

ORF 474: High Frequency Trading
Spring 2020
Robert Almgren

Lecture 1b

Feb 5, 2020

Today

- Issues in exchange design
- Data pipelines
- Beginning data analysis

Exchange design

- minimum price increment (tick size)
- trade hours
- matching algorithm (FIFO, pro rata, auction, ...)
- opening and closing rules

Tick size

- Equities: heavily regulated
 - US:
 - April 2001 decimalization (max tick \$0.01)
 - 2005 Reg NMS (min tick \$0.01)
 - subsequently: tick size pilot programs
 - Japan and Europe: complicated schemes
- Treasuries: competitive exchanges
- Futures: decision of exchange

US stocks: decimalization April 2001



[Home](#) | [Previous Page](#)

U.S. Securities and Exchange Commission

Testimony Concerning Decimal Pricing in the Securities and Options Markets

By Chairman Arthur Levitt
U.S. Securities & Exchange Commission

Subcommittee On Finance and Hazardous Materials
Committee on Commerce
United States House of Representatives

June 13, 2000

I. Executive Summary

The convention of quoting stock prices in fractions dates back more than two hundred years. Currently, the United States securities markets are the only major markets not to price stocks in decimals. As the securities markets become more global, with many stocks traded in multiple jurisdictions, the U.S. securities markets need to adopt the international convention of decimal pricing to remain competitive. And the overall benefits of decimal pricing are likely to be significant. Investors may benefit from lower transaction costs due to narrower spreads.¹ Moreover, the markets will be easier to understand for the average investor, who is used to dealing in dollars and cents for every-day transactions. It is time for the U.S. securities markets to make this change.

<https://www.sec.gov/news/testimony/ts092000.htm>

Reg NMS (2005): no subpenny quoting

Tick size cannot be less than \$0.01



[Home](#) | [Previous Page](#)

U.S. Securities and Exchange Commission

Division of Market Regulation: Responses to Frequently Asked Questions Concerning Rule 612 (Minimum Pricing Increment) of Regulation NMS

I. Introduction

On April 6, 2005, the Commission adopted Regulation NMS, a series of initiatives designed to modernize and strengthen the national market system for equity securities. Regulation NMS was published in Securities Exchange Act Release No. 51808 (Jun. 9, 2005), 70 FR 37496 (Jun. 29, 2005) ("NMS Release"). One part of Regulation NMS is Rule 612 ("the Rule"), which specifies minimum pricing increments for NMS stocks. In general, the Rule prohibits market participants from displaying, ranking, or accepting quotations, orders, or indications of interest in any NMS stock priced in an increment smaller than \$0.01 if the quotation, order, or indication of interest is priced equal to or greater than \$1.00 per share. If the quotation, order, or indication of interest is priced less than \$1.00 per share, the minimum pricing increment is \$0.0001. The Commission has extended the initial compliance date for the Rule from August 29, 2005 until January 31, 2006 in Securities Exchange Act Release No. 52196 (Aug. 2, 2005), 70 FR 45529 (Aug. 8, 2005).

- [Question 1:](#) Orders Explicitly Priced in Sub-Penny Increments
- [Question 2:](#) Orders Not Explicitly Priced in Sub-Penny Increments
- [Question 3:](#) Orders That Must Be Repriced Due to a Corporate Action
- [Question 4:](#) Orders with Performance Targets
- [Question 5:](#) Orders Above \$1.00 Placed When the Last Sale Is Below \$1.00
- [Question 6:](#) No Requirement to Accept Sub-Penny Orders Below \$1.00
- [Question 7:](#) Rounding Orders Below \$1.00 to Conform to Exchange Rules
- [Question 8:](#) Rounding Orders Above \$1.00 to Conform to Rule 612
- [Question 9:](#) Representing a Sub-Penny Order When Some Markets Won't Accept It
- [Question 10:](#) Considering Sub-Penny Quotations in a Best Execution Analysis
- [Question 11:](#) Sub-Penny Quotations Under the ITS Plan
- [Question 12:](#) Rounding Sub-Penny Quotations for Display
- [Question 13:](#) Sub-Penny Price Improvement
- [Question 14:](#) Warrants and Rights May Be NMS Stocks
- [Question 15:](#) Applying the "Madoff Exemption" for Short Sales in a Sub-Penny Market
- [Question 16:](#) Quoting Increments for Non-NMS Stocks
- [Question 17:](#) Using a Service Bureau to Accept Orders

<https://www.sec.gov/divisions/marketreg/subpenny612faq.htm>

Ongoing regulatory review



U.S. SECURITIES AND
EXCHANGE
COMMISSION

Search SEC.gov

COMPANY FILINGS | MORE SEARCH
OPTIONS

ABOUT | DIVISIONS | ENFORCEMENT | REGULATION | EDUCATION | FILINGS | NEWS

Trading and
Markets

About

Banks and
Other
Depository
Institutions

Clearing
Agencies

Exchanges

Market
Structure

ECNs/Alternative
Trading
Systems

National Market
System

Security-Based
Swap Data
Repositories

Contact Us

Tick Size Pilot Program



What is it?

The Tick Size Pilot Program is a national market system (NMS) plan designed to allow the Commission, market participants, and the public to study and assess the impact of wider minimum quoting and trading increments – or tick sizes – on the liquidity and trading of the common stocks of certain small-capitalization companies. The Tick Size Pilot began in October 2016 and will be in effect for two years.

What is included in the study?

Securities that are included in the Tick Size Pilot are NMS common stocks that have a market capitalization of \$3.0 billion or less, a closing price of at least \$2.00, and a consolidated average daily volume of one million shares or less ("Pilot Securities"). The lists of Pilot Securities are available on the websites of the listing exchanges and FINRA.

The Pilot Securities have been divided into one control group and three test groups. Each test group contains approximately 400 Pilot Securities and the remaining Pilot Securities are in the control group. The groups are defined

Public Comments

[Submit a Comment](#)

[Comment Letters on File](#)

Tick Pilot Documents

- [Order Directing the Exchanges and FINRA to Submit a Tick Size Pilot Plan](#)
- [Plan to Implement Tick Size Pilot Program](#)
- [Order approving Tick Size Pilot Plan](#)
- [Exemption requests and Exemption Orders & Letters](#)

<https://www.sec.gov/ticksizepilot>

Tick size set by exchange

FEDERAL RESERVE BANK *of NEW YORK* *Serving the Second District and the Nation*



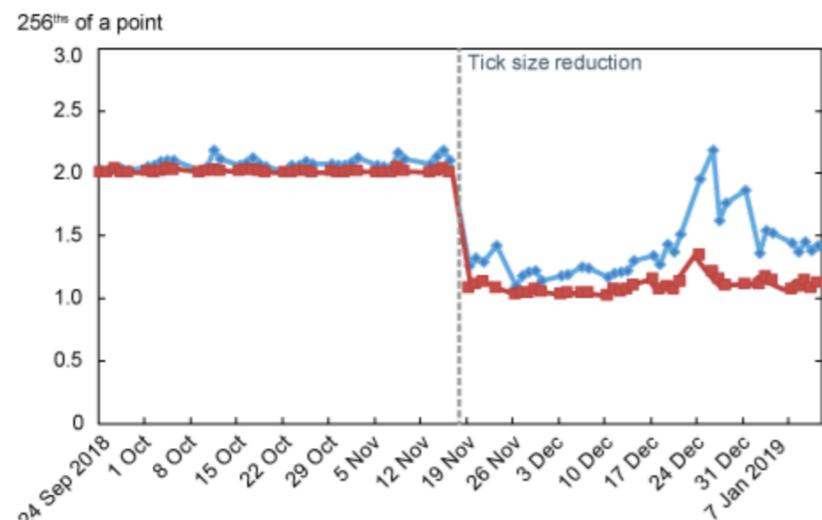
Liberty Street Economics

JANUARY 15, 2020

How Does Tick Size Affect Treasury Market Quality?

Michael Fleming, Giang Nguyen, and Francisco Ruela

The popularity of U.S. Treasury securities as a means of pricing other securities, managing interest rate risk, and storing value is, in part, due to the efficiency and liquidity of the U.S. Treasury market. Any structural changes that might affect these attributes of the market are therefore of interest to market participants and policymakers alike. In this post, we consider how a 2018 change in the minimum price increment, or **tick size**, for the 2-year U.S. Treasury note affected market quality, following our recently updated New York Fed [staff report](#).



<https://libertystreeteconomics.newyorkfed.org/2020/01/how-does-tick-size-affect-treasury-market-quality.html>

Fractional Pricing

Unlike U.S. equity markets, which switched to decimal pricing in 2001, U.S. Treasury securities still trade in fractions. In particular, prices are quoted in 32nds of a point, where a point equals one percent of par, with the 32nds themselves split into fractions. On the BrokerTec platform, for example, 3- and 5-year notes trade in quarters of 32nds, whereas 7-, 10-, and 30-year securities trade in halves of 32nds. The quoted price for a 5-year note might be 98-15¼, for example, indicating a price in decimal form of 98.4765625 (that is, $98 + 15\frac{1}{32} + \frac{1}{4}\frac{1}{32}$).

The Tick Size Change

Effective November 19, 2018, BrokerTec halved the tick size in the 2-year note from $\frac{1}{4}$ to $\frac{1}{8}$ of a 32nd. This was the first tick size change since the platform's inception in 2000. Interestingly, the tick size on the 2-year note futures contract was also halved to $\frac{1}{8}$ of a 32nd, albeit eight weeks after the BrokerTec change, on January 14, 2019. (Note that BrokerTec is owned by the CME Group, which operates the Treasury futures market.) Tick sizes of other securities remained unchanged, allowing us to use those securities as controls in our analysis.

US Treasuries

FENICS

Help Library > Setup > Exchanges: Americas > Fenics

DESCRIPTION	-
Fenics Overview	
Fenics supported order types	

FENICS OVERVIEW

Fenics U.S. Treasuries (“Fenics UST”) is a fully electronic U.S Government securities (“USTs”) trading venue owned and operated by BGC Financial, L.P. (“BGC”). Fenics UST operates a Central Limit Order Book (“CLOB”). Subscribers to the platform are banks, broker-dealers, professional trading firms (PTFs) and institutional investors.

The CLOB will compete with well-entrenched incumbents in a marketplace designed for low-latency hedgers and arbitrageurs utilizing the tightest tick size in the U.S. Treasury Market, 1/16th of 1/32nd.

Fenics UST is the fastest, most stable, and most reliable ultra-low latency matching engine in the cash UST marketplace. All modern forms of electronic connectivity are supported, including binary protocol, FIX protocol and high-speed internet access with secured VPN.

Other US Treasury electronic markets are
1/2 or 1/4 of 1/32nd

Available Products

TT supports trading of all On-The-Run U.S. Treasury instruments will trade in increments of 1/16th of 1/32nd.

Trading hours

Opening/closing hours Gaps in trading

Choose trading hours to
Aggregate liquidity
Avoid information event disruptions

Trading hours (equities)

The New York Times

‘Why Does the Stock Market Have Pre-Market and After-Hours Trading?’

By NEIL IRWIN JAN. 18, 2017

A reader asks: “Why does the stock market have pre-market and after-hours trading? If trades can occur outside the 9:30 a.m. to 4 p.m. market hours, then why not extend the trading hours altogether, or allow for 24/7 trading?”

