# The Participant Timestamp: Get The Most Out Of TAQ Data*

**Abstract:**

Starting in 2015, the participant timestamp is available alongside the primary SIP timestamp in TAQ data. This paper shows that all trades and quote updates triggered in the execution of the same marketable order receive the same participant (but not SIP) timestamp. Using this insight, TAQ can now be extended by a marketable order execution identifier. The new identifier is applied to accurately obtain otherwise biased prevailing NBBO prices and depths, to improve trade signing accuracy, to consolidate trades of marketable orders executed in parts, and to identify trades executed against hidden liquidity directly in TAQ data.

*Sander Schwenk-Nebbe, Aarhus Universitet, Department of Economics and Business Economics

# Prelude

U.S. SECURITIES AND EXCHANGE COMMISSION

Search SEC

Testimony on "Examining the SEC's Agenda, Operations and FY 2016 Budget Request"

Chair Mary Jo White

**Before the United States House of Representatives Committee on Financial Services**

**March 24, 2015**

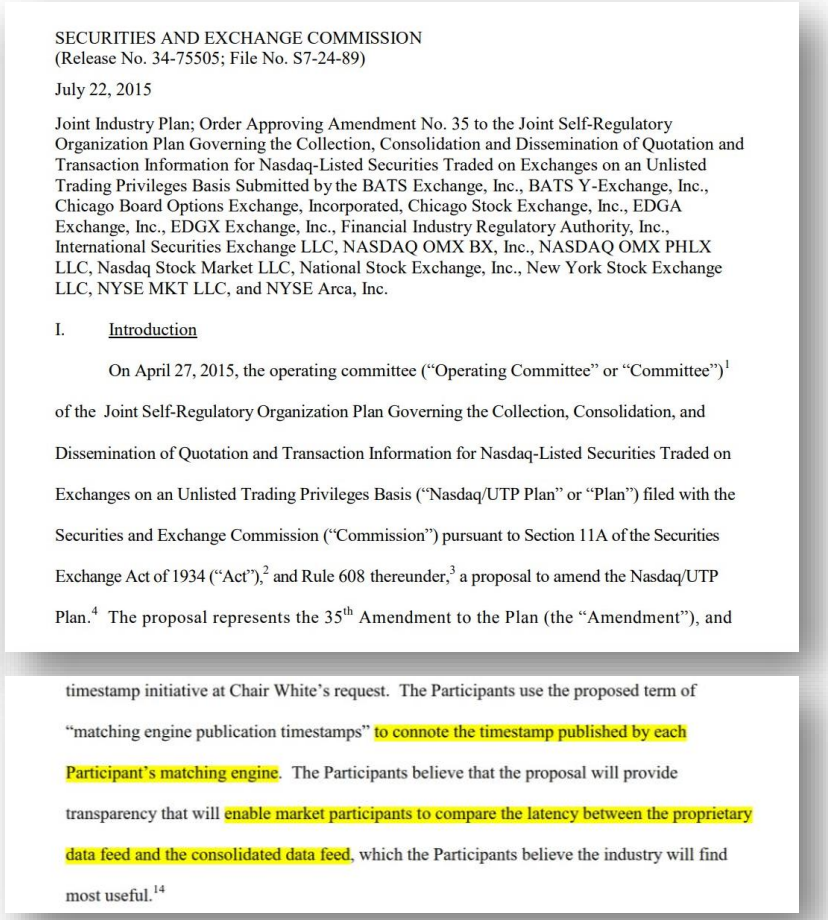
Chairman Hensarling, Ranking Members Waters, and members of the Committee:

Thank you for inviting me to testify regarding the recent activities and current initiatives of the U.S. Securities and Exchange Commission (SEC), ...

...

Substantial progress has also been made in our assessment of U.S. equity market structure to ensure that our markets remain the deepest and fairest in the world and optimally serve investors and companies of all sizes seeking to raise capital. ... I have also asked the exchanges and SIPs to incorporate a time stamp in their data feeds to facilitate greater transparency on the issue of data latency, which I expect will be operationalized this summer. ...

...



SECURITIES AND EXCHANGE COMMISSION
(Release No. 34-75505; File No. S7-24-89)

July 22, 2015

Joint Industry Plan; Order Approving Amendment No. 35 to the Joint Self-Regulatory Organization Plan Governing the Collection, Consolidation and Dissemination of Quotation and Transaction Information for Nasdaq-Listed Securities Traded on Exchanges on an Unlisted Trading Privileges Basis Submitted by the BATS Exchange, Inc., BATS Y-Exchange, Inc., Chicago Board Options Exchange, Incorporated, Chicago Stock Exchange, Inc., EDGA Exchange, Inc., EDGX Exchange, Inc., Financial Industry Regulatory Authority, Inc., International Securities Exchange LLC, NASDAQ OMX BX, Inc., NASDAQ OMX PHLX LLC, Nasdaq Stock Market LLC, National Stock Exchange, Inc., New York Stock Exchange LLC, NYSE MKT LLC, and NYSE Arca, Inc.

I. Introduction

On April 27, 2015, the operating committee ("Operating Committee" or "Committee")[1] of the Joint Self-Regulatory Organization Plan Governing the Collection, Consolidation, and Dissemination of Quotation and Transaction Information for Nasdaq-Listed Securities Traded on Exchanges on an Unlisted Trading Privileges Basis ("Nasdaq/UTP Plan" or "Plan") filed with the Securities and Exchange Commission ("Commission") pursuant to Section 11A of the Securities Exchange Act of 1934 ("Act"),[2] and Rule 608 thereunder,[3] a proposal to amend the Nasdaq/UTP Plan.[4] The proposal represents the 35th Amendment to the Plan (the "Amendment"), and

timestamp initiative at Chair White's request. The Participants use the proposed term of "matching engine publication timestamps" to connote the timestamp published by each Participant's matching engine. The Participants believe that the proposal will provide transparency that will enable market participants to compare the latency between the proprietary data feed and the consolidated data feed, which the Participants believe the industry will find most useful.[14]

- Study issues raised in Easley, O'Hara, and Yang (2016): *Differential Access to Price Information in Financial Markets*
- First to use the new timestamp in the intended way:
  - Bartlett & McCrary (2019): *How Rigged Are Stock Markets? Evidence from Microsecond Timestamps*
  - Hasbrouck (2019): *Price Discovery in High Resolution*
- This study reveals a side-effect of including the new timestamp:
  - the timestamp can be used to identify marketable order executions

[1] https://financialservices.house.gov/uploadedfiles/03.24.2015_mary_jo_white_testimony.pdf

[2] https://www.sec.gov/rules/sro/nms/2015/34-75505.pdf

# Prelude

Securities Acts Amendments of 1975[1]:

SEC. 7. The Securities Exchange Act of 1934 is amended by inserting after section 11 (15 U.S.C. 78k) the following new section: 15 USC 78k.

"NATIONAL MARKET SYSTEM FOR SECURITIES; SECURITIES INFORMATION PROCESSORS

"SEC. 11A. (a) (1) The Congress finds that— 15 USC 78k-1.

"(A) The securities markets are an important national asset which must be preserved and strengthened.

"(B) New data processing and communications techniques create the opportunity for more efficient and effective market operations.

89 STAT. 112 PUBLIC LAW 94-29—JUNE 4, 1975

"(C) It is in the public interest and appropriate for the protection of investors and the maintenance of fair and orderly markets to assure—

"(i) economically efficient execution of securities transactions;

"(ii) fair competition among brokers and dealers, among exchange markets, and between exchange markets and markets other than exchange markets;

"(iii) the availability to brokers, dealers, and investors of information with respect to quotations for and transactions in securities;

"(iv) the practicability of brokers executing investors' orders in the best market; and

"(v) an opportunity, consistent with the provisions of clauses (i) and (iv) of this subparagraph, for investors' orders to be executed without the participation of a dealer.

"(D) The linking of all markets for qualified securities through communication and data processing facilities will foster efficiency, enhance competition, increase the information available to brokers, dealers, and investors, facilitate the offsetting of investors' orders, and contribute to best execution of such orders.

National market system for securities, establishment.

"(2) The Commission is directed, therefore, having due regard for the public interest, the protection of investors, and the maintenance of fair and orderly markets, to use its authority under this title to facilitate the establishment of a national market system for securities (which may include subsystems for particular types of securities with

[TERMINOLOGY]
Securities Information Processor (SIP)[1]:

PUBLIC LAW 94-29—JUNE 4, 1975

"(22) (A) The term 'securities information processor' means any person engaged in the business of (i) collecting, processing, or preparing for distribution or publication, or assisting, participating in, or coordinating the distribution or publication of, information with respect to transactions in or quotations for any security (other than an exempted security) or (ii) distributing or publishing (whether by means of a ticker tape, a communications network, a terminal display device, or otherwise) on a current and continuing basis, information with respect to such transactions or quotations. The term 'securities

[SIP TIMESTAMP IS ALSO KNOWN AS]
WRDS: time_m; TAQ: utcsec

[TERMINOLOGY]
Participant[2]:

**§ 242.600 NMS security designation and definitions.**

**(b)** For purposes of Regulation NMS (§§ 242.600 through 242.612), the following definitions shall apply:

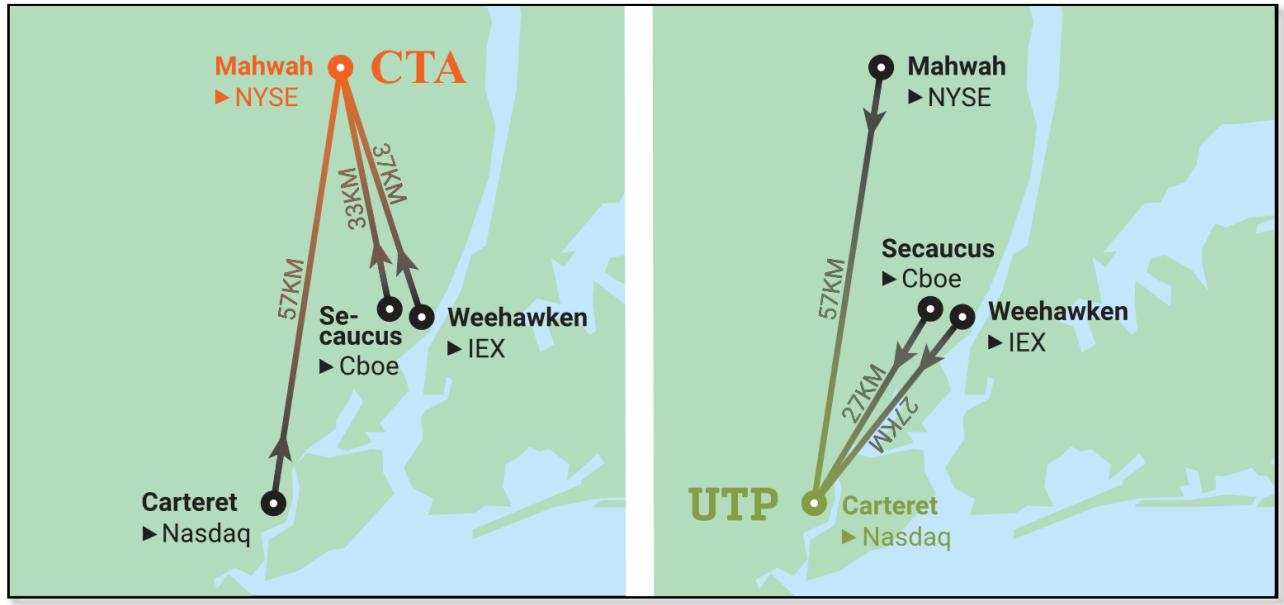
...

**(65)** ***Participants***, when used in connection with a national market system plan, means any self-regulatory organization which has agreed to act in accordance with the terms of the plan but which is not a signatory of such plan.

...

[PARTICIPANT TIMESTAMP IS ALSO KNOWN AS]
WRDS: part_time; TAQ: participant time/timestamp; exchange timestamp

Current system, Participants report to CTA SIP (left) or UTP SIP (right), depicts geographical location and distance:



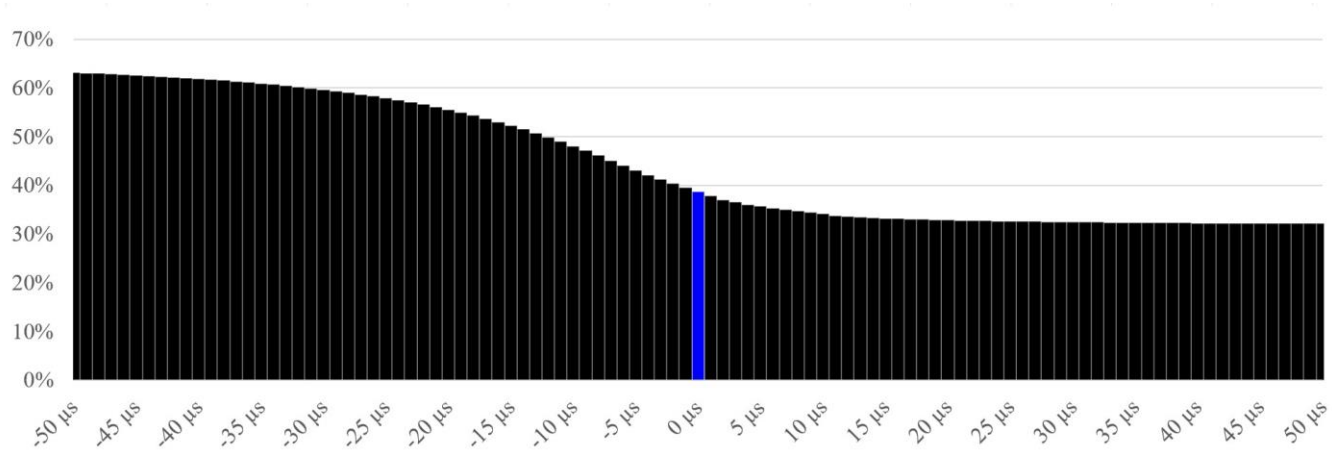

[1] https://www.govinfo.gov/content/pkg/STATUTE-89/pdf/STATUTE-89-Pg97.pdf

[2] https://www.law.cornell.edu/cfr/text/17/242.600

# Observation that sparked the study

- Asynchronicity between trade and quote reporting in US equity markets has been studied for decades[1]
- To shed new light on the old issue: gauge trade and quote agreement in time around trades
- Short timeframe, ±50 microseconds (0.05 milliseconds)
- There appears to be delay: quotes respond on trades before those trades are reported/timestamped
- In line with Carrion and Kolay (2020), who find that lagging trade SIP time relative to quote SIP time continues to improve trade signing

