\top X_{t_j-} + \gamma_k^\top f(B_{j-})\big)\,. \end{aligned} \tag{12}$$

$$\text{interior stop:} \quad \mathcal{L}_{K,j} = \pi_{jk} \textstyle\prod_{r<k}(1-\pi_{jr}), \tag{13}$$

$$\text{full clear:} \quad \mathcal{L}_{K,j}^{\text{cens}} = \textstyle\prod_{r \leq K_j^{\max}}(1-\pi_{jr}). \tag{14}$$

- This is the same censoring rule as the descriptive measurement, now carrying information about the latent state.
- Where an episode *stops* is an observation about the state that generated it.

What makes the decomposition identified

1. V is a martingale given public information and flow innovations.
2. The response of V to $\widetilde{g}$ does **not** decay; the response of U **must** decay, $\kappa_r > 0$.
3. Execution price relative to the pre-event centre is a transitory cushion, not information.
4. Multi-horizon quote responses (100 ms to 30 min) separate persistence from reversal.
5. Regimes are labelled *after* estimation from fitted persistence and reach signatures, never named in advance.

IEX alone identifies the local non-decaying component. Calling it the national efficient price requires validation against consolidated quotes.

The decisive object

$$\boxed{\Delta_k(h) = \Delta_k^{\text{cushion}}(h) + \Delta_k^{\text{permanent}}(h) + \Delta_k^{\text{temporary}}(h)} \tag{15}$$

Recovery horizon for rank k in state (X, Z):

$$h_k^* = \inf\{h : \Delta_k(h \mid X, Z) \geq 0\}, \qquad \text{PIR}_{k,r} = \frac{|\Delta_{k,r}^{\text{perm}}|}{|\Delta_{k,r}^{\text{perm}}| + |\Delta_{k,r}^{\text{temp}}|}. \tag{16}$$

- A rank negative at 1 s that crosses zero by 30 minutes is *pressure-toxic*, not information-toxic.
- Low PIR with strong short-horizon toxicity means queue depth is horizon-dependent insurance rather than an adverse-selection tax.

The likelihood has five blocks

$$\ell(\Theta) = \ell_{\text{state}} + \ell_{\text{price}} + \ell_{\text{events}} + \ell_{\text{marks}} + \ell_{\text{reach}}, \tag{17}$$

$$\ell_{\text{events}} = \sum_n \log \lambda_{e_n}(\tau_n) - \int_0^T \sum_{a \in \mathcal{A}} \lambda_a(t)\, dt, \tag{18}$$

$$\ell_{\text{reach}} = \sum_j \Big[(1 - C_j) \log \mathcal{L}_{K,j} + C_j \log \mathcal{L}^{\text{cens}}_{K,j} \Big]. \tag{19}$$

- ℓ_{price} scores the midpoint, VWAP, spread and depth equations; ℓ_{marks} scores sizes, distances and lifetimes.
- The observed-data likelihood integrates out (V, U) and sums over regime paths $Z_{1:T}$.
- Every message contributes: adds and cancels through the intensity term, executions through reach and price.

Estimation: EM with a switching Kalman filter

E-step

Run a Kim/Hamilton switching filter, add $\log \mathcal{L}_{K,j}$ to each regime's filtering weight, then smooth to obtain $\gamma_j(r) = \mathbb{P}(Z_j = r \mid O_{1:T})$, $\xi_j(r, r')$, and the smoothed moments of (V_j, U_j).

M-step

Transition matrix from $\sum_j \xi_j(r, r')$; $\Theta_r, \Psi_r, \Omega_r$ by regime-weighted least squares on smoothed states; $\kappa_r > 0$ by constrained line search; rank hazards by $\gamma_j(r)$-weighted logistic survival regression.

- Ten or more random starts; retain the best *validation* likelihood, not the best training likelihood.
- The Kim filter collapses the Gaussian mixture, so the estimates are validated against a Rao-Blackwellised particle smoother.