In the old days, showing up at the same time was more obviously necessary — the traders at the [New York Stock Exchange](#) needed to be physically in the same place to buy and sell shares. But even now that most trades happen electronically, it benefits everyone if the exchange sets certain hours as the business day, so that buyers and sellers know exactly when everyone else will be ready to buy and sell. (For the New York Stock Exchange, those hours are 9:30 a.m. to 4 p.m., Monday through Friday.)

In response to new technologies and increased demands (particularly global demands), the stock market began offering extended hours that now allow you to trade shares as early as 4 a.m. and as late as 6:30 p.m. — but there are fewer buyers and sellers at those times. Most traders are busy having dinner with their spouses or asleep or whatever. So even small orders can distort the price; trying to sell just a few thousand shares of a stock might make its price plummet in after-hours trading, whereas, during the trading day, a similar order might find a buyer without affecting the price much at all.

Business

Traders in World's Longest Working Day Plead for Shorter Hours

By Ksenia Galouchko, William Canny, and Jan-Patrick Barnert

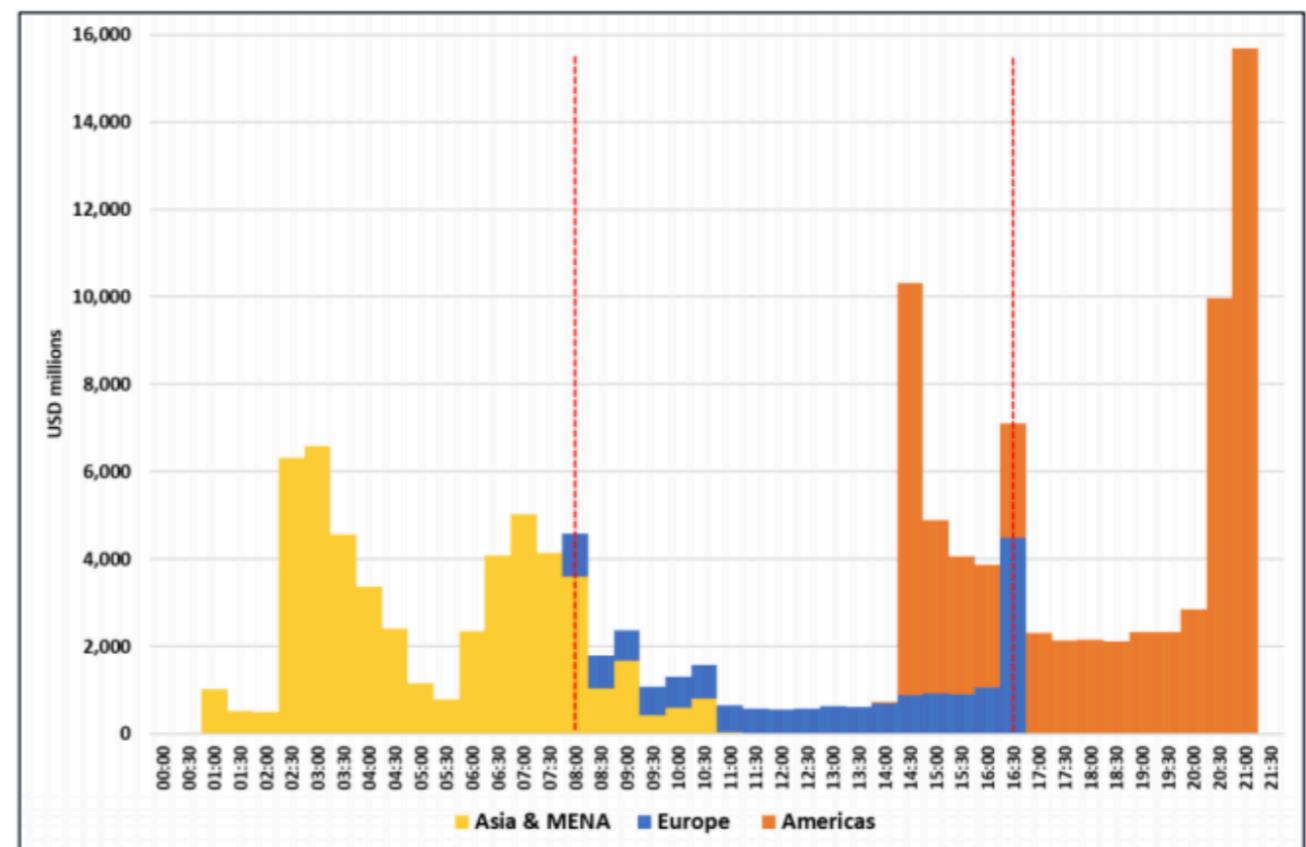
February 5, 2020, 9:37 AM UTC

Across Europe, 74% of finance professionals, including traders, analysts and fund managers, said they favor reducing the current 8 1/2-hour stock market day -- which exceeds the U.S. by two hours.

The results of the survey lend weight to a proposal by two professional bodies to shorten market hours for reasons including the concentration of liquidity in the first and last hours of trading, the short time window between corporate news releases and the market open, and the need to improve the work-life balance and diversity at financial firms. The Association for Financial Markets in Europe (AFME) and the Investment Association of U.K. asset managers asked bourses in November to consider cutting 90 minutes from their trading day.

"Making trading hours shorter will improve market efficiency," said Matthew McLoughlin, a London-based partner and head of trading at Liontrust Investment Partners LLP and a member of the Investment Association's buy-side trading committee. "There's no benefit in spreading liquidity out over a longer period than necessary."

Figure 3:



Above figure represents liquidity across global exchanges, as per September 2019 daily average value traded (continuous + auctions) across Equities. Exchanges include: Asia & MENA (ASX, HKSE, NSE, QSE, SGX, SSE, TSE), Europe (AEX, ISE, MAD, MTA, PAR, LSE, XETRA) and Americas (NASDAQ, NYSE, TSX)

Source: Bloomberg

<https://www.bloomberg.com/news/articles/2020-02-05/traders-in-world-s-longest-working-day-plead-for-shorter-hours>



Extended trading hours. Extended opportunities.

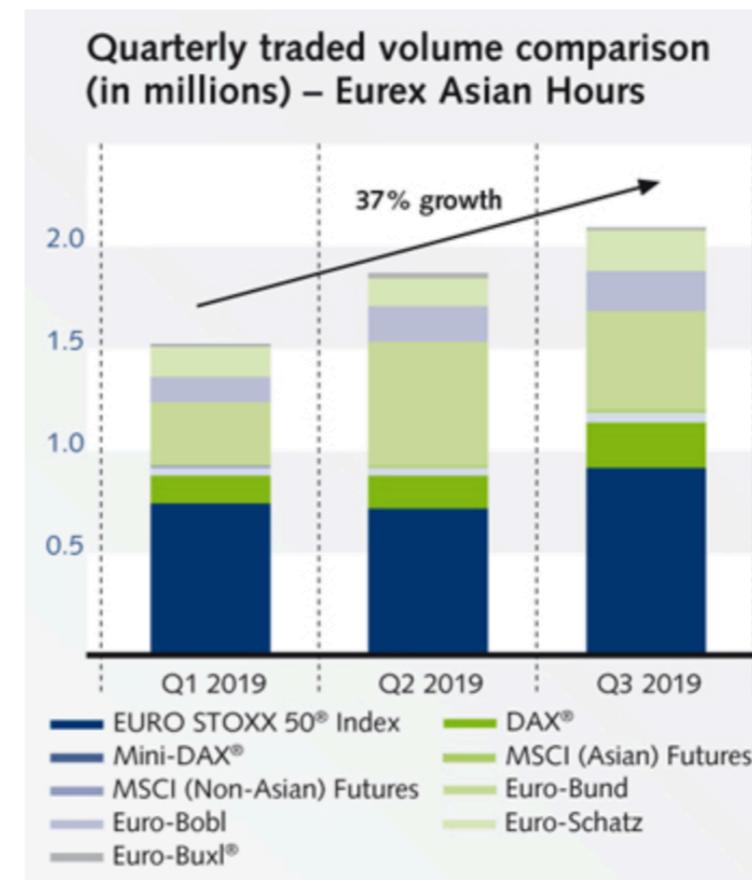
Eurex offered extended trading hours for select liquid futures to support the increasing demand of Eurex clients for trading on a global market place.

Launched on 10 December 2018, more than 7.7 million contracts traded since.

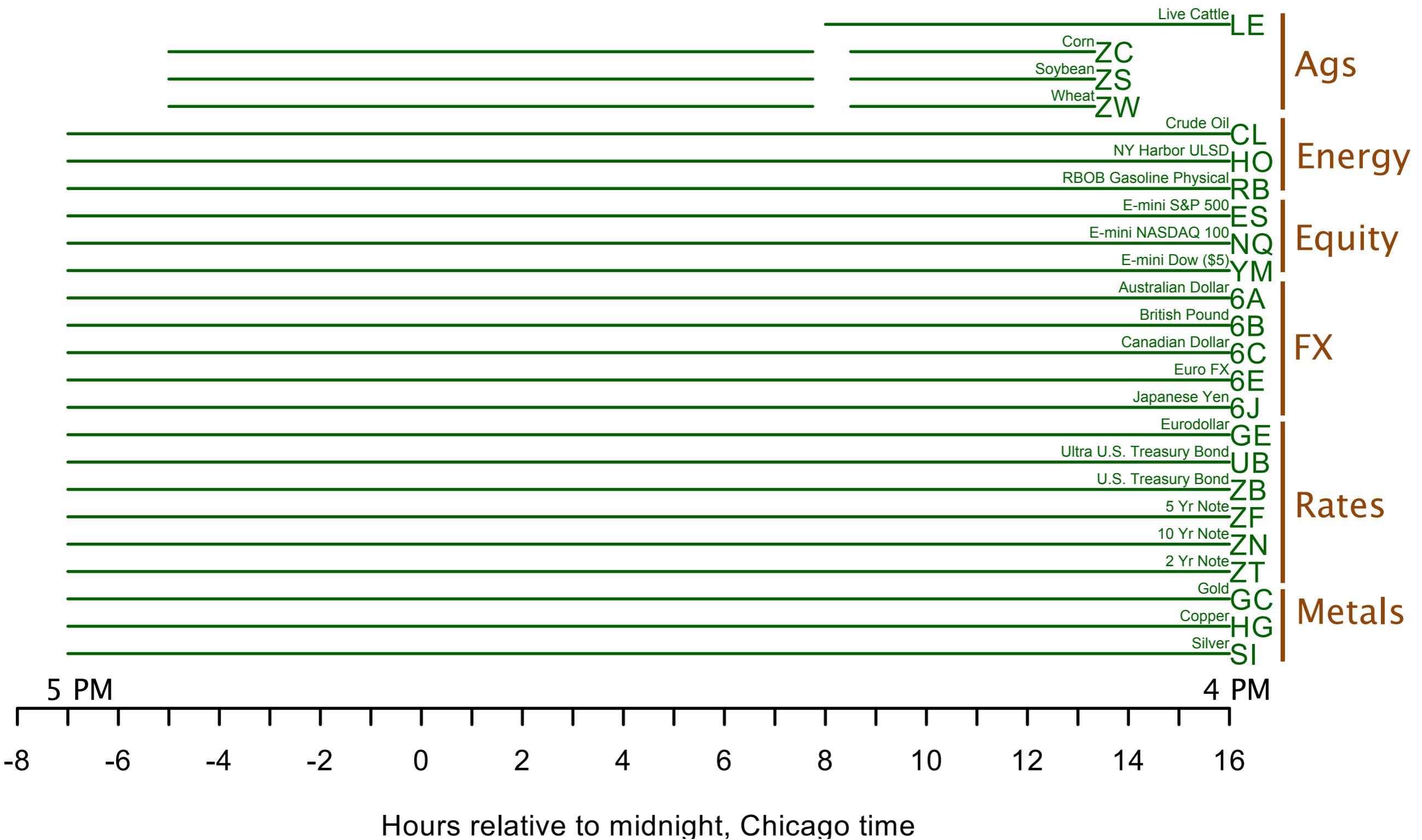
Key highlights

- **Extended opportunities during the Asian hours**
Trading starts from
8am (Singapore / Hong Kong Time)
9am (South Korea / Japan Time)
10am (Sydney / Melbourne Time)
- Spread trading opportunities:
 - Euro-Bund vs JGB
 - Euro-Bund / Euro-Schatz vs ASX 10s and ASX 3s
- **Major trading, clearing and risk management functionalities are available** to all Eurex Clearing/Non-Clearing members during the extended hours without additional requirements
- **Healthy distribution of activities**
 - More than 230 active members
 - Attractive spread - 1 tick (EURO STOXX 50 and Euro-Bund, Euro-Bobl, Euro-Schatz)