Proportion of trades traded at "prevailing" quoted prices at various offsets, using SIP timestamps:

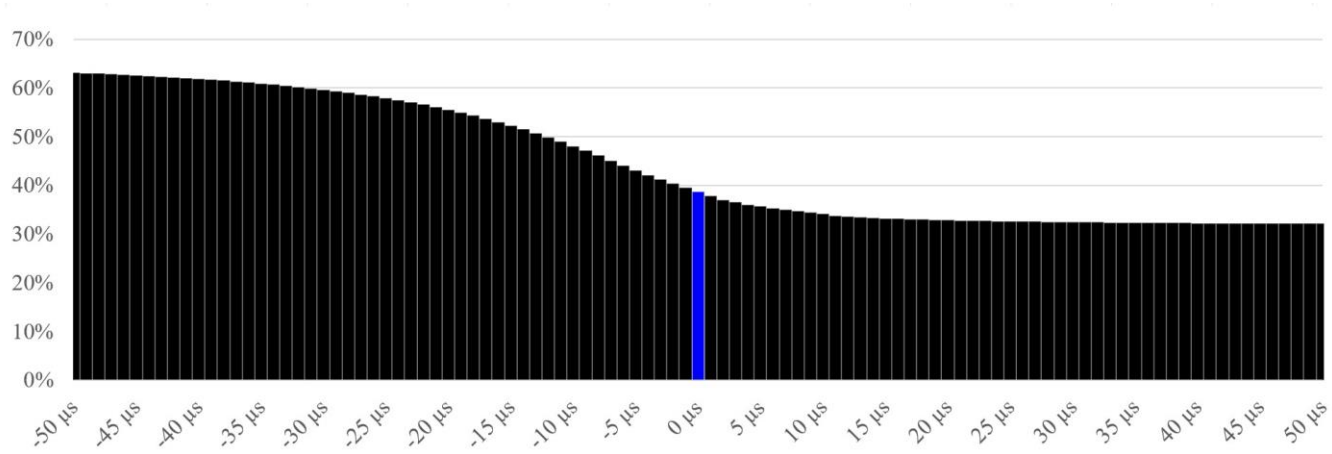



[1] See, for example, Lee and Ready (1991), Hasbrouck et al. (1993), Blume and Goldstein (1997), Odders-White (2000), Peterson and Sirri (2003), Vergote (2005), Piwowar and Wei (2006), Henker and Wang (2006), Chakrabarty et al. (2012), and Carrion and Kolay (2020)
[2] Carrion and Kolay (2020): *Trade signing in fast markets*

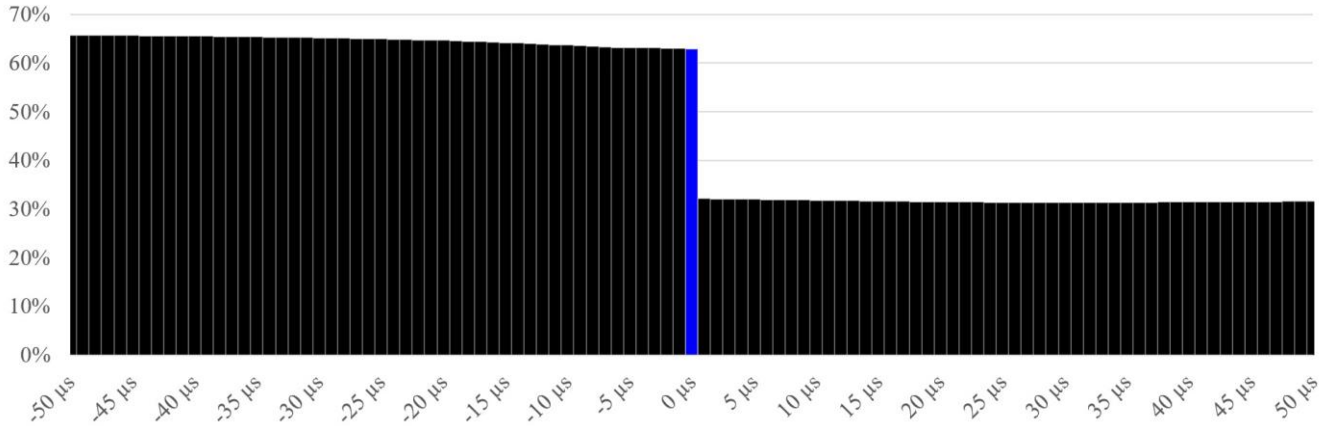
# Observation that sparked the study

- Asynchronicity between trade and quote reporting in US equity markets has been studied for decades[1]
- To shed new light on the old issue: gauge trade and quote agreement in time around trades
- Short timeframe, ±50 microseconds (0.05 milliseconds)
- There appears to be delay: quotes respond on trades before those trades are reported/timestamped
- In line with Carrion and Kolay (2020), who find that lagging trade SIP time relative to quote SIP time continues to improve trade signing
- In participant time, quotes respond immediately

Proportion of trades traded at "prevailing" quoted prices at various offsets, using SIP timestamps:




Proportion of trades traded at "prevailing" quoted prices at various offsets, using Participant timestamps:




[1] See, for example, Lee and Ready (1991), Hasbrouck et al. (1993), Blume and Goldstein (1997), Odders-White (2000), Peterson and Sirri (2003), Vergote (2005), Piwowar and Wei (2006), Henker and Wang (2006), Chakrabarty et al. (2012), and Carrion and Kolay (2020)
[2] Carrion and Kolay (2020): *Trade signing in fast markets*

# Data

- The issue is systemic, all (on-exchange) TAQ data is affected
- [1] covers August 2015 to April 2021
- [1] excludes securities that:
  - cannot be matched with CRSP
  - have zero mcap, volume, or median traded price of less than $1
  - all else (including ETFs) is included
- Leaves 2,871 securities; to reduce computational load all but Wednesdays are excluded in [1]
  - except in comparison to [4]
- Although issue is systemic, some securities may be more affected than others, hence the sub-grouping

| Ex/Group | Primary (TAQ) [1] | Nasdaq ITCH & NYSE IF [2] | TSPP Appendix B.II [3] | SEC Market Structure Data [4] |
|---|---|---|---|---|
| Nasdaq | 1421.2M | 19.3M | 66.4M | 6894.5M |
| Nasdaq BX | 206.6M | 1.7M | 19.6M | 1090.6M |
| Nasdaq PLX | 63.9M | .7M | 2.3M | 322.8M |
| NYSE | 510.1M | 6.3M | 12.3M | 2499.6M |
| NYSE Arca | 634.7M | 7.2M | 22.1M | 2747.3M |
| Cboe EDGA | 197.4M | | 4.9M | 1012.5M |
| Cboe EDGX | 456.1M | | 15.3M | 2170.7M |
| Cboe BYX | 318.5M | | 20.8M | 1694.5M |
| Cboe BZX | 601.9M | | 15.9M | 2858.4M |
| CTA | 2962.9M | 23.6M | | |
| UTP | 1447.4M | 11.7M | | |
| Equities | 3857.2M | 30.2M | | |
| ETFs | 553.1M | 5.0M | | |
| Low MCap | 201.4M | 1.6M | | |
| Med MCap | 822.3M | 6.8M | | |
| High MCap | 3386.6M | 26.8M | | |
| Low DVol | 110.9M | 1.1M | | |
| Med DVol | 615.8M | 5.5M | | |
| High DVol | 3683.6M | 28.5M | | |
| Low Price | 247.7M | 1.4M | | |
| Med Price | 3175.8M | 22.3M | | |
| High Price | 986.8M | 11.5M | | |
| All | 4410.3M | 35.2M | 179.6M | 21.3B |

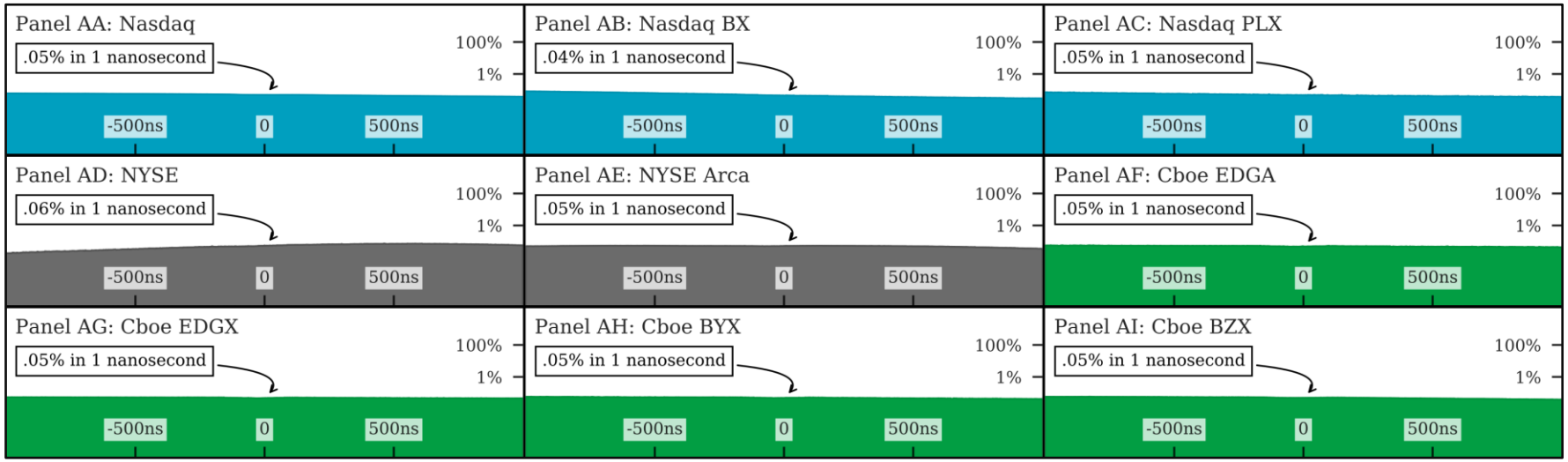
[2] Available at ftp://ftp.nysedata.com/ and ftp://ftp.nasdaqtrader.com/; Nasdaq ITCH parser from https://github.com/martinobdl/ITCH; NYSE IF parser based on https://github.com/martinobdl/ITCH
[3] Available at https://www.sec.gov/page/dera_ticksizepilot_appb
[4] Available at https://www.sec.gov/opa/data/market-structure/market-structure-data-security-and-exchange

# The Key Difference

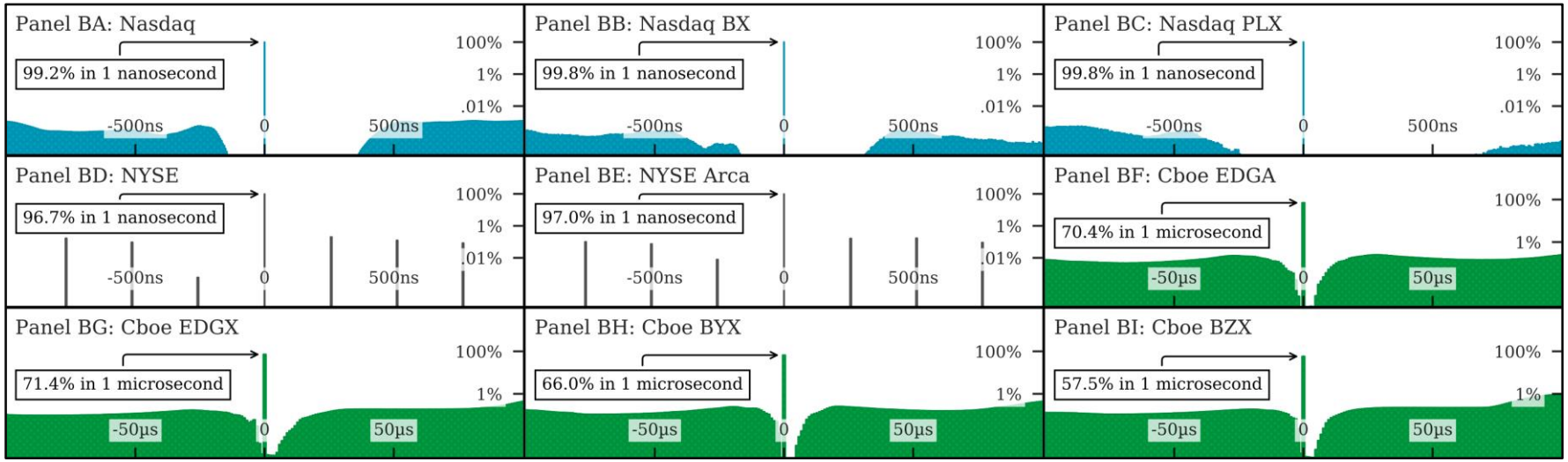
## SIP vs Participant timestamps

- Gauge the occurrence of quote updates in time around trades in the same security × exchange
- X-axis: time relative to trade
- Y-axis: [# quote updates in bin] over [# of quote updates in all bins]
- # bins in nanoseconds: 2000+1
- # bins in microseconds: 200+1
- Note the log scale on y-axis

Proportion of quote updates in bins around trades, using SIP timestamps:




Proportion of quote updates in bins around trades, using Participant timestamps:

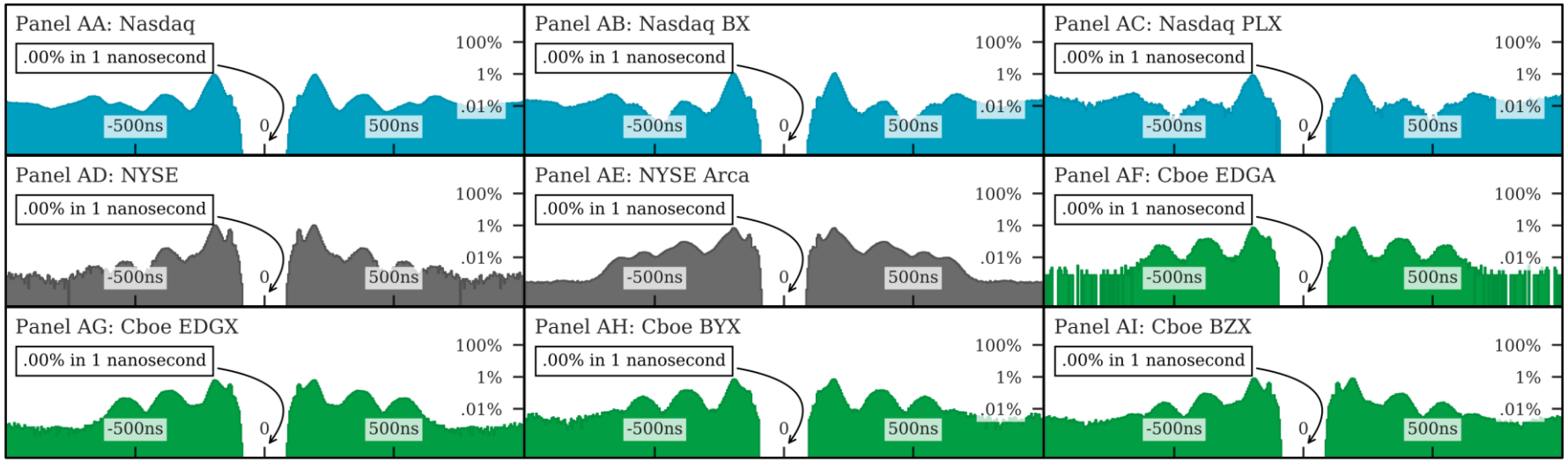


# The Key Difference

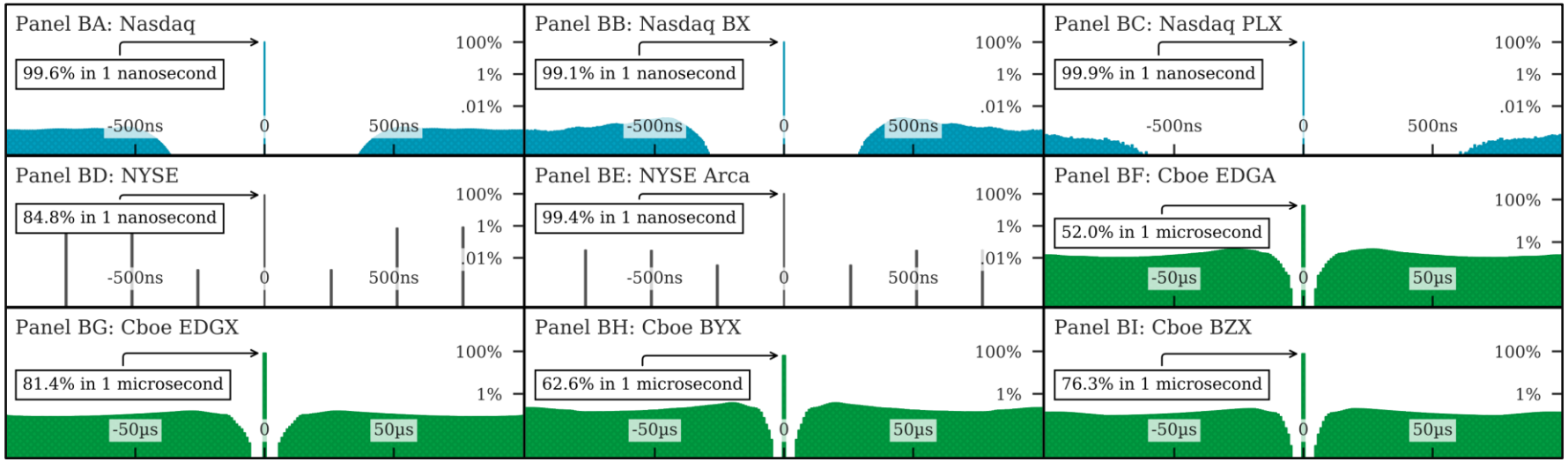
## SIP vs Participant timestamps

- Gauge the occurrence of other trades in time around trades in the same security × exchange
- Otherwise, same exact methodology

Proportion of other trades in bins around trades, using SIP timestamps:




Proportion of other trades in bins around trades, using Participant timestamps:



# The Key Difference

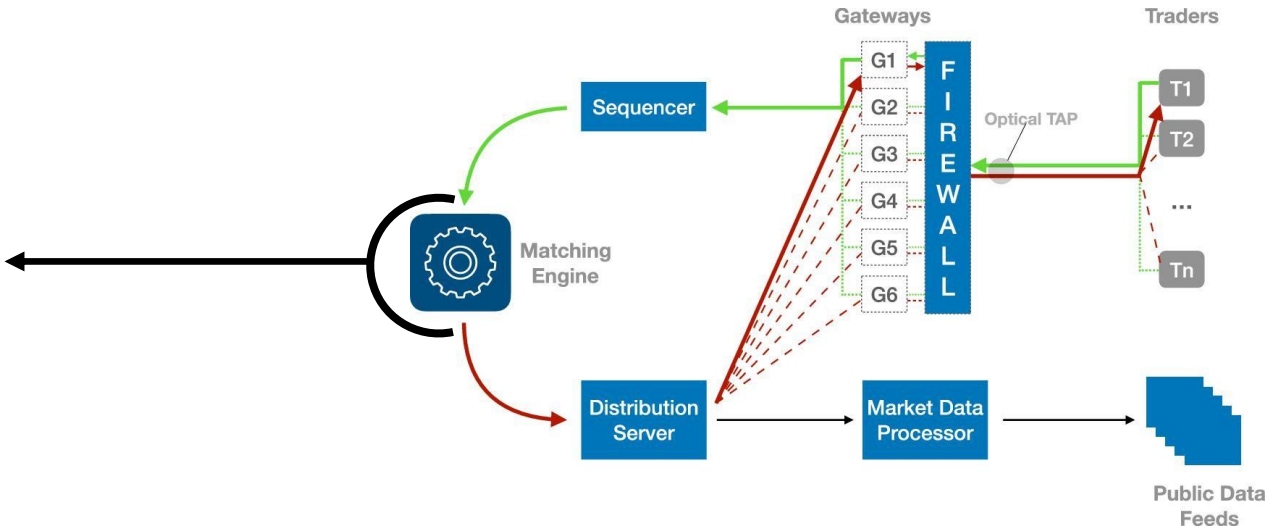
## SIP vs Participant timestamps

- The key difference between the timestamps are where they are set
- The SIP timestamp is exclusively set by the SIP and the Participant timestamp is exclusively set by the exchange's matching engine[1]
- We won't get to see what's inside the matching engine, but we can type up some pseudo code
- The SIPs receive trades and quotes in separate messages, which contain no information about the originating marketable order[1]

```
Handle_Message_From_Sequencer (In_Message):
{
  // lock order book of security
  Message := Parser(In_Message);
  Timestamp, Out_Messages := Handle_Parsed_Message(Message);
  // unlock order book of security, send market data
}

Handle_Parsed_Message (Message):
{
  // allocate returned variables:
  Out_Messages := [];
  Timestamp := Get_Timestamp_From_Precise_Clock();
  // .. ⌛ .. match and execute the order:
  if Message is Market or Marketable Limit Order:
  {
    for each matching Resting Limit Order:
    {
      // execute:
      // - generate trade, add to Out_Messages
      // - update state of book, add to Out_Messages
    }
  }
  return Timestamp, Out_Messages;
}
```

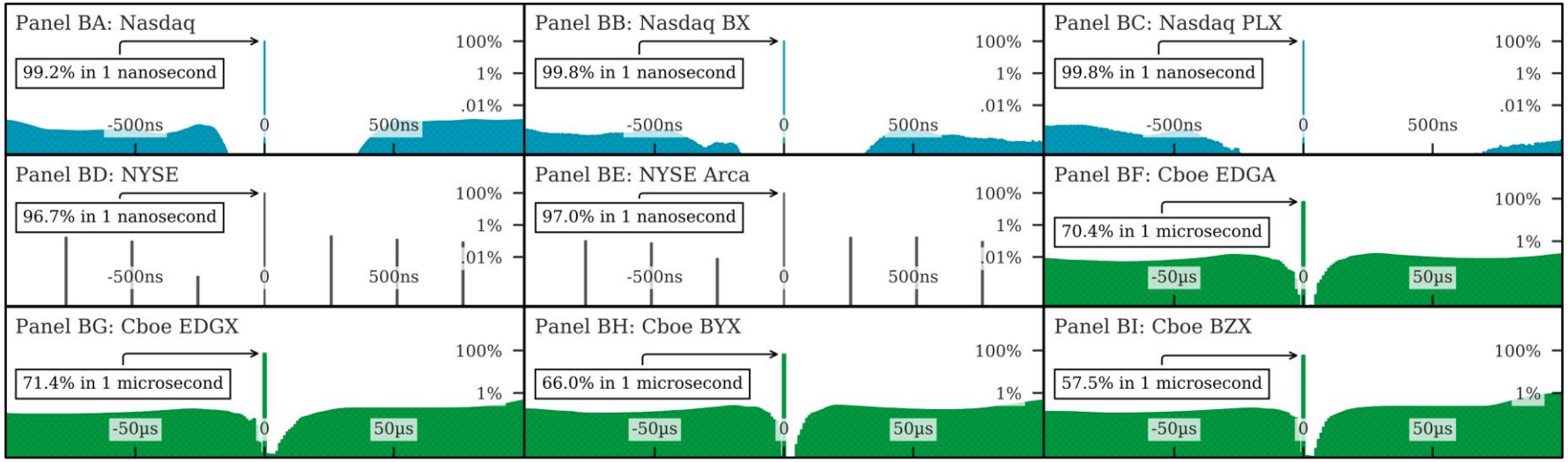
Figure 1 from Aquilina, Budish, and O'Neill (2022):



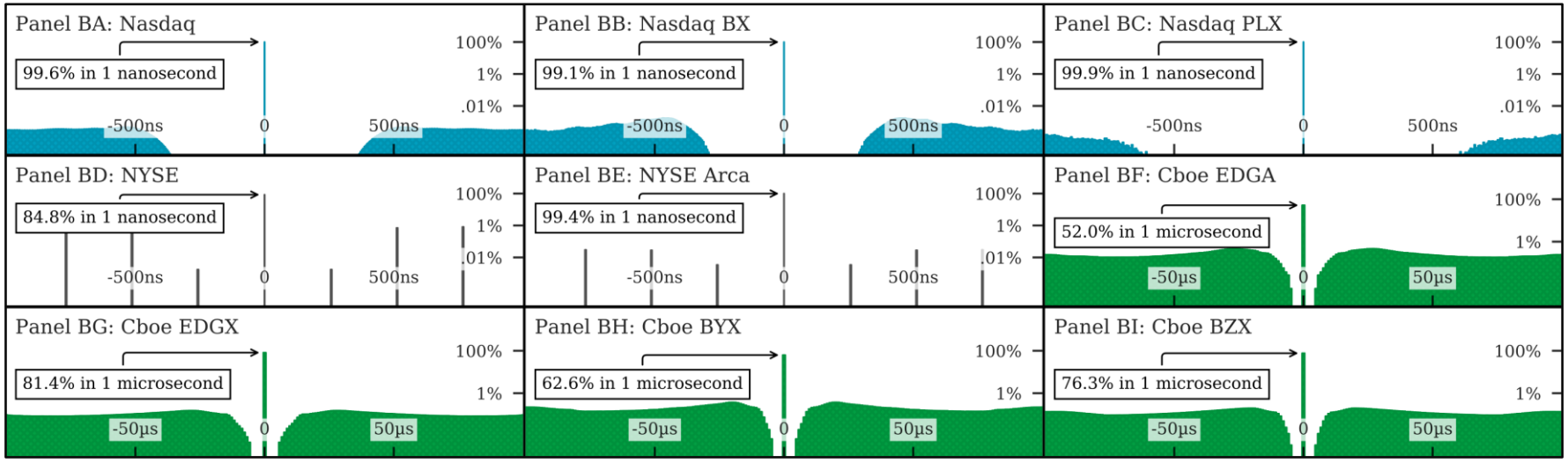

[1] inbound protocol from CTA and UTP SIP, available at: https://utpplan.com/DOC/UtpBinaryInputSpec.pdf; https://www.ctaplan.com/publicdocs/ctaplan/CTS_Pillar_Input_Specification.pdf; https://www.ctaplan.com/publicdocs/ctaplan/CQS_Pillar_Input_Specification.pdf
[2] Aquilina, Budish, and O'Neill (2022): *Quantifying the High-Frequency Trading "Arms Race"*

# The Key Difference
## SIP vs Participant timestamps

Quotes around trades in Participant time:



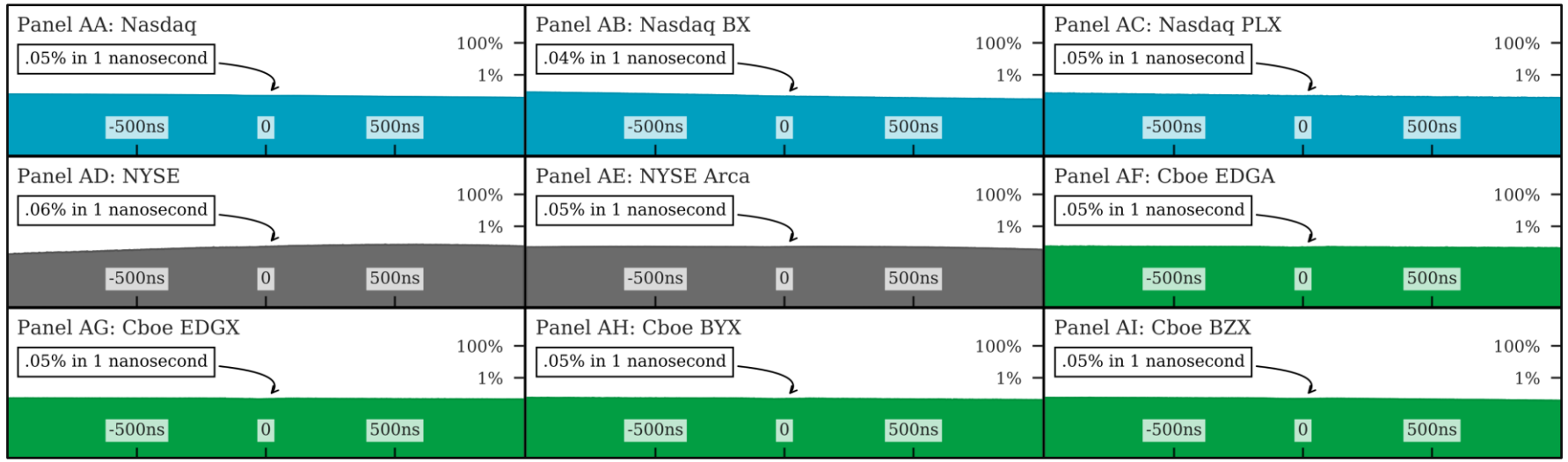

Other trades around trades in Participant time:



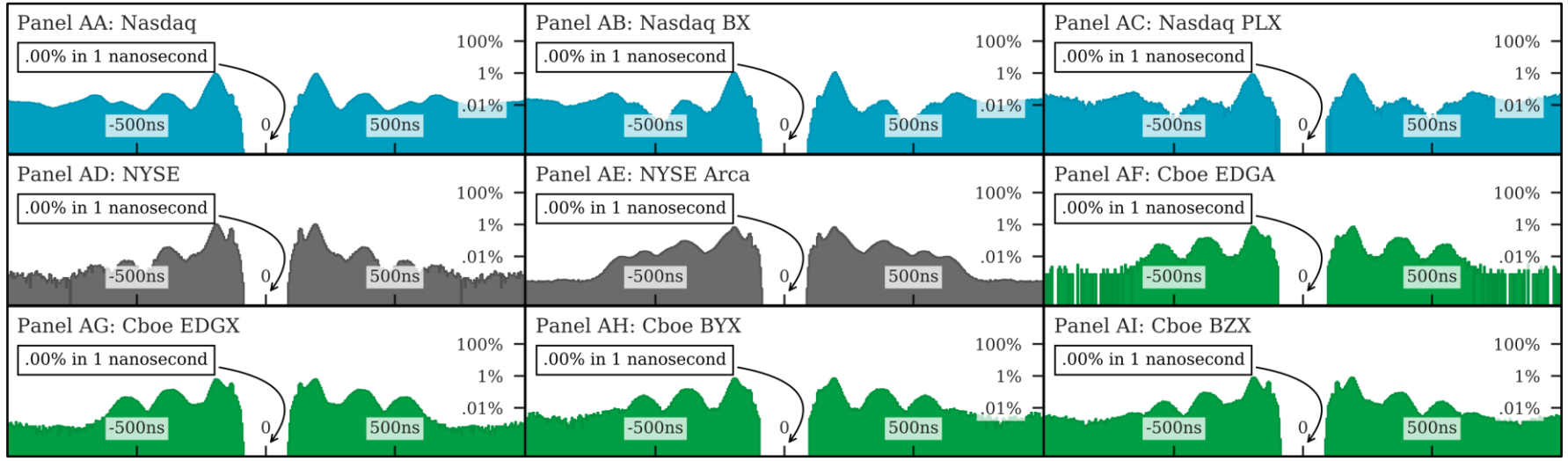

- ✓ Trades share Participant timestamp at high resolution with quotes/other trades in the same sec×ex
- ✓ No quotes can be updated immediately before and immediately after trades in the same sec×ex
- ✓ Some processing is done before timestamp is set, more is done afterwards

# The Key Difference
## SIP vs Participant timestamps

Quotes around trades in Participant time:



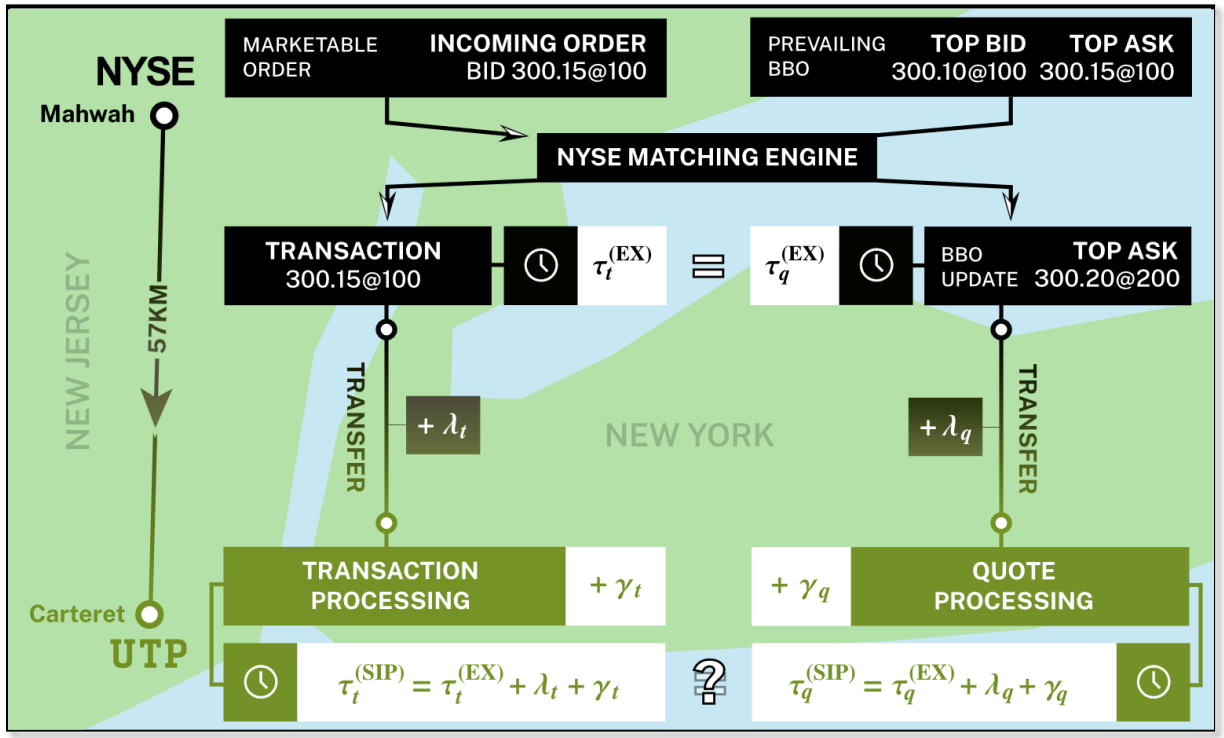

Other trades around trades in SIP time:




- ✓ Quotes do not respond on trades; they are handled entirely separately
- ✓ Some quotes receive the same timestamp as trades, but at uniform expectation (1/2001 ≈ 0.05%)
- ✓ Because quotes are not locked, there is no guarantee that quotes with the same timestamps have anything to do with the trades
- ✓ Trades are still handled sequentially, but one-by-one
- ✓ The simpler task for the SIPs translates into lower processing time

# In conjunction with Dissemination Latency…

- SIP timestamp = [participant timestamp] + [dissemination latency]
- Dissemination latency is stochastic and consists of:
  - network latency: time spent transmitting events from exchange to SIP
  - processing latency: time spent processing events at the SIP
- We may assume that the average transmission latency (λ) does not differ between trades and quotes
- We cannot assume that the average processing latency (γ) of trades equals that of quotes
- SIPs report their processing latency: trade latency is consistently larger than that of quotes[1]
- The concurrent quote tends to receive an earlier SIP timestamp than the trade




[1] see “Average Latency” and “Median Latency” under “Processor Metrics” on www.utpplan.com/ and www.ctaplan.com/

# Example 1

## Trades and quotes in AAPL on Nasdaq in brief period on April 19, 2021

- Sequence arranged by SIP timestamp

| | | Transaction | | Top-of-Book Update | | |
|---|---|---|---|---|---|---|
| Type | SIP Timestamp | Trade ID | Trade | Quote ID | Bid | Ask |
| Quote | 11:50:19.031888803 | | | 43840785 | 1800@134.30 | 1400@134.31 |
| — | 11:50:19.031912255 | | | 43840786 | 1800@134.30 | 1000@134.31 |
| Trade | 11:50:19.031913039 | 61064 | 400@134.31 | | | |
| Quote | 11:50:19.031920421 | | | 43840790 | 1700@134.30 | 1000@134.31 |
| — | 11:50:19.031935216 | | | 43840795 | 1200@134.30 | 1000@134.31 |
| Trade | 11:50:19.031944590 | 61065 | 600@134.31 | | | |
| Quote | 11:50:19.031945358 | | | 43840796 | 1200@134.30 | 400@134.31 |
| — | 11:50:19.031945483 | | | 43840797 | 1200@134.30 | 300@134.32 |
| Trade | 11:50:19.031946321 | 61066 | 400@134.31 | | | |
| — | 11:50:19.031948306 | 61067 | 200@134.31 | | | |
| — | 11:50:19.031950918 | 61068 | 200@134.31 | | | |
| — | 11:50:19.031956938 | 61069 | 8@134.31 | | | |
| Quote | 11:50:19.031983974 | | | 43840801 | 1300@134.30 | 300@134.32 |

- Sequence arranged by Participant timestamp
- Events with same Participant timestamp sorted according to causal chain
- Executions against hidden liquidity inferred

| | | | | Transaction | | Top-of-Book Update | | |
|---|---|---|---|---|---|---|---|---|
| Type | SIP Timestamp | Participant Timestamp | *MOX Identifier* | Trade ID | Trade | Quote ID | Bid | Ask |
| Quote | 11:50:19.031888803 | 11:50:19.031869975 | | | | 43840785 | 1800@134.30 | 1400@134.31 |
| Trade | 11:50:19.031913039 | 11:50:19.031**898218** | 1. | 61064 | 400@134.31 | | | |
| Quote | 11:50:19.031912255 | 11:50:19.031**898218** | 1. | | | 43840786 | 1800@134.30 | 1000@134.31 |
| — | 11:50:19.031920421 | 11:50:19.031903863 | | | | 43840790 | 1700@134.30 | 1000@134.31 |
| — | 11:50:19.031935216 | 11:50:19.031918407 | | | | 43840795 | 1200@134.30 | 1000@134.31 |
| Trade | 11:50:19.031944590 | 11:50:19.031**924671** | 2. | 61065 | 600@134.31 | | | |
| Quote | 11:50:19.031945358 | 11:50:19.031**924671** | 2. | | | 43840796 | 1200@134.30 | 400@134.31 |
| Trade | 11:50:19.031946321 | 11:50:19.031**924671** | 2. | 61066 | 400@134.31 | | | |
| Quote | 11:50:19.031945483 | 11:50:19.031**924671** | 2. | | | 43840797 | 1200@134.30 | 300@134.32 |
| Trade | 11:50:19.031948306 | 11:50:19.031**924671** | 2. | 61067 | 200@134.31 | | | *hidden* |
| — | 11:50:19.031950918 | 11:50:19.031**924671** | 2. | 61068 | 200@134.31 | | | *hidden* |
| — | 11:50:19.031956938 | 11:50:19.031**931283** | 3. | 61069 | 8@134.31 | | | *odd-lot* |
| Quote | 11:50:19.031983974 | 11:50:19.031970315 | | | | 43840801 | 1300@134.30 | 300@134.32 |

# Example 2
## Trade and Quote excerpt from AFXZ[1]

Excerpt of Page 7 from AFXZ[1]:

A snapshot of the quote update data is illustrated in Table 3. Each row in the quote data corresponds to the NBBO at a certain timestamp. The third line of quote update is likely caused by the fourth transaction shown in Table 2. O'Hara et al. (2014) report possible issues with the lack of records of odd-lot trades when TAQ only recorded round-lot trades; TAQ started to include odd-lot trades since 2014, as we can see in Table 3. The quotes are still round-lot but this should have only a minimal impact on our response variables.