Rao-Blackwell EM

- Particles carry the discrete regime Z_t and the continuous liquidity and uncertainty states (L_t, H_t).
- Conditional on a particle, (V_t, U_t) stays **analytically** Kalman-filtered rather than sampled. This is the Rao-Blackwell step, and it is why the most persistent state does not degenerate.
- Weights combine the price, spread, depth, event-time, mark and censored-reach likelihoods:

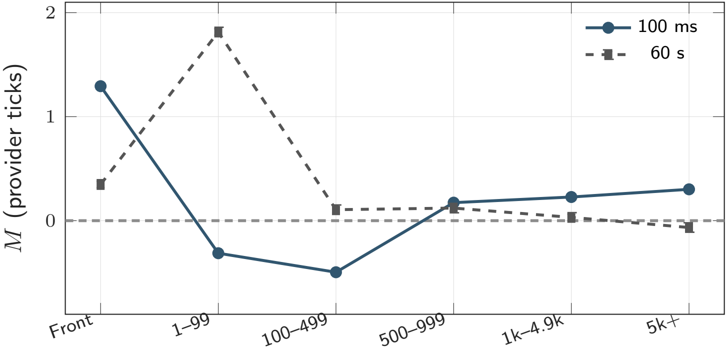
$$w_n^{(p)} \;\propto\; w_{n-1}^{(p)} \;\times\; p\Big(y_n^{\text{price}} \mid X_n^{(p)}, Z_n^{(p)}\Big) \;\times\; \lambda_{e_n}^{(p)}(\tau_n)\, f_{e_n}^{(p)}(v_n, d_n) \\ \times\; \exp\left[-\int_{\tau_{n-1}}^{\tau_n} \sum_{a \in \mathcal{A}} \lambda_a^{(p)}(u)\,\mathrm{d}u\right] \;\times\; \mathcal{L}_{K,n}^{(p)}. \tag{20}$$

It was not temporary, it was permanent

Horizon	M	Cushion	Drift
100 ms	-0.496	1.919	-2.415
1 s	-0.344	1.917	-2.261
10 s	-0.157	1.914	-2.071
60 s	$+0.106$	1.904	-1.799

- The 100–499 band recovers 0.60 ticks between 100 ms and 60 s.
- Essentially all of it is the drift unwinding. The cushion moves 0.015 ticks.
- The cushion is a level, not a dip: 1.92 ticks here against 10.17 at the front.

Fill quality by shares-ahead band



Source: 41 sessions, 28.2 million marginal stops. Realised markouts, provider-signed.

Four regimes, separating on information

Regime	Half-life (s)	$\sigma^2_{v,r}$	Permanent impact $\Theta_{r,0}$
Benign-thick	693.11	55.94	-0.013
Sniping-like	3.15	84.76	0.105
Sweep-like	3.13	76.89	0.101
Temporary-pressure-like	0.42	1058.38	-0.514

- Four regimes survive on held-out likelihood, and they are persistent.
- They differ mainly in how much information the flow carries, not in where it stops: dispersion in permanent impact is 4.5 times the dispersion in q.

The welfare context

- FIFO allocates marginal execution states by time rank.
- Toxic front priority can intensify speed, monitoring, and cancellation races.
- A queue shield can preserve displayed liquidity by filtering adverse marginal fills.
- Welfare also depends on taker execution, depth, price discovery, and participation.
- Current IEX evidence measures provider surplus locally; it does not identify total welfare.

Three findings to remember

1. **Queue rank reallocates execution exposure:** the front is reached most often, while marginal fill quality changes sign inside the queue.
2. **Directional toxicity predicts correction rather than continuation:** prices reverse over the next 1, 5, and 30 minutes.
3. **Market state reshapes the curve:** same-day return quintiles trace side-aligned M patterns, while lagged characteristics mainly reorganize q; robust $\Delta = qM$ differences remain thin.

The front bencher is not always the topper, and the identity of the topper changes with market state.

Thank you

Discussion