Volume development



Trading hours (CME futures)



Trading gaps and information

THE WALL STREET JOURNAL.

CME Group Cuts Grain, Soy Futures Trading to 17½ Hours

By IAN BERRY And JACOB BUNGE

Updated March 5, 2013 12:41 p.m. ET

CHICAGO—[CME Group](#) Inc. will reduce its trading day in grain and soybean futures to 17½ hours from 21 hours, largely undoing a controversial expansion done last year.

The new schedule trims afternoon hours and creates a 45-minute pause in trading in the morning before open-outcry, or "pit," trading begins, exchange officials said Tuesday.

The new hours come in response to criticism that the trading schedule had become bloated, increasing costs for grains traders and creating periods of low-volume, highly volatile trade.

Chicago-based CME's struggle over how to manage one of its oldest markets reflects the delicate balance the world's largest futures exchange operator must strike between keeping key customers happy and fending off competitors in Europe and the U.S.

While longer trading sessions meant less time for big grain merchandisers like [Bunge](#) Ltd. and Cargill Inc. to square their books between sessions, some smaller grain elevators had embraced a trading day that more closely tracked their grain-buying schedule, according to Mike Hall, president of brokerage MLH Futures Inc., based in Litchfield, IL.

Danny Murphy, president of the American Soybean Association, a trade group representing soybean farmers, said his members' top concern remains the release of key crop reports from the U.S. Department of Agriculture during trading hours. Before the lengthier sessions introduced by CME and ICE, brokers and grain companies had hours to parse the reports before trading began. Since the change, some have complained that releasing the reports amid live trading has triggered wild price swings.

"We still would like to see a pause during trading," said Mr. Murphy, who farms soybeans, corn and wheat near Canton, Miss. CME officials have said they would consider a pause if ICE implements one as well, while ICE has said that it makes more sense to have markets open for those who want to place trades immediately, or look to the futures market for a reference price.



Trading through a crop report

Fake Solar Generators Generated Fake Tax Breaks

Also latency arbitrage, asbestos and CLO connoisseurship.

By Matt Levine

January 28, 2020, 5:26 PM UTC

“Latency arbitrage”

A simple case is, you know, a company's stock is trading at \$100 per share, and at 2 p.m. one Tuesday the company puts out a press release saying “actually our company is great now and the stock should be trading at \$110.” And the press release is obviously true and trustworthy and it is instantly clear that the stock should really be trading at \$110. They found oil under their office or whatever. Obviously it doesn't work that way, but it's a straightforward toy story.

One question you might ask is: Is this good or bad or what? There are some arguments that it is good for the people who notice the press release first to make money, and for the people who don't notice it to lose money. One argument is sort of a moral argument. It's a free country: If I notice something first, why shouldn't I be able to trade on it ahead of people who notice it later? The people who trade first have in some sense proven their moral superiority; they are first because they read faster or understand more clearly or have a faster horse or have a faster computer or trained their faster computer to read press releases faster or whatever. Effort and attention and skill are rewarded; sloth and ignorance are punished; it is standard stuff.

<https://www.bloomberg.com/opinion/articles/2020-01-28/fake-solar-generators-generated-fake-tax-breaks>
<https://www.bloomberg.com/opinion/authors/ARbTQIRLRjE/matthew-s-levine>

Here is an occasional paper from the U.K. Financial Conduct Authority by Matteo Aquilina, Eric Budish and Peter O'Neill titled “Quantifying the High-Frequency Trading ‘Arms Race’: A Simple New Methodology and Estimates.” It is a good and interesting paper that can tell you some interesting things. The interesting things that it tells you are: How frequently do traders race to incorporate new information into stock prices, how fast are those races, and how much money is at stake?

Anyway this is all interesting but of course instead of finding it interesting you could find it outrageous. That's sort of the traditional approach to anything having to do with market structure and high-speed trading. It just seems to outrage everyone, I don't know. If you dislike high-frequency trading, you will find this tax outrageous and unacceptable. If you dislike criticism of high-frequency trading, you will find this characterization as a “tax” outrageous and unacceptable. Everyone who has read this far presumably strongly holds one of those views and is therefore either mad about this paper or mad at this paper. [6]

My guiding philosophy around here is that I am always going to favor anything that both (1) has a plausible economic basis and (2) promotes laziness. If the stock exchange was open only one hour per day, say in two half-hour sessions, almost all news would occur outside of market hours. Everyone would have plenty of time to reflect on the news at their leisure, and to prepare for the opening auction, and also generally to go home and think about other stuff. Then there'd be a brief trading period where high-frequency traders would claw at each other mercilessly, but it would be over quickly and everyone could go back to not picking each other off and spending time with their families. Market makers would have less exposure to being picked off—just because they'd be exposed for less time—and could charge less. Competition to make prices efficient would be intense for an hour, and chill the rest of the time. Of course you couldn't trade the rest of the time, and the great liquidity for 1 hour a day would be balanced by no liquidity the rest of the time, but you'd get used to it. Why were you trading so much anyway? Wasn't it exhausting?



Changes to Department of Labor Media Lockup Effective March 1, 2020

Last Modified Date: January 21, 2020

The [Office of Inspector General \(OIG\) report 17-14-001-03-315](#), dated July 14, 2014, states that the lock-up “unintentionally creates an unfair competitive advantage for certain news organizations and their clients”:

Pre-release access of DOL-generated economic data is intended to serve the general public by ensuring that news reports about the data are accurate. To that end, the media are given access to the data in advance of the public release to facilitate their ability to analyze and ask questions about the data as they prepare their news stories. However, the intended purpose of ensuring accurate news reports must be weighed against the inequitable trading advantage that a lock-up can potentially create. Several news organizations that participate in the DOL press lock-up are able to profit from their presence in the lock-up by selling, to traders, high speed data feeds of economic data formatted for computerized algorithmic trading. Because these news organizations have pre-release access, they are able to pre-load the data ... allowing their clients to get this information faster than the general public, which has to wait to download the data after it gets posted to the Department of Labor websites.

Moreover, even if ETA and OPA were able to post the data on their websites precisely at 8:30 AM, the approved news organizations that are in the press lock-up, and their clients, would still have a competitive trading advantage because the news organizations in the lock-up are able to pre-load the data so it is ready to be transmitted to their clients precisely at 8:30 AM. In contrast, those outside the lock-up would first have to download the data at 8:30 AM before they could use it. Even if this “latency period” only took fractions of seconds, the trading advantage could be quite significant, as noted above.⁴ In addition to the trading advantage, approved news organizations that are in the press lock-up have a competitive advantage over individuals and other organizations not permitted in the lock-up because they are in position to market their data services for profit.

In light of the OIG recommendations, the following lock-up procedures will go into effect on March 1, 2020.

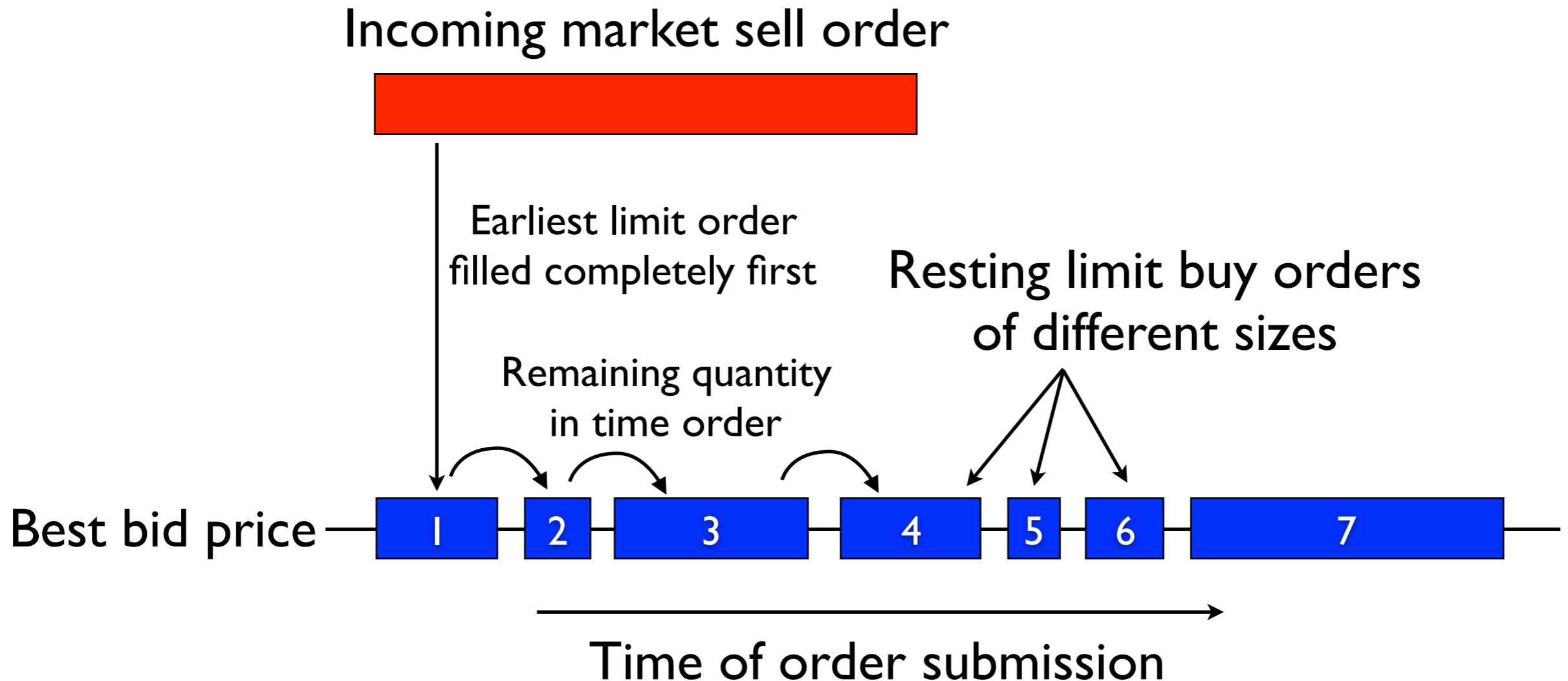
- BLS and/or ETA staff will be present in the lock-up rooms, will provide paper copies of releases and related material, and will add a new pen and pad briefing attributable to a BLS or ETA source.
- At release times, BLS and ETA will continue to distribute official news releases through various dissemination methods, such as the DOL or BLS websites, Twitter feeds, and e-mail subscription lists.