Excerpt of trade data from TAQ; Table 2 from AFXZ[1]:

Table 2: Examples of trade data: INTC on Jan. 3rd, 2019

| Time | Price | Size | Direction (Lee-Ready) |
|---|---|---|---|
| 10 : 07 : 48.956770900 | 45.18 | 100 | −1 |
| 10 : 07 : 48.956773554 | 45.18 | 300 | −1 |
| 10 : 07 : 48.956916983 | 45.18 | 100 | −1 |
| 10 : 07 : 48.956971093 | 45.18 | 100 | +1 |
| 10 : 07 : 48.957830128 | 45.18 | 66 | +1 |

Excerpt of quote data from TAQ; Table 3 from AFXZ[1]:

Table 3: Example of quote data: INTC on Jan. 3rd, 2019

| Time | Best Bid Price | Best Bid Size | Best Ask Price | Best Ask Size |
|---|---|---|---|---|
| 10 : 07 : 48.956906761 | 45.18 | 100 | 45.19 | 4800 |
| 10 : 07 : 48.956921135 | 45.18 | 100 | 45.19 | 4700 |
| 10 : 07 : 48.956970663 | 45.17 | 1600 | 45.19 | 4700 |
| 10 : 07 : 48.956980355 | 45.17 | 1600 | 45.19 | 4100 |
| 10 : 07 : 48.956991775 | 45.17 | 1600 | 45.19 | 4000 |

| Exchange | Time (SIP) | Price | Size | Participant Time |
|---|---|---|---|---|
| NYSE Arca | 10:07:48.956770900 | 45.18 | 100 | 10:07:48.956381184 |
| NYSE | 10:07:48.956773554 | 45.18 | 300 | 10:07:48.956400128 |
| Cboe BZX | 10:07:48.956916983 | 45.18 | 100 | 10:07:48.956700000 |
| Nasdaq | 10:07:48.956971093 | 45.18 | 100 | **10:07:48.956953405** |
| FINRA | 10:07:48.957830128 | 45.18 | 66 | 10:07:48.956000000 |

| Exchange | Time (SIP) | Participant Time | Bid | Bid Size | Ask | Ask Size |
|---|---|---|---|---|---|---|
| Nasdaq | 10:07:48.956906760 | 10:07:48.956886483 | 45.18 | 100 | 45.19 | 4800 |
| Nasdaq | 10:07:48.956921134 | 10:07:48.956905260 | 45.18 | 100 | 45.19 | 4700 |
| Nasdaq | 10:07:48.956970662 | **10:07:48.956953405** | 45.17 | 1600 | 45.19 | 4700 |
| Nasdaq | 10:07:48.956980354 | 10:07:48.956964125 | 45.17 | 1600 | 45.19 | 4100 |
| Nasdaq | 10:07:48.956991774 | 10:07:48.956975830 | 45.17 | 1600 | 45.19 | 4000 |

- Note that the SIP timestamp of the quote is earlier than the SIP timestamp of the trade

[1] Ait-Sahalia, Fan, Xue, and Zhou (2022): *How and When are High-Frequency Stock Returns Predictable?*

# Example 2
## Trade and Quote excerpt from AFXZ[1]

Table 2 from AFXZ[1]:

Table 2: Examples of trade data: INTC on Jan. 3rd, 2019

| Time | Price | Size | Direction (Lee-Ready) |
|---|---|---|---|
| 10 : 07 : 48.956770900 | 45.18 | 100 | −1 |
| 10 : 07 : 48.956773554 | 45.18 | 300 | −1 |
| 10 : 07 : 48.956916983 | 45.18 | 100 | −1 |
| 10 : 07 : 48.956971093 | 45.18 | 100 | +1 |
| 10 : 07 : 48.957830128 | 45.18 | 66 | +1 |

Extended excerpt from TAQ with Participant timestamp:

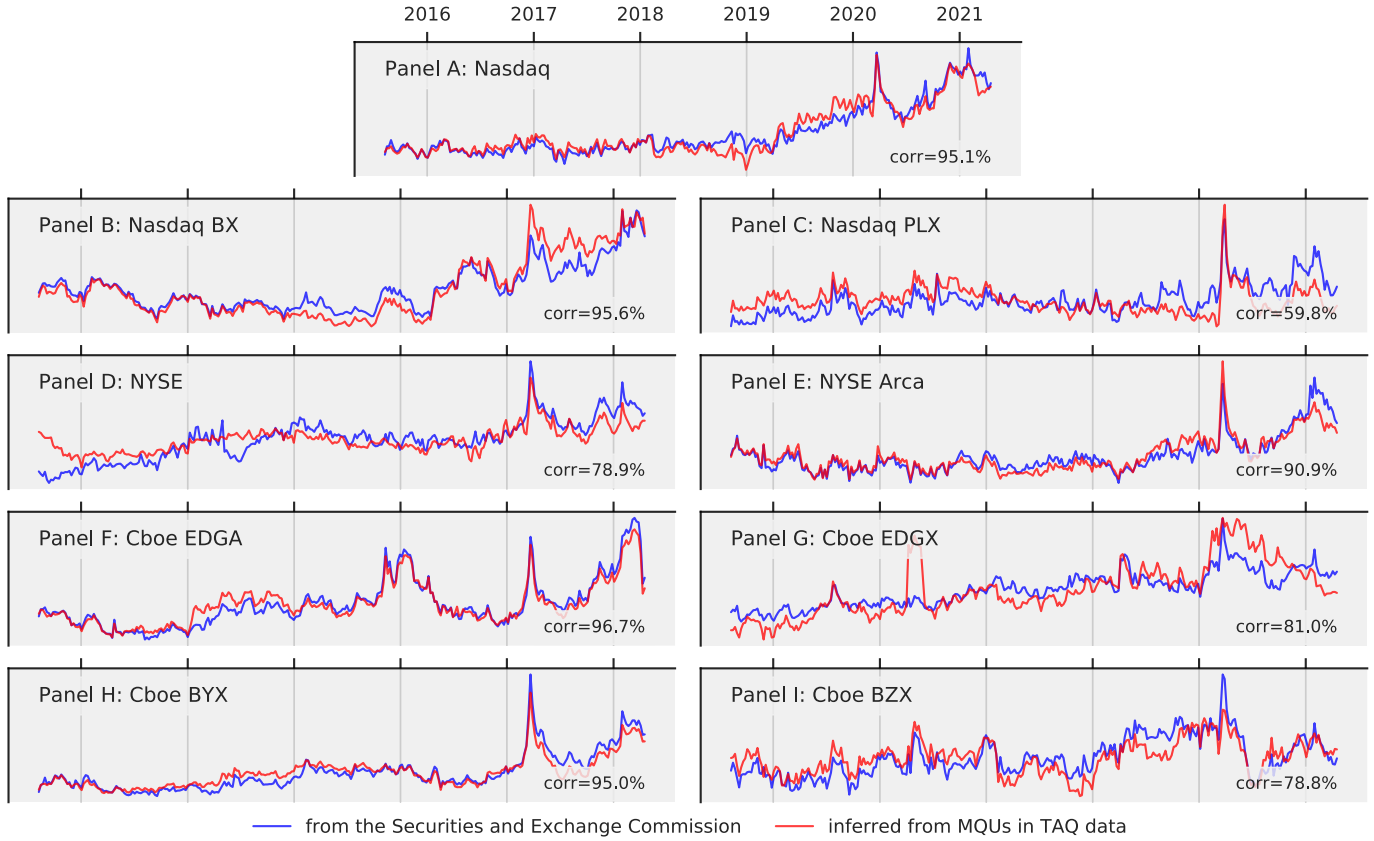
| Exchange | Time (SIP) | Price | Size | Participant Time | Mkt. Order # |
|---|---|---|---|---|---|
| Cboe BYX | 10:07:48.956663088 | 45.18 | 100 | 10:07:48.956447000 | 1 |
| NYSE Arca | 10:07:48.956761225 | 45.18 | 100 | **10:07:48.956381184** | **2 [200 shares]** |
| NYSE Arca | 10:07:48.956770900 | 45.18 | 100 | **10:07:48.956381184** | |
| NYSE | 10:07:48.956773554 | 45.18 | 300 | 10:07:48.956400128 | 3 |
| Cboe BZX | 10:07:48.956916983 | 45.18 | 100 | 10:07:48.956700000 | 4 |
| Nasdaq | 10:07:48.956971093 | 45.18 | 100 | 10:07:48.956953405 | 5 |
| FINRA | 10:07:48.957830128 | 45.18 | 66 | 10:07:48.956000000 | 6 |
| FINRA | 10:07:48.958053076 | 45.18 | 100 | 10:07:48.956000000 | 7 |
| FINRA | 10:07:48.963251524 | 45.18 | 100 | 10:07:48.956000000 | 8 |
| Nasdaq | 10:07:48.980768357 | 45.17 | 100 | **10:07:48.980744964** | **9 [1400 shares]** |
| Nasdaq | 10:07:48.980771589 | 45.17 | 100 | **10:07:48.980744964** | |
| Nasdaq | 10:07:48.980775480 | 45.17 | 100 | **10:07:48.980744964** | |
| Nasdaq | 10:07:48.980779225 | 45.17 | 200 | **10:07:48.980744964** | |
| Nasdaq | 10:07:48.980783822 | 45.17 | 3 | **10:07:48.980744964** | |
| Nasdaq | 10:07:48.980787586 | 45.17 | 82 | **10:07:48.980744964** | |
| Nasdaq | 10:07:48.980792109 | 45.17 | 815 | **10:07:48.980744964** | |
| Nasdaq BX | 10:07:48.980793655 | 45.17 | 3 | **10:07:48.980773076** | **10 [100 shares]** |
| Nasdaq BX | 10:07:48.980796505 | 45.17 | 82 | **10:07:48.980773076** | |
| Nasdaq BX | 10:07:48.980799256 | 45.17 | 15 | **10:07:48.980773076** | |
| Nasdaq PSX | 10:07:48.980824382 | 45.17 | 200 | 10:07:48.980802975 | 11 |

[1] Ait-Sahalia, Fan, Xue, and Zhou (2022): *How and When are High-Frequency Stock Returns Predictable?*

# Final Validation

## Replicating proportion of hidden executions

- TAQ data does not report odd-lot quotes
- When does a trade NOT update the top-of-book:
  - If it is too small to update round-lot liquidity
  - If it is executed against odd-lot liquidity
  - If it is executed against hidden liquidity
- We can infer SOME information on executions against hidden liquidity
- For each marketable order execution:
  - compare the [# quote updates] with [# of trades ≥ 100 shares]
- SEC publishes proportion of [# hidden executions] over [# total trades] per day × security × exchange
  - sourced from proprietary data of each exchange
- Compare TAQ-inferred with SEC data over entire sample period (every day, not just Wednesdays)



# How Widespread



Proportion of all trades in primary sample with quotes and/or other trades (same Participant timestamp on same day, same security, same exchange);
% of trades with at least one quote update; % of trades with at least one additional trade:

Panel A: Nasdaq

| #Quotes (MQUs) | # Trades | | | | |
|---|---|---|---|---|---|
| | 1 | 2 | 3 | ≥ 1 | ≥ 2 |
| 0 | 23.4 | 4.2 | 1.0 | 29.3 | 5.9 |
| 1 | 25.1 | 9.9 | 2.3 | 38.4 | 13.3 |
| 2 | 0.3 | 8.2 | 3.6 | 13.7 | 13.3 |
| 3 | 0.0 | 0.2 | 3.8 | 6.7 | 6.7 |
| ≥ 0 | 48.8 | 22.4 | 10.8 | (100) | **51.2** |
| ≥ 1 | 25.5 | 18.2 | 9.8 | **70.7** | 45.3 |
| ≥ 2 | 0.4 | 8.4 | 7.5 | 32.3 | 32.0 |

Panel B: Nasdaq BX

| #Quotes (MQUs) | # Trades | | | | |
|---|---|---|---|---|---|
| | 1 | 2 | 3 | ≥ 1 | ≥ 2 |
| 0 | 30.7 | 2.2 | 0.4 | 33.7 | 3.0 |
| 1 | 48.9 | 5.0 | 0.5 | 54.6 | 5.7 |
| 2 | 0.1 | 6.0 | 0.9 | 7.2 | 7.1 |
| 3 | 0.0 | 0.0 | 1.7 | 2.1 | 2.1 |
| ≥ 0 | 79.7 | 13.3 | 3.4 | (100) | **20.3** |
| ≥ 1 | 49.0 | 11.1 | 3.0 | **66.3** | 17.3 |
| ≥ 2 | 0.1 | 6.1 | 2.5 | 11.7 | 11.6 |

Panel C: Nasdaq PLX

| #Quotes (MQUs) | # Trades | | | | |
|---|---|---|---|---|---|
| | 1 | 2 | 3 | ≥ 1 | ≥ 2 |
| 0 | 18.1 | 1.6 | 0.4 | 20.5 | 2.4 |
| 1 | 54.5 | 3.7 | 0.4 | 58.8 | 4.3 |
| 2 | 0.4 | 9.9 | 1.1 | 11.7 | 11.3 |
| 3 | 0.0 | 0.1 | 3.4 | 4.1 | 4.1 |
| ≥ 0 | 73.0 | 15.4 | 5.4 | (100) | **27.1** |
| ≥ 1 | 54.9 | 13.8 | 5.0 | **79.5** | 24.6 |
| ≥ 2 | 0.4 | 10.1 | 4.6 | 20.7 | 20.3 |

Panel D: NYSE

| #Quotes (MQUs) | # Trades | | | | |
|---|---|---|---|---|---|
| | 1 | 2 | 3 | ≥ 1 | ≥ 2 |
| 0 | 39.9 | 0.5 | 0.1 | 40.6 | 0.7 |
| 1 | 57.7 | 1.4 | 0.2 | 59.4 | 1.7 |
| 2 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| 3 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| ≥ 0 | 97.7 | 2.0 | 0.2 | (100) | **2.3** |
| ≥ 1 | 57.7 | 1.4 | 0.2 | **59.4** | 1.7 |
| ≥ 2 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |

Panel E: NYSE Arca

| #Quotes (MQUs) | # Trades | | | | |
|---|---|---|---|---|---|
| | 1 | 2 | 3 | ≥ 1 | ≥ 2 |
| 0 | 23.9 | 4.1 | 0.9 | 29.5 | 5.6 |
| 1 | 32.8 | 16.1 | 7.5 | 70.4 | 37.7 |
| 2 | 0.0 | 0.0 | 0.0 | 0.1 | 0.1 |
| 3 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| ≥ 0 | 56.6 | 20.2 | 8.4 | (100) | **43.4** |
| ≥ 1 | 32.8 | 16.1 | 7.5 | **70.5** | 37.7 |
| ≥ 2 | 0.0 | 0.0 | 0.0 | 0.1 | 0.1 |