<https://www.bls.gov/bls/changes-to-dol-media-lockup-effective-march-1-2020.htm>

Matching algorithms

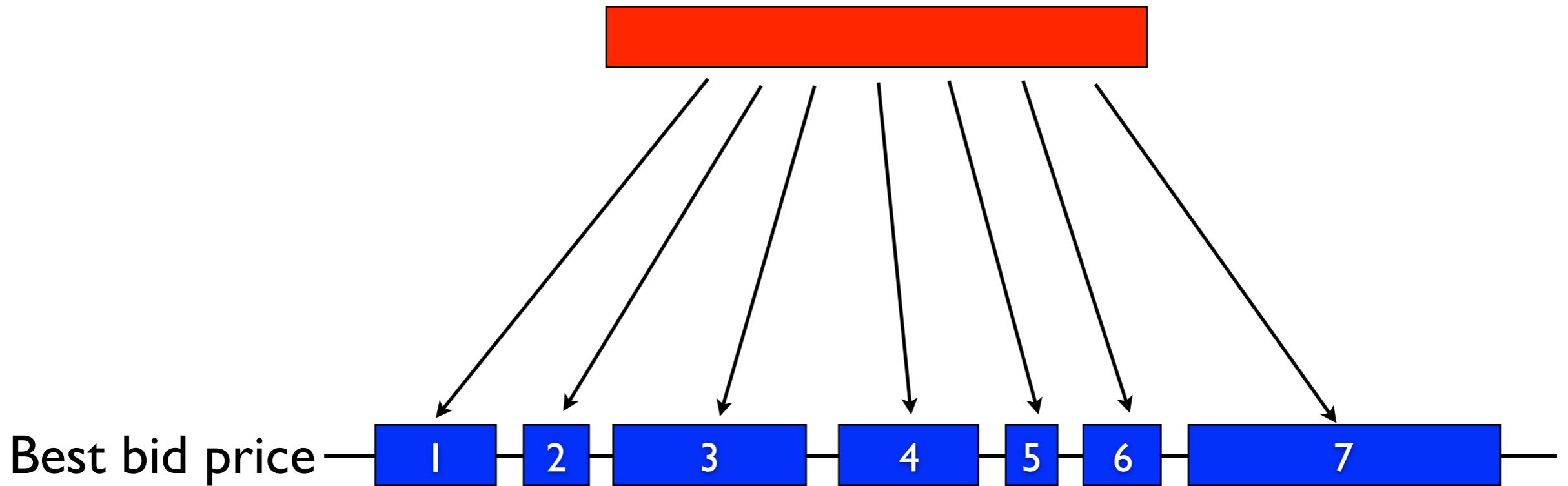
- How market orders are matched to limit
- Algorithm fixed by exchange
 - to attract more volume
 - to attract correct mix of participants
 - etc

First-in first-out (FIFO) order matching



Pro rata order matching

Incoming market sell order



Incoming volume divided
among *all* resting orders
at best price

Matching Algorithms

This topic explains the different matching algorithms supported on the CME Globex Platform. These matching algorithms support order management functionalities offered to market participants and ensure that each market participant is given the best possible execution at the fairest price.

A matching algorithm is a technique to allocate matched quantities, used when an [aggressor order](#) matches with one or multiple resting orders. Algorithms apply to both outright and [implied matching](#).

Order matching follows three steps:

1. Determine the current prices opposite of the aggressor order, outright and implied prices
2. Determine quantity at best price opposite the aggressor order, based on business rule priorities
3. Allocate resting quantity at best price to trade with aggressor order using the market algorithm

The following are the supported matching algorithms:

Algorithm	tag 1142-MatchAlgorithm Value
Allocation	A
Eurodollar Option	Y
FIFO	F
FIFO with LMM	T
FIFO with Top Order and LMM	S
Pro-Rata	C
Split FIFO and Pro-Rata	K
Threshold Pro-Rata	O
Threshold Pro-Rata with LMM	Q

9 different ways to match market and limit orders, on CME alone

"The FIFO algorithm uses price and time as the only criteria for filling an order. In this algorithm, all orders at the same price level are filled according to time priority; the first order at a price level is the first order matched."

Opening and closing auctions

THE WALL STREET JOURNAL.

Home World

WSJ. Magazine



MARKETS



What's the Biggest Trade on the New York Stock Exchange? The Last One

The NYSE operates between 9:30 a.m. and 4 p.m., but much of the action has moved to the final moments, thanks to index funds and others that flock to the day's closing auction

By [Corrie Driebusch](#), [Alexander Osipovich](#) and [Gregory Zuckerman](#)

March 14, 2018 10:38 a.m. ET

Most Popular Videos

Competition between exchanges

THE WALL STREET JOURNAL.

English Edition ▾ | January 28, 2020 | Print Edition | Video

MARKETS

SEC Decision on 4 p.m. Closing Trades Deals Blow to NYSE, Nasdaq

Regulator approves Cboe end-of-day auction proposal despite opposition from two largest exchanges

By [Alexander Osipovich](#)

Updated Jan. 21, 2020 8:41 pm ET

 SAVE  PRINT  TEXT

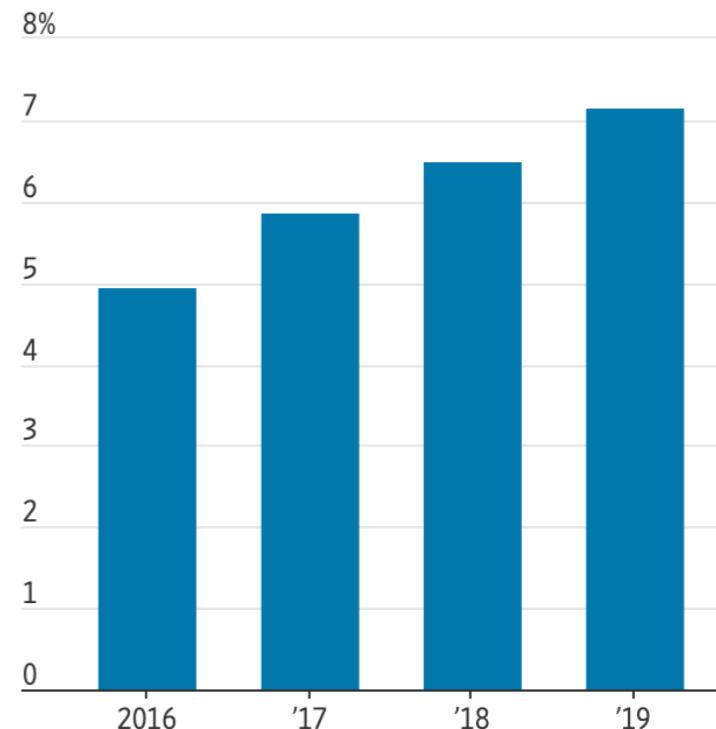
Regulators have dealt a blow to the New York Stock Exchange and [Nasdaq](#) Inc.

[NDAQ +0.81% ▲](#) by loosening their grip on the 4 p.m. closing auctions that determine end-of-day prices for thousands of stocks.

The Securities and Exchange Commission said Tuesday that it would allow a rival exchange group, [Cboe Global Markets](#) Inc., [CBOE +1.34% ▲](#) to implement a mechanism that lets traders bypass paying fees to NYSE and Nasdaq when they seek to buy or sell shares at daily closing prices.

The move could hurt the revenue of NYSE, owned by Intercontinental Exchange Inc., and Nasdaq at a time closing auctions are playing an increasingly important role in the U.S. stock market.

Percentage of U.S. stock-trading volume executed in exchanges' closing auctions

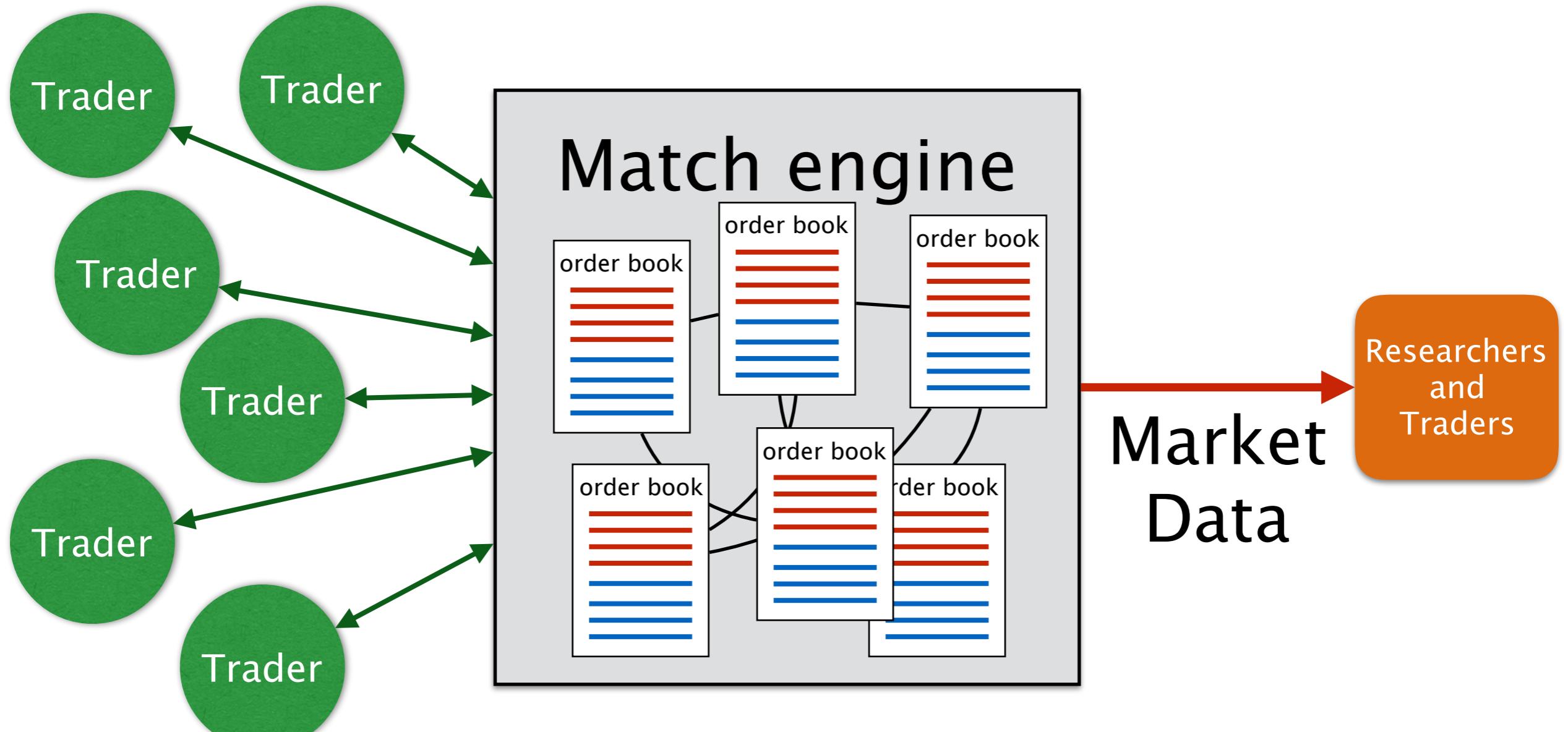


Source: Cboe Global Markets

Data for this course

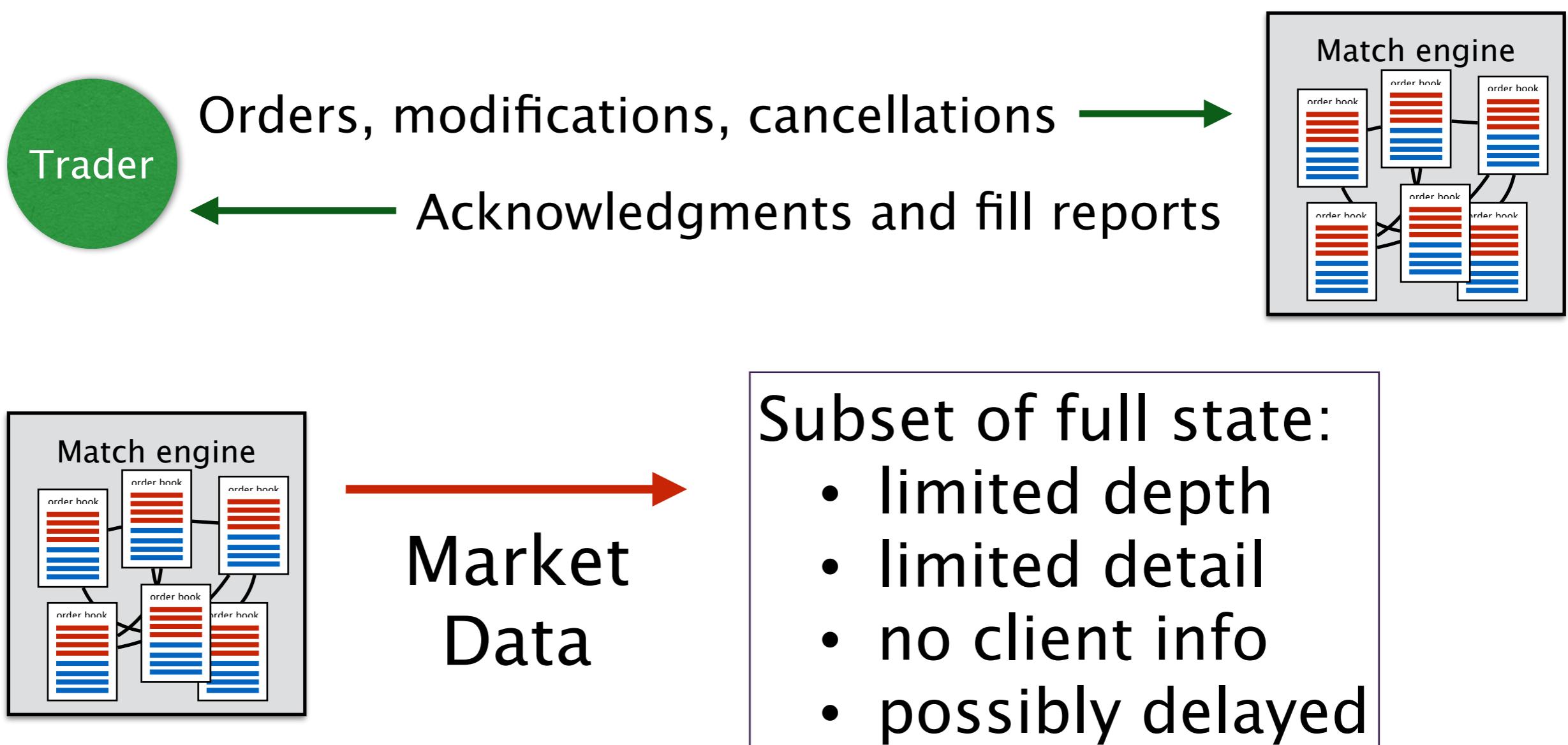
- Equities: TAQ (Trade-and-Quote)
Accessible through WRDS Web interface
Source: SIP consolidated data feed
Advantages: equities are familiar, interface is easy
Disadvantages: fragmentation, inadequate software
- Futures: CME
Accessible through course data server
Source: direct market data feed
Advantages: very clean data, good software tools
Disadvantages: need to learn Q/Kdb+ languages

Exchange structure (single exchange)

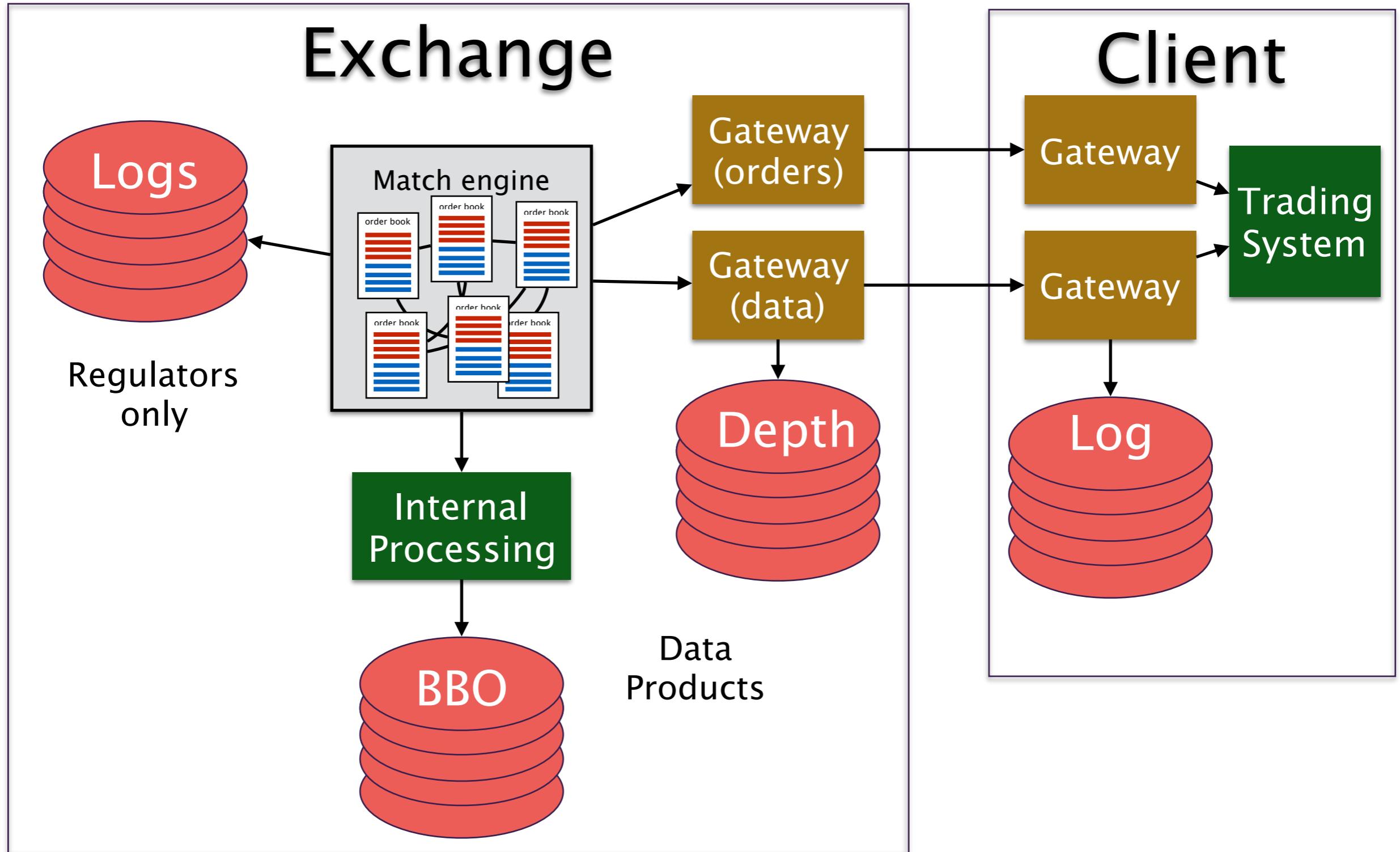


Want to use market data
to understand what is
happening inside match engines

Data content



Data recording from futures exchange



Consequences of single exchange recording

- Can be very clean because single pathway
- Data for research is not necessarily trading data
- Internal data has client identification
not available to external researchers

What is possible with client-labelled data

The Flash Crash: The Impact of High Frequency Trading on an Electronic Market*

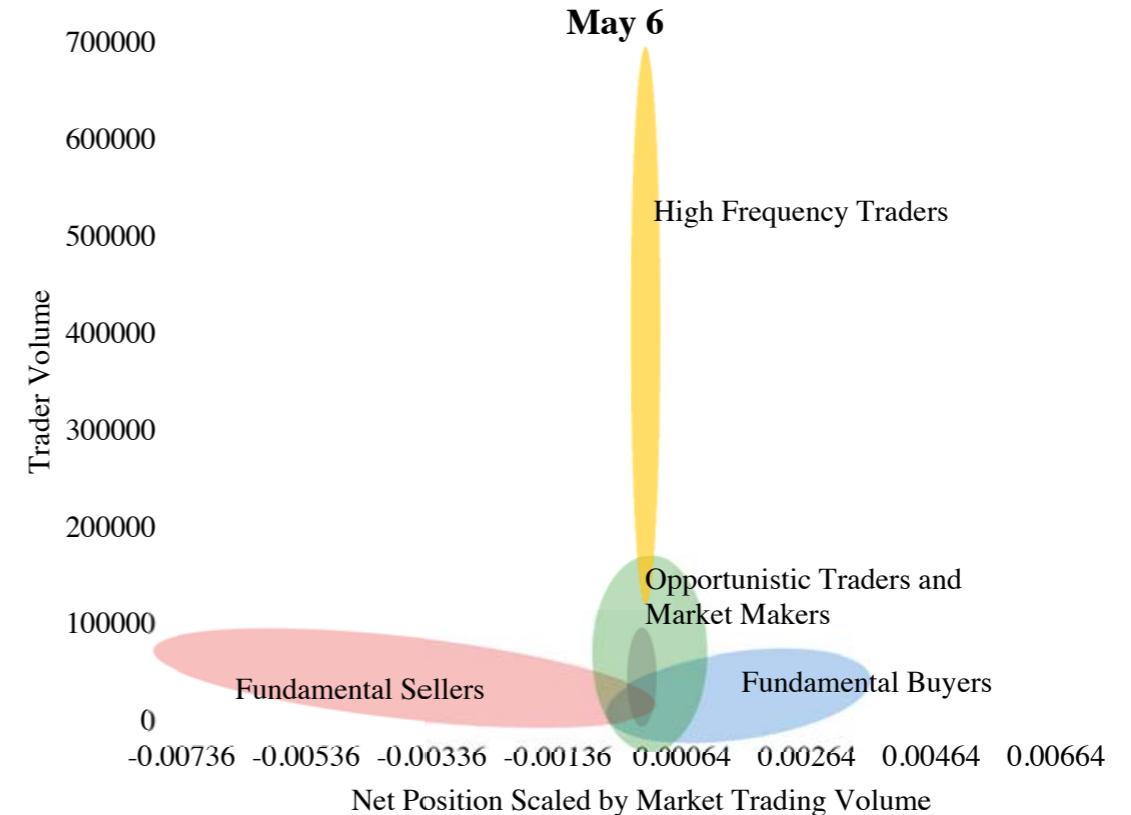
& CFTC Chief Economist
Andrei Kirilenko—MIT Sloan School of Management