Panel F: Cboe EDGA

| #Quotes (MQUs) | # Trades | | | | |
|---|---|---|---|---|---|
| | 1 | 2 | 3 | ≥ 1 | ≥ 2 |
| 0 | 32.7 | 2.2 | 0.4 | 35.4 | 2.8 |
| 1 | 49.1 | 6.9 | 1.1 | 57.3 | 8.2 |
| 2 | 0.0 | 3.1 | 1.7 | 5.4 | 5.4 |
| 3 | 0.0 | 0.0 | 0.0 | 1.0 | 1.0 |
| ≥ 0 | 81.8 | 12.2 | 3.2 | (100) | **18.2** |
| ≥ 1 | 49.1 | 10.0 | 2.8 | **64.6** | 15.5 |
| ≥ 2 | 0.0 | 3.1 | 1.8 | 7.3 | 7.3 |

Panel G: Cboe EDGX

| #Quotes (MQUs) | # Trades | | | | |
|---|---|---|---|---|---|
| | 1 | 2 | 3 | ≥ 1 | ≥ 2 |
| 0 | 24.9 | 3.0 | 0.6 | 28.8 | 3.9 |
| 1 | 36.7 | 9.2 | 2.4 | 49.2 | 12.4 |
| 2 | 0.1 | 5.8 | 4.1 | 12.1 | 12.0 |
| 3 | 0.0 | 0.0 | 0.2 | 3.9 | 3.9 |
| ≥ 0 | 61.7 | 17.9 | 7.3 | (100) | **38.3** |
| ≥ 1 | 36.8 | 15.0 | 6.7 | **71.2** | 34.3 |
| ≥ 2 | 0.1 | 5.8 | 4.3 | 22.0 | 21.9 |

Panel H: Cboe BYX

| #Quotes (MQUs) | # Trades | | | | |
|---|---|---|---|---|---|
| | 1 | 2 | 3 | ≥ 1 | ≥ 2 |
| 0 | 29.6 | 2.5 | 0.4 | 32.8 | 3.2 |
| 1 | 43.9 | 8.6 | 1.7 | 54.6 | 10.7 |
| 2 | 0.0 | 4.2 | 2.8 | 8.2 | 8.2 |
| 3 | 0.0 | 0.0 | 0.1 | 2.1 | 2.1 |
| ≥ 0 | 73.6 | 15.3 | 5.0 | (100) | **26.4** |
| ≥ 1 | 43.9 | 12.8 | 4.5 | **67.2** | 23.2 |
| ≥ 2 | 0.0 | 4.2 | 2.9 | 12.6 | 12.5 |

Panel I: Cboe BZX

| #Quotes (MQUs) | # Trades | | | | |
|---|---|---|---|---|---|
| | 1 | 2 | 3 | ≥ 1 | ≥ 2 |
| 0 | 25.8 | 3.4 | 0.7 | 30.3 | 4.5 |
| 1 | 34.5 | 8.8 | 2.3 | 46.2 | 11.7 |
| 2 | 0.2 | 6.8 | 4.6 | 13.3 | 13.1 |
| 3 | 0.0 | 0.0 | 0.3 | 4.1 | 4.1 |
| ≥ 0 | 60.5 | 19.0 | 7.9 | (100) | **39.5** |
| ≥ 1 | 34.7 | 15.6 | 7.2 | **69.7** | 35.0 |
| ≥ 2 | 0.2 | 6.9 | 4.9 | 23.5 | 23.3 |

# Particularity of the NYSE

- In TAQ (not in proprietary data, anymore) NYSE reports trades in the perspective of marketable orders[1,2,3]
- The results reflect this exactly
- Does not imply that marketable order execs on NYSE cannot report two or more trades:
  - one trade is reported per price level the marketable order is executed against
  - the proportion of 2.3% coincides with the proportion of marketable orders executed at two or more prices

Excerpt from previous table:

Panel A: Nasdaq

| #Quotes (MQUs) | # Trades | | | | |
|---|---|---|---|---|---|
| | 1 | 2 | 3 | ≥ 1 | ≥ 2 |
| 0 | 23.4 | 4.2 | 1.0 | 29.3 | 5.9 |
| 1 | 25.1 | 9.9 | 2.3 | 38.4 | 13.3 |
| 2 | 0.3 | 8.2 | 3.6 | 13.7 | 13.3 |
| 3 | 0.0 | 0.2 | 3.8 | 6.7 | 6.7 |
| ≥ 0 | 48.8 | 22.4 | 10.8 | (100) | **51.2** |
| ≥ 1 | 25.5 | 18.2 | 9.8 | **70.7** | 45.3 |
| ≥ 2 | 0.4 | 8.4 | 7.5 | **32.3** | 32.0 |

Panel B: NYSE

| #Quotes (MQUs) | # Trades | | | | |
|---|---|---|---|---|---|
| | 1 | 2 | 3 | ≥ 1 | ≥ 2 |
| 0 | 39.9 | 0.5 | 0.1 | 40.6 | 0.7 |
| 1 | 57.7 | 1.4 | 0.2 | 59.4 | 1.7 |
| 2 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| 3 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| ≥ 0 | 97.7 | 2.0 | 0.2 | (100) | **2.3** |
| ≥ 1 | 57.7 | 1.4 | 0.2 | **59.4** | 1.7 |
| ≥ 2 | 0.0 | 0.0 | 0.0 | **0.0** | 0.0 |

[1] SEC (2014): *Order Book Reporting Methods and Their Impact on Some Market Activity Measures*, https://www.sec.gov/node/327371
[2] Upson, McInish, and Hardy Johnson IV (2021): *Order based versus level book trade reporting: An empirical analysis*
[3] https://www.reuters.com/article/us-interconti-exc-nyse-data-iex/nyse-plan-to-update-private-data-feed-draws-criticism-from-iex-idUSKCN0QP2CE20150820

# Fixing Sequence Errors (BBO)

- How to get correct prevailing quotes despite sequence errors – let's start at BBO
- Simple: entirely disregard SIP timestamp, accurate prevailing quote is last quote with earlier timestamp as trade
- Importantly: quotes with same participant timestamp as trades are NOT prevailing (before the execution of the marketable order)

- SIP BBO: prevailing quote at [SIP timestamp] - 1
- Direct BBO: prevailing quote at [Part timestamp] - 1
- Fix BBO:
  - SIP BBO if prevailing quote's Participant timestamp is earlier than trade's
  - Otherwise, Direct BBO

Prevailing quotes, agreement between proprietary round-lot BBO and TAQ:

| | | SIP BBO | | | Fix BBO | | |
|---|---|---|---|---|---|---|---|
| Ex/Group | # | Price | Depth | Both | Price | Depth | Both |
| Nasdaq | 19.3M | 61.5 | 53.2 | 41.0 | 98.6 | 96.2 | 95.6 |
| Nasdaq BX | 1.7M | 58.1 | 72.4 | 50.7 | 99.1 | 98.8 | 98.5 |
| Nasdaq PLX | .7M | 50.4 | 53.8 | 34.4 | 99.0 | 97.1 | 96.6 |
| NYSE | 6.3M | 78.8 | 77.4 | 69.6 | 98.9 | 99.1 | 98.4 |
| NYSE Arca | 7.2M | 67.6 | 65.3 | 56.0 | 99.6 | 99.3 | 99.1 |
| Equities | 30.2M | 66.1 | 64.1 | 51.6 | 98.9 | 98.0 | 97.4 |
| ETFs | 5.0M | 62.0 | 42.8 | 38.2 | 98.9 | 94.7 | 94.5 |
| Low MCap | 1.6M | 66.4 | 59.4 | 48.4 | 99.6 | 98.4 | 98.3 |
| Med MCap | 6.8M | 66.6 | 62.3 | 51.0 | 99.2 | 98.0 | 97.6 |
| High MCap | 26.8M | 65.2 | 61.0 | 49.5 | 98.8 | 97.4 | 96.8 |
| Low DVol | 1.1M | 68.8 | 68.6 | 55.3 | 99.6 | 99.3 | 99.0 |
| Med DVol | 5.5M | 67.3 | 66.6 | 53.8 | 99.1 | 98.5 | 98.0 |
| High DVol | 28.5M | 65.1 | 59.8 | 48.7 | 98.8 | 97.3 | 96.7 |
| Low Price | 1.4M | 67.1 | 45.6 | 39.8 | 99.6 | 97.0 | 97.0 |
| Med Price | 22.3M | 63.8 | 55.2 | 45.2 | 99.3 | 97.1 | 96.9 |
| High Price | 11.5M | 68.6 | 74.0 | 59.2 | 98.0 | 98.4 | 97.2 |
| All | 35.2M | 65.6 | 61.2 | **49.7** | 98.9 | 97.5 | **97.0** |

# Fixing Sequence Errors (BBO)

- How to get correct prevailing quotes despite sequence errors – let's start at BBO
- Simple: entirely disregard SIP timestamp, accurate prevailing quote is last quote with earlier timestamp as trade
- Importantly: quotes with same participant timestamp as trades are NOT prevailing (before the execution of the marketable order)

- SIP BBO: prevailing quote at [SIP timestamp] - 1
- Direct BBO: prevailing quote at [Part timestamp] - 1
- Fix BBO:
  - SIP BBO if prevailing quote's Participant timestamp is earlier than trade's
  - Otherwise, Direct BBO

Prevailing quotes, agreement between proprietary BBO (w/ odd-lot) and TAQ:

| | | SIP BBO | | | Fix BBO | | |
|---|---|---|---|---|---|---|---|
| Ex/Group | # | Price | Depth | Both | Price | Depth | Both |
| Nasdaq | 19.3M | 37.3 | 9.7 | 6.5 | 66.1 | 22.4 | 22.3 |
| Nasdaq BX | 1.7M | 37.2 | 43.2 | 25.4 | 75.9 | 63.0 | 62.9 |
| Nasdaq PLX | .7M | 36.1 | 31.0 | 15.6 | 81.2 | 61.6 | 61.5 |
| NYSE | 6.3M | 41.8 | 16.0 | 13.3 | 57.5 | 25.0 | 24.8 |
| NYSE Arca | 7.2M | 36.0 | 13.3 | 10.2 | 62.2 | 25.1 | 25.0 |
| Equities | 30.2M | 35.7 | 13.4 | 9.2 | 61.3 | 24.0 | 23.9 |
| ETFs | 5.0M | 50.8 | 15.7 | 12.8 | 85.1 | 40.6 | 40.5 |
| Low MCap | 1.6M | 44.3 | 18.2 | 13.0 | 72.8 | 33.5 | 33.4 |
| Med MCap | 6.8M | 39.0 | 15.5 | 11.0 | 65.7 | 28.5 | 28.4 |
| High MCap | 26.8M | 37.1 | 13.0 | 9.2 | 63.7 | 25.3 | 25.2 |
| Low DVol | 1.1M | 38.5 | 18.0 | 12.2 | 63.4 | 29.2 | 29.1 |
| Med DVol | 5.5M | 35.7 | 15.3 | 10.5 | 60.8 | 26.6 | 26.5 |
| High DVol | 28.5M | 38.2 | 13.3 | 9.4 | 65.3 | 26.1 | 26.0 |
| Low Price | 1.4M | 57.9 | 14.9 | 11.5 | 88.3 | 32.7 | 32.6 |
| Med Price | 22.3M | 43.1 | 15.0 | 10.7 | 73.8 | 29.8 | 29.8 |
| High Price | 11.5M | 25.6 | 11.4 | 7.6 | 44.6 | 19.0 | 18.8 |
| All | 35.2M | 37.8 | 13.8 | 9.7 | 64.5 | 26.3 | 26.2 |

- Note that even after fixing seq errs, TAQ is inferior due to prevalence of odd-lot liquidity at better prices; see high price group; also see [1]

[1] Bartlett, McCrary, and O'Hara (2022): *The Market Inside the Market: Odd-Lot Quotes*

# Fixing Sequence Errors (NBBO)

- To obtain prevailing NBBO, most commonly used method is Holden and Jacobsen (2014), updated in 2018[2]
  - SIP NBBO: method disregards participant timestamp and matches trades with the NBBO in force at [SIP timestamp] - 1
- Fix NBBO:
  - best prices and depth at best prices according to Fix BBO at each exchange
  - implicitly contains some latency
- Bartlett & McCrary (2019) use Direct NBBO:
  - NBBO in force at [Participant timestamp] - 1
  - assumes zero latency (not observed in real time)

Excerpt from SAS code from [2]:

Holden-and-Jacobsen-Daily-TAQ-Code-2018-03-16.sas

```
238
239  /* STEP 5: CLEAN DTAQ QUOTES DATA */
240
241  data quoteAB;
242
243      ...
244
245      drop Sym_Suffix Bidex Askex Qu_Cancel RPI SSR LULD_BBO_CQS
246           LULD_BBO_UTP FINRA_ADF_MPID SIP_Message_ID Part_Time RRN TRF_Time
247           Spread NATL_BBO_LULD;
248  run;
249
250  /* STEP 6: CLEAN DAILY TRADES DATA - DELETE ABNORMAL TRADES */
251
252  data trade2;
253      set DailyTrade;
254      where Tr_Corr eq '00' and price gt 0;
255      drop Tr_Corr Tr_Source TR_RF Part_Time RRN TRF_Time Sym_Suffix Tr_SCond
256           Tr_StopInd;
257  run;
258
```

Excerpt from SAS code from [2]:

Holden-and-Jacobsen-Daily-TAQ-Code-2018-03-16.sas

```
299
300  /* STEP 8: INTERLEAVE TRADES WITH NBBO QUOTES. DTAQ TRADES AT NANOSECOND
301     TMMMMMMMMM ARE MATCHED WITH THE DTAQ NBBO QUOTES STILL IN FORCE AT THE
302     NANOSECOND TMMMMMMMM(M-1) */;
303
304  data OfficialCompleteNBBO;
305      set OfficialCompleteNBBO;type='Q';
306      time_m=time_m+.000000001;
307      drop Qu_SeqNum;
308  run;
309
```

[1] Holden and Jacobsen (2014): *Liquidity Measurement Problems in Fast, Competitive Markets: Expensive and Cheap Solutions*
[2] SAS code available at https://host.kelley.iu.edu/cholden/
[3] Bartlett & McCrary (2019): *How Rigged Are Stock Markets? Evidence from Microsecond Timestamps*