Albert S. Kyle—University of Maryland

Mehrdad Samadi—University of North Carolina

Tugkan Tuzun—Board of Governors of the Federal Reserve System

Original Version: October 1, 2010
This version: May 5, 2014



B. The Data

For the day of the Flash Crash and three days prior to that, May 3-6, 2010, we examine transaction-level, “audit-trail” data for all regular transactions in the front-month June 2010 E-mini S&P 500 futures contract. These data come from the Trade Capture Report (TCR) dataset, which the CME provides to the Commodity Futures Trading Commission (CFTC) - the U.S. federal regulator of futures, options, and swaps markets.

For each transaction, we utilize fields with the account numbers for the buyer and the seller, the price and quantity transacted, the date and time (to the nearest second), a sequence ID number which sorts trades into chronological order even within one second, order type (market order or limit order), and an “aggressiveness” indicator stamped by the CME Globex matching engine - “N” for the resting order and “Y” for the order that executed against a resting order.

The source data is confidential. This means that the results we present often provide a deliberately obscured illustration of what we have actually rigorously established and validated. Moreover, even though we have checked and re-checked our results, they are unlikely to be ever independently validated by other researchers. Even with these limitations though, we still believe that we owe to the public to provide the most informative analysis of the extraordinary stressful events that unfolded in the E-mini on May 6, 2010 and the lessons for market design that we can learn from these events.

THE WALL STREET JOURNAL.

'Flash Crash' Investigators Likely Missed Clues
Strategies of Navinder Sarao, trader accused of helping cause selloff, could have been spotted sooner, oversight committee members say

By **BRADLEY HOPE** and **ANDREW ACKERMAN**

April 26, 2015 7:03 p.m. ET

While investigators had access to the full set of data from that day, they focused on a subset related to actual trades, the committee members said. Had investigators delved deeper into the bigger set that included all the bids and offers entered, they said, they likely would have noticed that Mr. Sarao single-handedly put enormous pressure on a key futures contract tied to the S&P 500 stock index by making bids and quickly canceling them in a bluffing tactic known as “spoofing.”

Equity data

- Consequences of fragmentation
- WRDS and how to use it

Central clearing enables fragmentation

The image shows the top navigation bar of the DTCC website. It features the DTCC logo with the tagline "Securing Today. Shaping Tomorrow." followed by five main service categories: INSTITUTIONAL TRADE PROCESSING, CLEARING SERVICES, SETTLEMENT & ASSET SERVICES, WEALTH MANAGEMENT SERVICES, COLLATERAL MANAGEMENT, DERIVATIVES SERVICES, and DATA SERVICES. Below this is a large banner with the word "CLEARING SERVICES" in gold letters.

Home / **Clearing Services**

Share Print

DTCC subsidiaries National Securities Clearing Corporation (NSCC) and Fixed Income Clearing Corporation (FICC) deliver highly efficient clearing services across the U.S. equities and fixed income markets, reducing risk and cost for clients, while ensuring safety and reliability in the marketplace.



EQUITIES CLEARING SERVICES

DTCC clears and settles virtually all broker-to-broker equity, listed corporate and municipal bond and unit investment trust (UIT) transactions in the U.S. equities markets, advancing new initiatives and driving development of products and services that mitigate risk, reduce costs and enhance processing efficiencies for market participants.



MEMX turns up the heat on US stock exchanges

Traders join forces to create a challenger aiming to take on the likes of NYSE and Nasdaq

Philip Stafford in London JANUARY 10, 2019



US stock exchanges are facing a new yet very familiar threat: some of their biggest users cranking up the pressure to cut the costs of trading shares.

Nine of the biggest participants in daily trading on American markets, including [Morgan Stanley](#), Fidelity and [E*Trade](#), are [joining forces](#) to create Members Exchange, or MEMX, in a bid to rival the New York Stock Exchange, Nasdaq and CBOE Global Markets.

Memx's arrival, likely in 2020, will throw it into the world's deepest equity market, where [more than a dozen exchanges](#) compete for business with a [host of alternative trading venues](#).

Equity market fragmentation

But one issue MEMX has already targeted is [the cost of vital market data](#), like shares quotes and size of an order book. It is a flashpoint between exchanges and their users.

Critics say NYSE, Nasdaq and CBOE are all listed companies answerable to shareholders and not their users. That, say the critics, creates a potential conflict of interest because [exchanges also act as semi-government agencies](#). They have raised fees on proprietary data with little resistance from the SEC and also supply and control the nearest low-cost competitor, a consolidated data feed known as the SIP.

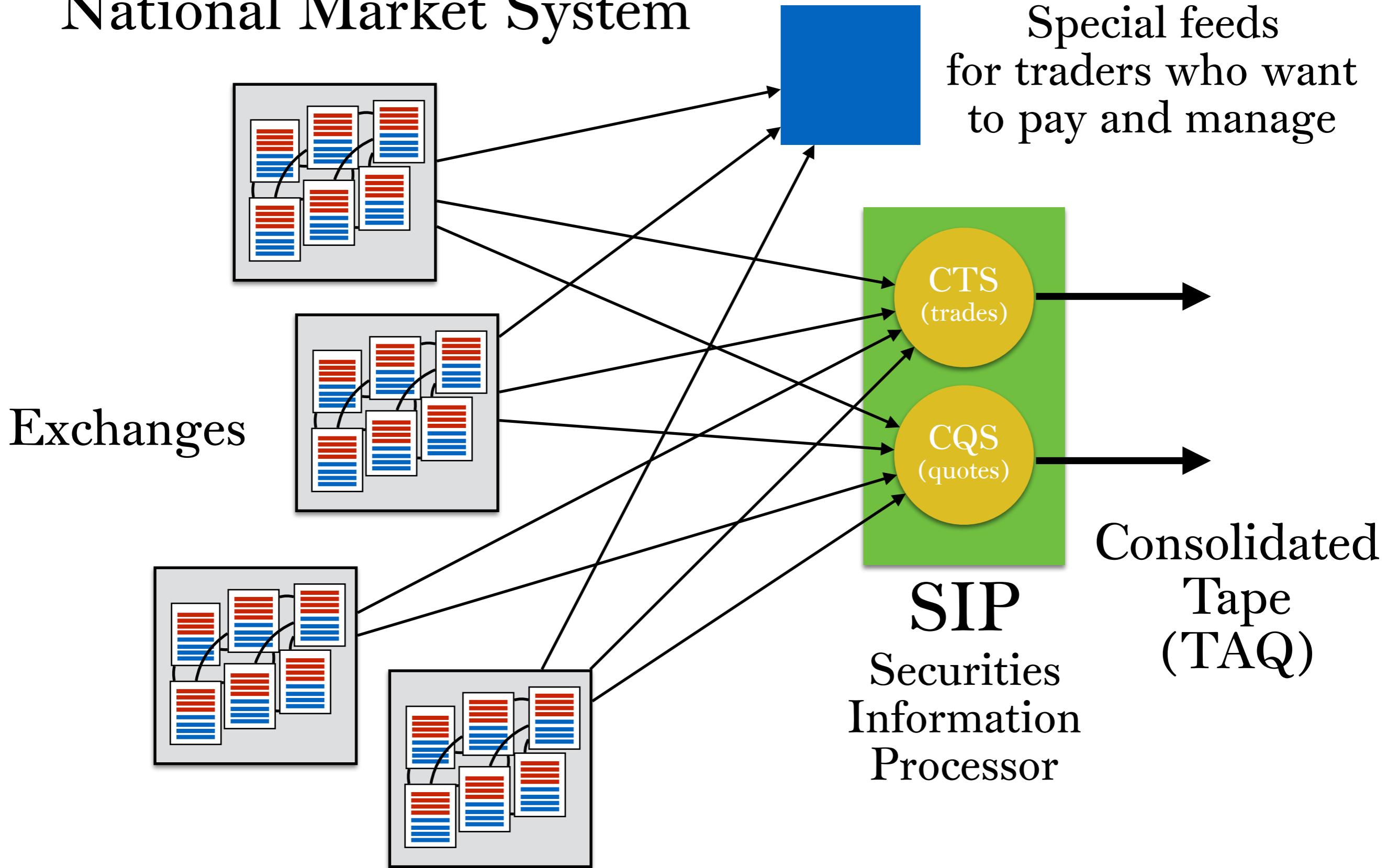
Last year Sifma, the trade body, complained that its members were charged up to 2,900 times more than they were in 2010 just to get the same data. In a rare move last October, the [SEC ruled](#) NYSE and Nasdaq had to show more justification for their fee increases.

Consequences of fragmented equity market

- Multiple exchanges in quote and trade stream
must merge to calculate NBBO
- Different latencies between quotes and trades
not always easy to merge
- Does not include depth (quotes below best)
- Professional participants may get direct feeds

Equity: exchanges feed to central facility

National Market System



The Consolidated Tape Association (CTA) oversees the dissemination of real-time trade and quote information in New York Stock Exchange LLC (Network A) and Bats, NYSE Arca, NYSE MKT and other regional exchange (Network B) listed securities. Since the late 1970s, all SEC-registered exchanges and market centers that trade Network A or Network B securities send their trades and quotes to a central consolidator where the Consolidated Tape System (CTS) and Consolidated Quote System (CQS) data streams are produced and distributed worldwide.

Facts about the Security Information Processor (SIP)

- The Security Information Processor (SIP) links the U.S. markets by processing and consolidating all protected bid/ask quotes and trades from every trading venue into a single, easily consumed data feed.
 - The SIP disseminates and calculates critical regulatory information including the National Best Bid and Offer (NBBO) and Limit Up Limit Down (LULD) price bands among other important information such as short sale restriction and regulatory halts.
 - CTA metrics on SIP performance (system availability, capacity, latency) and Subscriber use (non-professionals, professionals, quote usage, households) is available monthly and historically at SIP Metrics.
-

Consolidated Tape Association

From Wikipedia, the free encyclopedia

Both the CTS and the CQS systems are operated on fault tolerant computer platforms at different physical computer sites; CTS is operated at CTA's complex in Brooklyn; CQS is operated at CTA's complex in lower Manhattan, thereby providing redundancy in the event of a site disaster. If a site disaster should occur at either location, all of the computer processing would be transferred to the surviving site at reduced capacity.

TAQ data

- Capture of SIP feed: all exchanges merged
- Trade and Quote: top of book
- Available through WRDS
 - "millisecond TAQ" from 2015 onward
- Complication: have to deal with fragmentation

Proprietary Market Data

Real-Time

Indices

Historical

Reference



Daily TAQ (Trade and Quote) provides users access to all trades and quotes for all issues traded on NYSE, Nasdaq and the regional exchanges for the previous trading day.

It's a comprehensive history of daily activity from NYSE markets and the U.S. Consolidated Tape covering all U.S. Equities instruments (including all CTA and UTP participating markets).

The Daily TAQ files also include the NBBO, Master file, Quote and Trade Admin Messages, and, CTA and UTP Admin Message files.

Historical Data available as of 1993-present.

Resources

[Daily TAQ Client Spec v3.0 \[PDF\]](#)

[Fact Sheet \[PDF\]](#)

[Previous Specs](#)

[Sample Data](#)

[Fees \[PDF\]](#)

[Academics](#)

<https://www.nyse.com/market-data/historical/daily-taq>

WRDS is the delivery mechanism for many different datasets

<https://wrds.wharton.upenn.edu>

Wharton UNIVERSITY of PENNSYLVANIA wrds WHARTON RESEARCH DATA SERVICES

DEMO

Home / Wharton Research Data Services

Sign In

Username

Password

Welcome to WRDS!

Wharton Research Data Services (WRDS) is the award-winning research platform and business intelligence tool for over 40,000+ corporate, academic, government and nonprofit clients at over 400+ institutions in 30+ countries.

WRDS provides the user with one location to access over 250 terabytes of data across multiple disciplines including Accounting, Banking, Economics, Finance, ESG, and Statistics.

Flexible data delivery options include a powerful web query method that reduces research time, the WRDS Cloud for executing research and strategy development, and the WRDS client server using PCSAS, Matlab, Python and R.

Our Analytics team, doctoral-level support and rigorous data review and validation give clients the confidence to tailor research within complex databases and create a wide range of reliable data models.

WRDS provides access to S&P Capital IQ, CRSP, NYSE, Thomson Reuters, Bureau van Dijk, Global Insight, OptionMetrics and other important business research databases.

WRDS class account

Walra5ian–Auc10

- Login name orf474
- Password: based on "Walrasian Auction"
- Please use only for class purposes
or at least for academic purposes
(do not connect from corporate IP address or VPN)

ADDITIONAL TERMS APPLICABLE TO ACADEMIC AND GOVERNMENT SUBSCRIBERS

Academic and government subscribers:

- The WRDS services are for academic and non-commercial research purposes only. Users may not use data downloaded from the WRDS database for any non-academic or commercial endeavor.

Academic subscribers:

- Individual WRDS accounts are offered to standing faculty, full-time research staff, currently-enrolled Ph.D. students, and Masters students at Subscriber institutions. Access and individual accounts for research assistants, visiting scholars, and other non-permanent faculty is offered at Wharton Research Data Service's discretion. Access by other types of students is limited to class accounts, and special IP authenticated access within the library or computer lab, as provided by Wharton Research Data Services. Please note: Our WRDS policy prohibits undergraduate/Masters students from accessing WRDS during the extended break between semesters.

NYSE Trade and Quote

For more about this dataset, see the [Dataset List](#), [Manuals and Overviews](#) or [FAQs](#).

Trade and Quote

"Monthly Product". 1993 - 2014

[Consolidated Quotes](#)

[Consolidated Trades](#)

[Reg SHO \(NYSE Short Sales\)](#)

Millisecond Trade and Quote

"Daily Product". 2001 - present, updated daily.

[Consolidated Quotes](#)

[Consolidated Trades](#)

To subscribe: MarketDataHelp@nyse.com

Please note: WRDS charges an additional fee to access the NYSE TAQ data.

[Access pricing information HERE>>](#)

Description

- **Trade and Quote:** Contains intraday transactions data (trades and quotes) for all securities listed on the New York Stock Exchange (NYSE), American Stock Exchange (AMEX), as well as Nasdaq National Market System (NMS) and SmallCap issues.
- **Reg SHO - NYSE Short Sales:** Contains data related to **SEC Regulation SHO** per the terms of the SRO requirement to make Short Sale transactions data available on a monthly basis for trade dates beginning January 2005 through the pilot period ending 2007. (SROs are self-regulatory organizations in the securities industry, such as the NYSE, that create and enforce rules for its members based on the federal securities laws.)
- **Millisecond Trade and Quote:** Trades and Bids are tracked at the millisecond level (HHMMSSXXX).

Getting Started

- **Using TAQ Efficiently:** [Documentation](#)
- **WRDS NBBO:** WRDS has stored all the output datasets that contain the NBBOs at each second for all stocks with quotes data in TAQ since 1993, in the following directory on the wrds-cloud.wharton.upenn.edu server: `/wrds/nyse/sasdata/wrds_taqs_nbbo/`
- **Lee Ready:** In the market microstructure study, Lee and Ready (1991) methodology is widely used to infer trade directions: whether a trade is buyer-initiated or seller-initiated. In order to expedite trade classification algorithms, WRDS has generated two sets of data that rely on original TAQ Trades and Quotes datasets: One is the WRDS NBBO described above. The second are WRDS derived WCT files that have matched trades and NBBO quotes. They are located on wrds-cloud.wharton.upenn.edu under `/wrds/nyse/sasdata/wrds_taqs_ct` directory.



Listings

Markets

Data

Regulation

Insights



Proprietary Market Data

Real-Time

Indices

Historical

Reference

HISTORICAL DATA

Daily TAQ

Daily TAQ (Trade and Quote) provides users access to all trades and quotes for all issues traded on NYSE, Nasdaq and the regional exchanges for the previous trading day.

It's a comprehensive history of daily activity from NYSE markets and the U.S. Consolidated Tape covering all U.S. Equities instruments (including all CTA and UTP participating markets).

The Daily TAQ files also include the NBBO, Master file, Quote and Trade Admin Messages, and, CTA and UTP Admin Message files.

Historical Data available as of 1993-present.

Resources

[Daily TAQ Client Spec v3.0d \[PDF\]](#)

[Fact Sheet \[PDF\]](#)

[Previous Specs](#)

[Sample Data](#)

[Fees \[PDF\]](#)

[Academics](#)

https://www.nyse.com/publicdocs/nyse/data/Daily_TAQ_Client_Spec_v3.0d.pdf
(Version 3.0d from May 2019)

<https://www.nyse.com/market-data/historical/daily-taq>

Different feeds, depending on
where stock is listed
(all stocks trade on all exchanges)

ANNOUNCEMENTS CONTRACTS MEETINGS PLANS PRICING
TOOLS COMPLIANCE TECHNICAL POLICY GUIDELINES

<https://www.ctaplan.com>

CTA=NYSE
(3 letters)

Advisory Committee

CTA Chair

Participants

Cboe BYX
Exchange, Inc.

The Consolidated Tape Association (CTA) oversees the dissemination of real-time trade and quote information in New York Stock Exchange LLC (Network A) and Bats, NYSE Arca, NYSE American and other regional exchange (Network B) listed securities. Since the late 1970s, all SEC-registered exchanges and market centers that trade Network A or Network B securities send their trades and quotes to a central consolidator where the Consolidated Tape System (CTS) and Consolidated Quote System (CQS) data streams are produced and distributed worldwide.

UTP=Nasdaq
(4 letters)



<http://www.utpplan.com>

OVERVIEW | DATA ADMIN | TECHNICAL | ALERTS | METRICS | UTP PLAN | PARTICIPANTS | MEETING SUMMARY | SUPPORT

Overview

The UTP Plan (Plan) governs the collection, consolidation and dissemination of Quotation and Transaction Information for Nasdaq-listed Securities Traded on Exchanges on an Unlisted Trading Privilege (UTP) Basis. The Plan is administered by the participant exchanges and associations listed under Participant section of this web page. These organizations determine policy matters and oversee system operations. The Plan has been approved by the Securities and Exchange Commission (SEC).

TAQ trade information

- Date, time--NY local, nanosecond (!) precision
- Symbol, sequence number (specific to symbol)
- Trade size, price
- Exchange that trade printed on
- Various other information:
 - condition code (O,Q=open, F=sweep, etc
 - stopped stock indicator
- Missing: aggressiveness or "trade direction"
 - buy market order against sell limit order or conversely
 - must determine this by matching trades against quotes

Example: LECO in Nov 2018

1000th largest stock in Dec 2018

CSV trade file as downloaded from WRDS (.csv.gz)

```
europia: gunzip -c trade-LECO-20181114.csv.gz | head
DATE,TIME_M,EX,SYM_R00T,SYM_SUFFIX,TR_SCOND,SIZE,PRICE,TR_CORR,TR_SEQNUM,TR_SOURCE,TR_RF
20181114,9:30:00.361488291,Q,LECO,,@0 X,2123,83.96,00,6989,N,
20181114,9:30:00.361503399,Q,LECO,,@ Q,2123,83.96,00,6990,N,
20181114,9:30:00.824299258,B,LECO,,@F I,9,83.57,00,8396,N,
20181114,9:35:00.007732410,X,LECO,,@F I,8,83.58,00,47943,N,
20181114,9:35:55.214374911,D,LECO,,@ I,10,84.16,00,55886,N,Q
20181114,9:36:55.724860358,Q,LECO,,@ I,20,83.995,00,66391,N,
20181114,9:38:02.042408127,J,LECO,,@F I,17,84,00,75776,N,
20181114,9:38:02.042433721,J,LECO,,@F I,10,83.82,00,75777,N,
20181114,9:38:56.750385428,B,LECO,,@ I,20,83.985,00,82830,N,
```

Loaded into R as a data.table

```
> head(T,9)
```

	date	time	exch	sym	prc	siz	cond	corr	seq	src	rf
1:	2018-11-14	34200.3614883	Q	LECO	83.960	2123	@0 X	0	6989	N	
2:	2018-11-14	34200.3615034	Q	LECO	83.960	2123	@ Q	0	6990	N	
3:	2018-11-14	34200.8242993	B	LECO	83.570	9	@F I	0	8396	N	
4:	2018-11-14	34500.0077324	X	LECO	83.580	8	@F I	0	47943	N	
5:	2018-11-14	34555.2143750	D	LECO	84.160	10	@ I	0	55886	N	Q
6:	2018-11-14	34615.7248604	Q	LECO	83.995	20	@ I	0	66391	N	
7:	2018-11-14	34682.0424082	J	LECO	84.000	17	@F I	0	75776	N	
8:	2018-11-14	34682.0424337	J	LECO	83.820	10	@F I	0	75777	N	
9:	2018-11-14	34736.7503855	B	LECO	83.985	20	@ I	0	82830	N	

Time in seconds after midnight NY
(09:30:00 = 9.5*60*60 = 34200)

Exchange where the trade printed
Condition code:

(O=opening, @=normal, I=odd lot, etc)

Looks like opening print is duplicated

Timestamps

Timestamp History

The accuracy of timestamps published by the SIPs has improved in steps over time, with CTA and UTP on different schedules from each other. Timestamps in Historical Daily TAQ reflect this evolution. For consumers of historical Daily TAQ, here is the dated change list for SIP timestamps.