# Quantifying Sequence Errors



- At BBO:
  - every third trade receives different prev quoted price
  - more than half of trades receive different prev quote
- At NBBO:
  - every fifth trade receives different prev quoted price
  - more than 2/3rd of trades receive different prev quote
- Note: the NBBO partly shields against price sequence errors when:
  - quotes are updated to inferior prices after a trade
  - are confused as prevailing quotes
  - but another exchange continues to hold the better price
  - (depths are impacted either way)
- Prevailing quoted depths are also important:
  - Hagströmer (2021) advocates the imbalance-weighted effective spread

Disagreement between Fix BBO/NBBO and SIP BBO/NBBO, proportion of all trades for which price/depth of prevailing quotes do not match:

| | | Fix BBO | | | Fix NBBO | | |
|---|---|---|---|---|---|---|---|
| Ex/Group | # | Price | Depth | Any | Price | Depth | Any |
| Nasdaq | 1421.2M | 33.7 | 44.6 | 54.5 | 30.9 | 67.5 | 71.3 |
| Nasdaq BX | 206.6M | 40.2 | 29.5 | 51.1 | 10.8 | 70.0 | 71.2 |
| Nasdaq PLX | 63.9M | 43.5 | 40.9 | 57.7 | 23.5 | 75.7 | 77.9 |
| NYSE | 510.1M | 30.8 | 37.9 | 47.6 | 13.0 | 48.7 | 51.3 |
| NYSE Arca | 634.7M | 31.5 | 34.4 | 43.0 | 19.7 | 58.1 | 60.5 |
| Cboe EDGA | 197.4M | 42.6 | 29.1 | 52.1 | 10.5 | 71.0 | 72.5 |
| Cboe EDGX | 456.1M | 37.1 | 44.7 | 56.2 | 22.6 | 73.8 | 76.0 |
| Cboe BYX | 318.5M | 40.5 | 41.0 | 55.8 | 7.5 | 73.1 | 74.1 |
| Cboe BZX | 601.9M | 37.6 | 44.0 | 55.8 | 24.0 | 71.8 | 74.2 |
| CTA | 2962.9M | 37.3 | 42.7 | 54.3 | 22.8 | 69.3 | 71.4 |
| UTP | 1447.4M | 31.2 | 36.2 | 48.2 | 19.6 | 59.4 | 63.0 |
| Equities | 3857.2M | 35.4 | 38.8 | 51.7 | 22.2 | 64.8 | 67.8 |
| ETFs | 553.1M | 34.8 | 53.0 | 56.9 | 18.6 | 74.3 | 74.8 |
| Low MCap | 201.4M | 34.8 | 40.2 | 52.7 | 20.7 | 63.2 | 66.1 |
| Med MCap | 822.3M | 35.6 | 39.6 | 52.3 | 22.4 | 64.7 | 67.8 |
| High MCap | 3386.6M | 35.3 | 40.8 | 52.3 | 21.7 | 66.5 | 69.0 |
| Low DVol | 110.9M | 33.8 | 31.8 | 47.9 | 21.7 | 54.5 | 59.1 |
| Med DVol | 615.8M | 35.5 | 36.1 | 50.7 | 23.5 | 61.7 | 65.5 |
| High DVol | 3683.6M | 35.3 | 41.6 | 52.7 | 21.5 | 67.1 | 69.5 |
| Low Price | 247.7M | 31.3 | 53.0 | 58.7 | 13.4 | 76.0 | 76.4 |
| Med Price | 3175.8M | 36.2 | 43.5 | 54.5 | 21.1 | 69.8 | 71.7 |
| High Price | 986.8M | 33.5 | 27.9 | 43.8 | 26.1 | 51.4 | 57.2 |
| All | 4410.3M | 35.3 | 40.6 | 52.3 | 21.8 | 66.0 | 68.7 |

[1] Hagströmer (2021): *Bias in the effective bid-ask spread*

# Impact of Sequence Errors



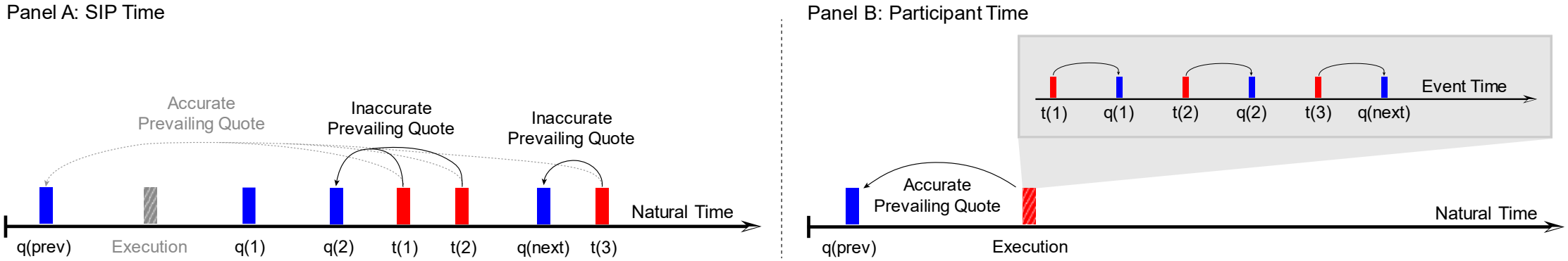
- ❑ Compute average percentage effective spread, price impact, realized spread, and imbalance-weighted effective spread[1]
- ❑ Obtain prevailing quotes with SIP NBBO and Fix NBBO (same trades, same future quotes for PI and RS)

| | | Price Impact (One Second) | | | | Price Impact (Five Minutes) | | | | Realized Spread (One Second) | | | | Realized Spread (Five Minutes) | | | | Effective Spread | | | | Effective Spread (IW) | | | | Avg % |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Ex/Group | # | SIP | Fix | % | +/- | SIP | Fix | % | +/- | SIP | Fix | % | +/- | SIP | Fix | % | +/- | SIP | Fix | % | +/- | SIP | Fix | % | +/- | (Excl. RS) |
| Nasdaq | 1421.2M | 2.21 | 2.91 | 32 | + | 2.68 | 3.40 | 27 | + | 0.12 | -0.25 | / | - | -0.36 | -0.73 | -104 | - | 2.29 | 2.65 | 15 | + | 1.63 | 1.96 | 20 | + | 24 |
| Nasdaq BX | 206.6M | 1.48 | 1.70 | 15 | + | 2.25 | 2.47 | 10 | + | 1.85 | 1.76 | -5 | - | 1.09 | 1.01 | -8 | - | 3.31 | 3.45 | 4 | + | 2.79 | 3.07 | 10 | + | 10 |
| Nasdaq PLX | 63.9M | 1.92 | 2.50 | 30 | + | 2.44 | 3.04 | 25 | + | 0.38 | 0.05 | -88 | - | -0.13 | -0.49 | -283 | - | 2.28 | 2.53 | 11 | + | 1.60 | 1.79 | 12 | + | 19 |
| NYSE | 510.1M | 2.21 | 2.49 | 13 | + | 2.88 | 3.15 | 9 | + | 0.22 | 0.08 | -64 | - | -0.45 | -0.58 | -30 | - | 2.40 | 2.54 | 6 | + | 1.64 | 1.80 | 10 | + | 10 |
| NYSE Arca | 634.7M | 2.23 | 2.76 | 23 | + | 2.76 | 3.31 | 20 | + | 0.00 | -0.29 | / | - | -0.52 | -0.84 | -62 | - | 2.21 | 2.44 | 10 | + | 1.52 | 1.69 | 11 | + | 16 |
| Cboe EDGA | 197.4M | 1.46 | 1.69 | 15 | + | 2.23 | 2.45 | 10 | + | 1.54 | 1.44 | -6 | - | 0.77 | 0.68 | -11 | - | 2.98 | 3.12 | 4 | + | 2.43 | 2.71 | 11 | + | 10 |
| Cboe EDGX | 456.1M | 2.27 | 2.88 | 26 | + | 2.87 | 3.51 | 22 | + | 0.06 | -0.27 | / | - | -0.54 | -0.90 | -68 | - | 2.31 | 2.58 | 12 | + | 1.58 | 1.91 | 21 | + | 20 |
| Cboe BYX | 318.5M | 1.41 | 1.61 | 14 | + | 2.36 | 2.55 | 8 | + | 1.94 | 1.86 | -4 | - | 1.00 | 0.93 | -6 | - | 3.33 | 3.46 | 4 | + | 2.67 | 2.93 | 10 | + | 9 |
| Cboe BZX | 601.9M | 2.06 | 2.62 | 28 | + | 2.45 | 3.03 | 24 | + | 0.10 | -0.20 | / | - | -0.29 | -0.60 | -107 | - | 2.13 | 2.40 | 13 | + | 1.45 | 1.74 | 20 | + | 21 |
| CTA | 2962.9M | 1.80 | 2.31 | 28 | + | 2.28 | 2.81 | 23 | + | 0.27 | -0.01 | / | - | -0.21 | -0.50 | -133 | - | 2.05 | 2.28 | 12 | + | 1.41 | 1.65 | 17 | + | 20 |
| UTP | 1447.4M | 2.61 | 3.15 | 21 | + | 3.36 | 3.90 | 16 | + | 0.62 | 0.37 | -41 | - | -0.11 | -0.38 | -232 | - | 3.20 | 3.49 | 9 | + | 2.45 | 2.78 | 14 | + | 15 |
| Equities | 3857.2M | 2.26 | 2.82 | 25 | + | 2.89 | 3.47 | 20 | + | 0.43 | 0.14 | -68 | - | -0.20 | -0.50 | -149 | - | 2.65 | 2.93 | 11 | + | 1.92 | 2.23 | 16 | + | 18 |
| ETFs | 553.1M | 0.76 | 0.96 | 26 | + | 0.88 | 1.07 | 22 | + | 0.09 | -0.04 | / | - | -0.03 | -0.15 | -341 | - | 0.83 | 0.91 | 9 | + | 0.54 | 0.58 | 7 | + | 16 |
| Low MCap | 201.4M | 6.86 | 8.65 | 26 | + | 9.76 | 11.66 | 19 | + | 3.37 | 2.45 | -27 | - | 0.50 | -0.50 | / | - | 10.13 | 11.03 | 9 | + | 7.63 | 8.65 | 13 | + | 17 |
| Med MCap | 822.3M | 3.53 | 4.45 | 26 | + | 4.62 | 5.57 | 20 | + | 0.77 | 0.29 | -62 | - | -0.32 | -0.81 | -154 | - | 4.25 | 4.71 | 11 | + | 3.05 | 3.55 | 16 | + | 18 |
| High MCap | 3386.6M | 1.43 | 1.77 | 24 | + | 1.73 | 2.08 | 20 | + | 0.11 | -0.07 | / | - | -0.19 | -0.37 | -97 | - | 1.53 | 1.69 | 11 | + | 1.08 | 1.26 | 16 | + | 18 |
| Low DVol | 110.9M | 6.98 | 8.97 | 28 | + | 10.47 | 12.56 | 20 | + | 5.74 | 4.79 | -17 | - | 2.32 | 1.30 | -44 | - | 12.62 | 13.70 | 9 | + | 9.87 | 11.08 | 12 | + | 17 |
| Med DVol | 615.8M | 3.91 | 4.97 | 27 | + | 5.09 | 6.17 | 21 | + | 1.05 | 0.50 | -52 | - | -0.13 | -0.68 | -425 | - | 4.91 | 5.44 | 11 | + | 3.52 | 4.09 | 16 | + | 19 |
| High DVol | 3683.6M | 1.61 | 1.99 | 24 | + | 1.99 | 2.38 | 20 | + | 0.11 | -0.09 | / | - | -0.27 | -0.48 | -79 | - | 1.70 | 1.89 | 11 | + | 1.21 | 1.40 | 16 | + | 18 |
| Low Price | 247.7M | 6.95 | 8.56 | 23 | + | 9.42 | 11.11 | 18 | + | 1.69 | 0.77 | -55 | - | -0.78 | -1.77 | -128 | - | 8.55 | 9.26 | 8 | + | 5.39 | 6.17 | 14 | + | 16 |
| Med Price | 3175.8M | 1.95 | 2.47 | 26 | + | 2.43 | 2.96 | 22 | + | 0.29 | 0.02 | -95 | - | -0.19 | -0.47 | -149 | - | 2.22 | 2.47 | 11 | + | 1.58 | 1.85 | 17 | + | 19 |
| High Price | 986.8M | 1.21 | 1.46 | 21 | + | 1.60 | 1.84 | 15 | + | 0.37 | 0.27 | -28 | - | -0.01 | -0.11 | -669 | - | 1.57 | 1.71 | 9 | + | 1.38 | 1.54 | 11 | + | 14 |
| All | 4410.3M | 2.07 | 2.59 | 25 | + | 2.64 | 3.17 | 20 | + | 0.38 | 0.11 | -70 | - | -0.18 | -0.46 | -154 | - | 2.43 | 2.68 | 10 | + | 1.75 | 2.02 | 16 | + | 18 |

[1] Hagströmer (2021): *Bias in the effective bid-ask spread*

# Impact of Sequence Errors

## Biased prevailing quoted prices (and depths)




- Example: marketable order executed against three resting limit orders, each time generating a trade and a top-of-book update
- Difference between q(prev) and q(1,2,next) follows simple rules
- A marketable buy (sell) order can only immediately increase (decrease) bid and asking prices by:
  - taking liquidity at the ask (bid)
  - providing liquidity at the bid (ask) when a part of the order is booked

- Bias in midquote, contingent on trade sign Q:

$$[M(1,2,next) \geq M(prev)]_{|Q=1}$$

$$[M(1,2,next) \leq M(prev)]_{|Q=-1}$$

- Downward bias in effective spread:

$$ES \equiv Q(P-M)$$

$$Q(P-M(1,2,next)) \leq Q(P-M(prev))$$

- Downward bias in price impact:

$$PI \equiv Q(M^{+}-M)$$

$$Q(M^{+}-M(1,2,next)) \leq Q(M^{+}-M(prev))$$

- Biased realized spread; bias in imbalance-weighted effective spread[1]

[1] Hagströmer (2021): *Bias in the effective bid-ask spread*

# Impact of Sequence Errors

## Trade signing accuracy

- We do not observe true trade signs in primary sample
- Use average percentage price impact as proxy for trade signing accuracy:
  - LR applied to prevailing quotes obtained in various ways
  - prevailing quotes that are used to sign trades are varied
  - same trades, same (accurate) prevailing midquotes, same future midquotes
  - due to Conrad and Wahal (2020): short horizon

Proprietary ITCH/IF data, accuracy of LR with given method:

| | | BBO | | NBBO | | | RNBBO | |
|---|---|---|---|---|---|---|---|---|
| Ex/Group | # | SIP BBO | Fix BBO | SIP NBBO | Direct NBBO | Fix NBBO | Direct RNBBO | SIP RNBBO |
| Nasdaq | 19.3M | 77.3 | 90.7 | 78.9 | 91.2 | 91.7 | 92.5 | 93.0 |
| Nasdaq BX | 1.7M | 68.0 | 89.9 | 94.9 | 97.8 | 97.9 | 98.1 | 97.9 |
| Nasdaq PLX | .7M | 72.5 | 92.9 | 80.0 | 91.6 | 92.2 | 93.2 | 93.9 |
| NYSE | 6.3M | 82.5 | 88.5 | 87.9 | 87.9 | 90.3 | 90.4 | 91.2 |
| NYSE Arca | 7.2M | 72.6 | 85.8 | 83.9 | 87.9 | 90.3 | 91.1 | 91.8 |
| All | 35.2M | 76.7 | 89.3 | 82.4 | 90.2 | 91.5 | 92.1 | 92.7 |

Proprietary ITCH/IF data, average percentage price impact proxy for accuracy:

| | | BBO | | NBBO | | | RNBBO | | True |
|---|---|---|---|---|---|---|---|---|---|
| Ex/Group | # | SIP BBO | Fix BBO | SIP NBBO | Direct NBBO | Fix NBBO | Direct RNBBO | SIP RNBBO | |
| Nasdaq | 19.3M | 1.71 | 2.61 | 2.00 | 2.75 | 2.81 | 2.81 | 2.80 | **2.88** |
| Nasdaq BX | 1.7M | 0.46 | 1.30 | 1.39 | 1.60 | 1.62 | 1.62 | 1.61 | **1.62** |
| Nasdaq PLX | .7M | 0.88 | 2.22 | 1.47 | 2.30 | 2.37 | 2.39 | 2.41 | **2.45** |
| NYSE | 6.3M | 1.95 | 2.29 | 2.33 | 2.32 | 2.49 | 2.49 | 2.50 | **2.54** |
| NYSE Arca | 7.2M | 1.14 | 1.97 | 1.98 | 2.30 | 2.46 | 2.47 | 2.47 | **2.50** |
| All | 35.2M | 1.56 | 2.35 | 2.01 | 2.52 | 2.61 | 2.62 | 2.62 | **2.67** |

TSPP Appendix B.II data, accuracy of LR with given method:

| | | BBO | | NBBO | | | RNBBO | |
|---|---|---|---|---|---|---|---|---|
| Ex/Group | # | SIP BBO | Fix BBO | SIP NBBO | Direct NBBO | Fix NBBO | Direct RNBBO | SIP RNBBO |
| Nasdaq | 66.4M | 81.7 | 95.1 | 88.8 | 96.1 | 96.9 | 96.9 | 97.9 |
| Nasdaq BX | 19.6M | 73.1 | 91.5 | 95.8 | 97.9 | 98.3 | 98.3 | 98.9 |
| Nasdaq PLX | 2.3M | 66.5 | 92.0 | 88.0 | 95.4 | 96.5 | 96.8 | 97.7 |
| NYSE | 12.3M | 83.3 | 94.5 | 90.6 | 94.2 | 95.7 | 95.6 | 97.9 |
| NYSE Arca | 22.1M | 83.5 | 93.0 | 87.7 | 93.4 | 94.3 | 95.9 | 96.5 |
| Cboe EDGA | 4.9M | 70.3 | 89.8 | 92.4 | 96.1 | 96.5 | 97.3 | 98.4 |
| Cboe EDGX | 15.3M | 65.9 | 90.3 | 86.3 | 95.2 | 95.6 | 96.7 | 97.7 |
| Cboe BYX | 20.8M | 77.8 | 92.7 | 95.4 | 97.6 | 97.8 | 98.2 | 98.6 |
| Cboe BZX | 15.9M | 71.8 | 91.8 | 87.0 | 94.7 | 95.2 | 96.2 | 96.8 |
| All | 179.6M | 77.9 | 93.2 | 90.0 | 95.8 | 96.5 | 96.9 | 97.8 |

TSPP Appendix B.II data, average percentage price impact proxy for accuracy:

| | | BBO | | NBBO | | | RNBBO | | True |
|---|---|---|---|---|---|---|---|---|---|
| Ex/Group | # | SIP BBO | Fix BBO | SIP NBBO | Direct NBBO | Fix NBBO | Direct RNBBO | SIP RNBBO | |
| Nasdaq | 66.4M | 4.53 | 6.87 | 5.88 | 7.02 | 7.22 | 7.16 | 7.22 | **7.35** |
| Nasdaq BX | 19.6M | 2.57 | 4.54 | 4.69 | 5.14 | 5.24 | 5.21 | 5.24 | **5.27** |
| Nasdaq PLX | 2.3M | 2.42 | 6.65 | 5.96 | 7.20 | 7.52 | 7.46 | 7.57 | **7.74** |
| NYSE | 12.3M | 3.53 | 5.38 | 4.74 | 5.34 | 5.65 | 5.60 | 5.71 | **5.80** |
| NYSE Arca | 22.1M | 4.45 | 6.16 | 5.15 | 6.29 | 6.50 | 6.61 | 6.62 | **6.76** |
| Cboe EDGA | 4.9M | 1.89 | 4.83 | 4.69 | 5.59 | 5.67 | 5.70 | 5.74 | **5.81** |
| Cboe EDGX | 15.3M | 1.99 | 6.82 | 5.65 | 7.65 | 7.75 | 7.81 | 7.85 | **8.01** |
| Cboe BYX | 20.8M | 2.52 | 4.65 | 4.56 | 5.18 | 5.23 | 5.24 | 5.24 | **5.30** |
| Cboe BZX | 15.9M | 2.98 | 6.62 | 5.51 | 7.11 | 7.23 | 7.26 | 7.26 | **7.48** |
| All | 179.6M | 3.55 | 6.09 | 5.34 | 6.42 | 6.58 | 6.58 | 6.62 | **6.74** |

[1] Conrad and Wahal (2020): *The term structure of liquidity provision*

# Impact of Sequence Errors
## Trade signing accuracy



Primary sample of TAQ data, average percentage price impact proxy for accuracy:

| | | BBO | | NBBO | | | RNBBO | |
|---|---|---|---|---|---|---|---|---|
| Ex/Group | # | SIP BBO | Fix BBO | SIP NBBO | Direct NBBO | Fix NBBO | Direct RNBBO | SIP RNBBO |
| Nasdaq | 1421.2M | 1.98 | 2.77 | 2.22 | 2.88 | 2.96 | 2.94 | 2.94 |
| Nasdaq BX | 206.6M | 0.63 | 1.30 | 1.51 | 1.67 | 1.71 | 1.69 | 1.70 |
| Nasdaq PLX | 63.9M | 1.01 | 2.28 | 1.84 | 2.42 | 2.53 | 2.52 | 2.53 |
| NYSE | 510.1M | 1.70 | 2.29 | 2.20 | 2.27 | 2.47 | 2.42 | 2.48 |
| NYSE Arca | 634.7M | 1.56 | 2.51 | 2.23 | 2.61 | 2.79 | 2.78 | 2.79 |
| Cboe EDGA | 197.4M | 0.53 | 1.29 | 1.50 | 1.66 | 1.70 | 1.68 | 1.68 |
| Cboe EDGX | 456.1M | 1.41 | 2.57 | 2.24 | 2.82 | 2.89 | 2.88 | 2.87 |
| Cboe BYX | 318.5M | 0.72 | 1.30 | 1.45 | 1.59 | 1.62 | 1.61 | 1.60 |
| Cboe BZX | 601.9M | 1.46 | 2.40 | 2.06 | 2.59 | 2.66 | 2.65 | 2.64 |
| Equities | 3857.2M | 1.65 | 2.56 | 2.27 | 2.74 | 2.85 | 2.83 | 2.83 |
| ETFs | 553.1M | 0.62 | 0.95 | 0.72 | 0.90 | 0.97 | 0.94 | 0.97 |
| Low MCap | 201.4M | 4.96 | 8.07 | 6.95 | 8.57 | 8.84 | 8.83 | 8.81 |
| Med MCap | 822.3M | 2.53 | 4.03 | 3.52 | 4.32 | 4.49 | 4.47 | 4.47 |
| High MCap | 3386.6M | 1.08 | 1.61 | 1.43 | 1.71 | 1.78 | 1.77 | 1.78 |
| Low DVol | 110.9M | 4.70 | 8.13 | 7.16 | 8.90 | 9.17 | 9.16 | 9.13 |
| Med DVol | 615.8M | 2.76 | 4.49 | 3.93 | 4.85 | 5.04 | 5.03 | 5.03 |
| High DVol | 3683.6M | 1.22 | 1.83 | 1.61 | 1.93 | 2.01 | 1.99 | 2.00 |
| Low Price | 247.7M | 5.63 | 8.37 | 6.93 | 8.48 | 8.78 | 8.78 | 8.82 |
| Med Price | 3175.8M | 1.43 | 2.26 | 1.95 | 2.39 | 2.49 | 2.47 | 2.48 |
| High Price | 986.8M | 0.81 | 1.18 | 1.23 | 1.41 | 1.46 | 1.44 | 1.44 |
| All | 4410.3M | 1.52 | 2.36 | 2.07 | 2.51 | 2.61 | 2.59 | 2.60 |

Proprietary ITCH/IF data, average percentage price impact proxy for accuracy:

| | | BBO | | NBBO | | | RNBBO | |
|---|---|---|---|---|---|---|---|---|
| Ex/Group | # | SIP BBO | Fix BBO | SIP NBBO | Direct NBBO | Fix NBBO | Direct RNBBO | SIP RNBBO |
| Nasdaq | 19.3M | 1.71 | 2.61 | 2.00 | 2.75 | 2.81 | 2.81 | 2.80 |
| Nasdaq BX | 1.7M | 0.46 | 1.30 | 1.39 | 1.60 | 1.62 | 1.62 | 1.61 |
| Nasdaq PLX | .7M | 0.88 | 2.22 | 1.47 | 2.30 | 2.37 | 2.39 | 2.41 |
| NYSE | 6.3M | 1.95 | 2.29 | 2.33 | 2.32 | 2.49 | 2.49 | 2.50 |
| NYSE Arca | 7.2M | 1.14 | 1.97 | 1.98 | 2.30 | 2.46 | 2.47 | 2.47 |
| All | 35.2M | 1.56 | 2.35 | 2.01 | 2.52 | 2.61 | 2.62 | 2.62 |

TSPP Appendix B.II data, average percentage price impact proxy for accuracy:

| | | BBO | | NBBO | | | RNBBO | |
|---|---|---|---|---|---|---|---|---|
| Ex/Group | # | SIP BBO | Fix BBO | SIP NBBO | Direct NBBO | Fix NBBO | Direct RNBBO | SIP RNBBO |
| Nasdaq | 66.4M | 4.53 | 6.87 | 5.88 | 7.02 | 7.22 | 7.16 | 7.22 |
| Nasdaq BX | 19.6M | 2.57 | 4.54 | 4.69 | 5.14 | 5.24 | 5.21 | 5.24 |
| Nasdaq PLX | 2.3M | 2.42 | 6.65 | 5.96 | 7.20 | 7.52 | 7.46 | 7.57 |
| NYSE | 12.3M | 3.53 | 5.38 | 4.74 | 5.34 | 5.65 | 5.60 | 5.71 |
| NYSE Arca | 22.1M | 4.45 | 6.16 | 5.15 | 6.29 | 6.50 | 6.61 | 6.62 |
| Cboe EDGA | 4.9M | 1.89 | 4.83 | 4.69 | 5.59 | 5.67 | 5.70 | 5.74 |
| Cboe EDGX | 15.3M | 1.99 | 6.82 | 5.65 | 7.65 | 7.75 | 7.81 | 7.85 |
| Cboe BYX | 20.8M | 2.52 | 4.65 | 4.56 | 5.18 | 5.23 | 5.24 | 5.24 |
| Cboe BZX | 15.9M | 2.98 | 6.62 | 5.51 | 7.11 | 7.23 | 7.26 | 7.26 |
| All | 179.6M | 3.55 | 6.09 | 5.34 | 6.42 | 6.58 | 6.58 | 6.62 |

# Conclusion



- The Participant timestamp is relevant beyond its intended application of gauging dissemination latency
- Prevailing quotes from SIP timestamps are impacted: prices (and depths) are biased – as a result:
  - central measures are biased
  - trade signing accuracy is impacted negatively
- Presented issues are by no means exhaustive:
  - as a ground rule, one should carefully consider whether the SIP timestamp is a good choice; usually the Participant timestamp should be preferred
  - other issues covered in the paper: larger marketable orders are impacted more by sequence errors; differences between NYSE and Nasdaq complicate comparisons; significant variation in prevalence of errors: across time, characteristics such as stock price, and exchanges
  - event time is particularly impacted
  - replacing SIP by Participant timestamp may not suffice, consolidating partly executions may be required
- TAQ has issues, but the Participant timestamp can improve the data immensely:
  - information on marketable order execs: the SEC releases daily – but not intra-daily – data on hidden executions
  - separate mechanic from organic quote revisions, gauge immediate (mechanic) price impacts
  - other creative ways to use this insight may exist…