Timestamps are durations since midnight, eg: seconds since midnight, nanoseconds since midnight

DATE OF CHANGE	UTP CHANGES	CTA CHANGES
January 1993	seconds HHMMSS	seconds HHMMSS
October 2003	milliseconds HHMMSSxxx	milliseconds HHMMSSxxx
July 27, 2015	microseconds HHMMSSxxxxxx	
August 3, 2015		microseconds HHMMSSxxxxxx
October 24, 2016	nanoseconds HHMMSSxxxxxxxxx	microseconds HHMMSSxxxxxx000
September 18, 2017		nanoseconds HHMMSSxxxxxxxxx

1 nanosecond = 1 foot
 They are certainly not this precise
 R does not represent time with this much precision

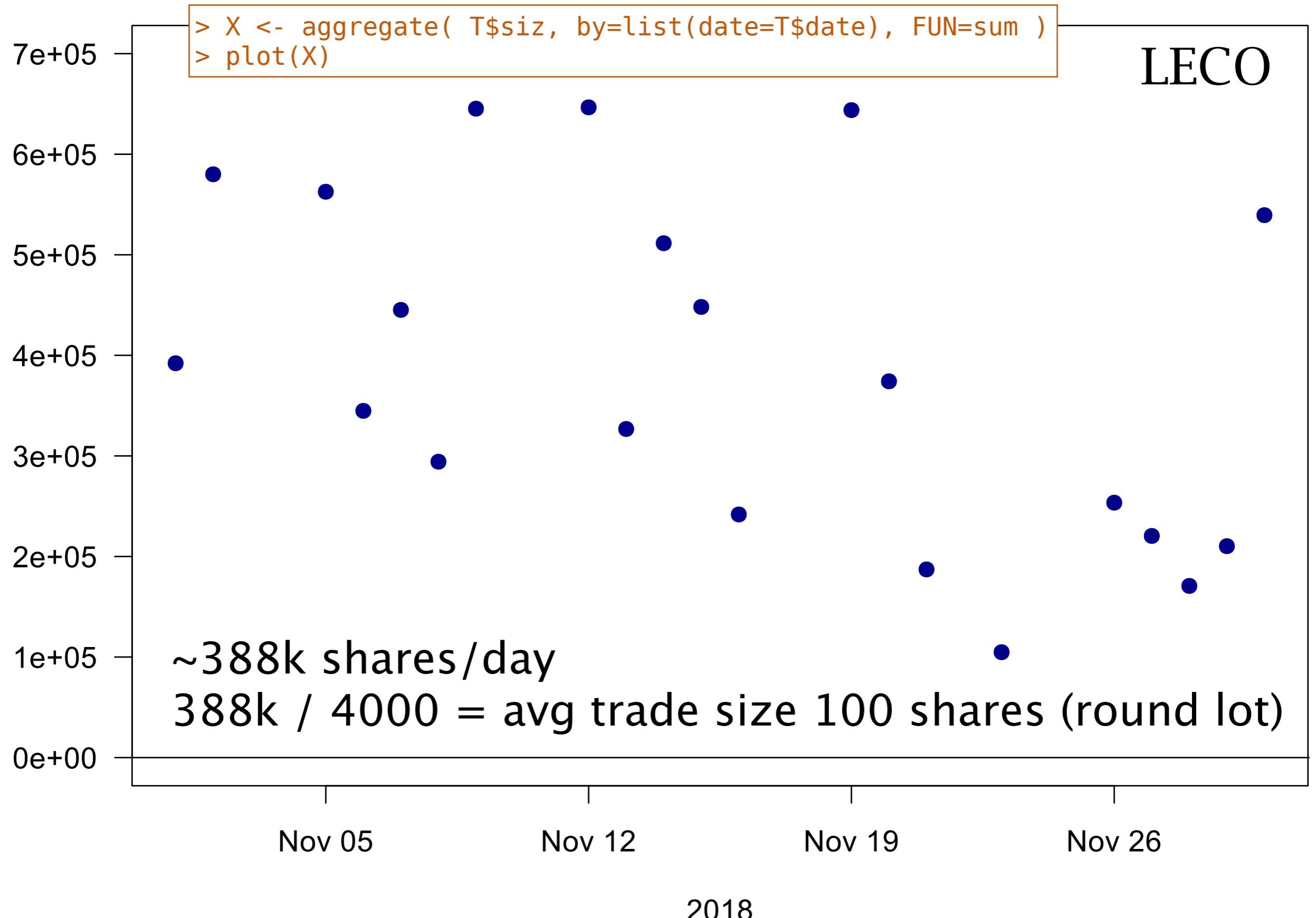
Number of trade data records per day (Very reasonable for this moderate stock)

```
> aggregate( T$siz, by=list(date=T$date), FUN=length )  
      date   x
```

1	2018-11-01	4581
2	2018-11-02	4316
3	2018-11-05	4761
4	2018-11-06	4224
5	2018-11-07	5071
6	2018-11-08	3401
7	2018-11-09	5116
8	2018-11-12	3566
9	2018-11-13	4108
10	2018-11-14	5808
11	2018-11-15	5323
12	2018-11-16	2677
13	2018-11-19	3968
14	2018-11-20	4189
15	2018-11-21	2649
16	2018-11-23	1429
17	2018-11-26	3167
18	2018-11-27	2988
19	2018-11-28	2329
20	2018-11-29	2787
21	2018-11-30	4898

~3900 rows/day

Traded volume per day



Volume by exchange (LECO, Nasdaq stock, Nov 2018)

```

> X <- aggregate( T$siz,
+                 by=list(T$exch), FUN=sum )
> X[ order(X$x, decreasing=TRUE)
  Group.1          x
10    Q 3664460   Nasdaq
4     D 2278048   FINRA ADF
9     P 659327   NYSE Arca
14    Z 363018   Bats
11    V 240317   IEX
13    Y 226292   Bats
2     B 203179   Nasdaq
6     K 183454   Bats
5     J 139553   Bats
12    X  93277   Nasdaq
8     N  44409   NYSE
3     C  37598   NSX
1     A  8987    Nyse
7     M  1532    Chicago
  
```

15. APPENDIX F: PARTICIPANT IDs

Table 16 The Participant ID is 1 text, Alphabetic character field, which identifies the Exchange or Processor that initiated the message:

CODE	Description
A	NYSE American, LLC (NYSE American)
B	NASDAQ OMX BX, Inc. (NASDAQ OMX BX)
C	NYSE National, Inc (NYSE National)
D	FINRA Alternative Display Facility (ADF)
I	International Securities Exchange, LLC (ISE)
J	Cboe EDGA Exchange (Cboe EDGA)
K	Cboe EDGX Exchange (Cboe EDGX)
M	Chicago Stock Exchange, Inc. (CHX)
N	New York Stock Exchange, LLC (NYSE)
P	NYSE Arca, Inc. (NYSE Arca)
S	Consolidated Tape System (CTS)
T	NASDAQ Stock Market, LLC (in Tape A, B securities) (NASDAQ)
Q	NASDAQ Stock Exchange, LLC (in Tape C securities)
V	The Investors' Exchange, LLC (IEX)
W	Chicago Broad Options Exchange, Inc.(CBSX)
X	NASDAQ OMX PSX, Inc. (NASDAQ OMX PSX)
Y	Cboe BYX Exchange, Inc (Cboe BYX)
Z	Cboe BZX Exchange, Inc (Cboe BZX)

Volume by condition code

(LECO, Nasdaq stock, Nov 2018)

```
> X <- aggregate( T$siz,
+                 by=list(T$cond), FUN=sum )
> X[ order(X$x, decreasing=TRUE), ]
  Group.1      x
1      @ 3161467 Regular
13     @F 1301950 Regular + Sweep
3      @ M 801037 Regular + Close
12     @6 X 796904 ...
2      @ I 628504
14     @F I 494574
5      @ T 474586
8      @ TW 346078
11     @4 W 40263
4      @ Q 38459
16     @0 X 37128
7      @ TP 11232
19     N T 3419
21     N TW 2685
9      @4 2393
6      @ TI 1681
10     @4 I 568
  ...
```

3. Daily TAQ Trades File

Every trade reported to the consolidated tape by all CTA and UTP participants.

The Daily TAQ Trades File is sorted by symbol, time, and message sequence number.

	FIELD NAME	MAX SIZE	FORMAT	DESCRIPTION
1	Time	15	Timestamp	Time the trade was published by the SIP
2	Exchange	1	Text	The ID of the exchange where the trade took place See Participant IDs for codes
3	Symbol	17	Symbol	Stock symbol
4	Sale Condition	4	Text	A special condition associated with the trade. Up to four codes are displayed per trade. <ul style="list-style-type: none"> ▪ @ = Regular Trade ▪ A = Acquisition ▪ B = Bunched Trade ▪ C = Cash Sale ▪ D = Distribution ▪ E = Placeholder for future use ▪ F = Intermarket Sweep ▪ G = Bunched Sold Trade ▪ H = Price Variation Trade ▪ I = Odd Lot Trade ▪ K = Rule 155 Trade (NYSE American Only) ▪ L = Sold Last ▪ M = Market Center Close Price ▪ N = Next Day ▪ O = Opening Prints ▪ P = Prior Reference Price ▪ Q = Market Center Open Price ▪ R = Seller ▪ S = Split Trade ▪ T = Form - T Trade ▪ U = Extended Trading Hours ▪ V = Stock-Option Trade ▪ W = Average Price Trade ▪ X = Cross Trade ▪ Y = Yellow Flag ▪ Z = Sold (Out of Sequence) ▪ 1 = Stopped Stock - Regular Trade ▪ 4 = Derivatively Priced ▪ 5 = Re-opening Prints ▪ 6 = Closing Prints ▪ 7 = Exempt Qualified Contingent Trade ("QCT") ▪ 8 = Placeholder for 611 Exempt ▪ 9 = Corrected Consolidated Close (per listing market)

Probably want to exclude exch='D' from microstructure studies



Financial Industry Regulatory Authority

About FINRA | Newsroom | Compliance Tools | Contact | Firm Gateway Sign In

Search Q

Industry Home	Oversight	Rules & Guidance	Registration & Qualification	Regulatory Filing & Reporting	Education & Training	Key Topics
---------------	-----------	------------------	------------------------------	-------------------------------	----------------------	------------

[Home](#) » [For Industry Professionals](#) » [Regulatory Filing & Reporting](#)

Alternative Display Facility (ADF)

The Alternative Display Facility (ADF) is an SRO display only facility that is operated by FINRA. The ADF provides members with a facility for the display of quotations, the reporting of trades, and the comparison of trades. ADF best bid and offer and trade reports are included in the consolidated data stream for Nasdaq and CQS listed securities. As an SRO display only facility, ADF does not provide automated order routing functionality, execution facilities, or linkages between ADF trading centers. ADF trading centers are required by FINRA rule to provide direct electronic access to all other ADF trading centers and provide such access to all FINRA members that request it. All FINRA members in good standing are eligible to participate in ADF, pending execution of appropriate contracts and meeting specific requirements as set forth by FINRA.

RELATED LINKS

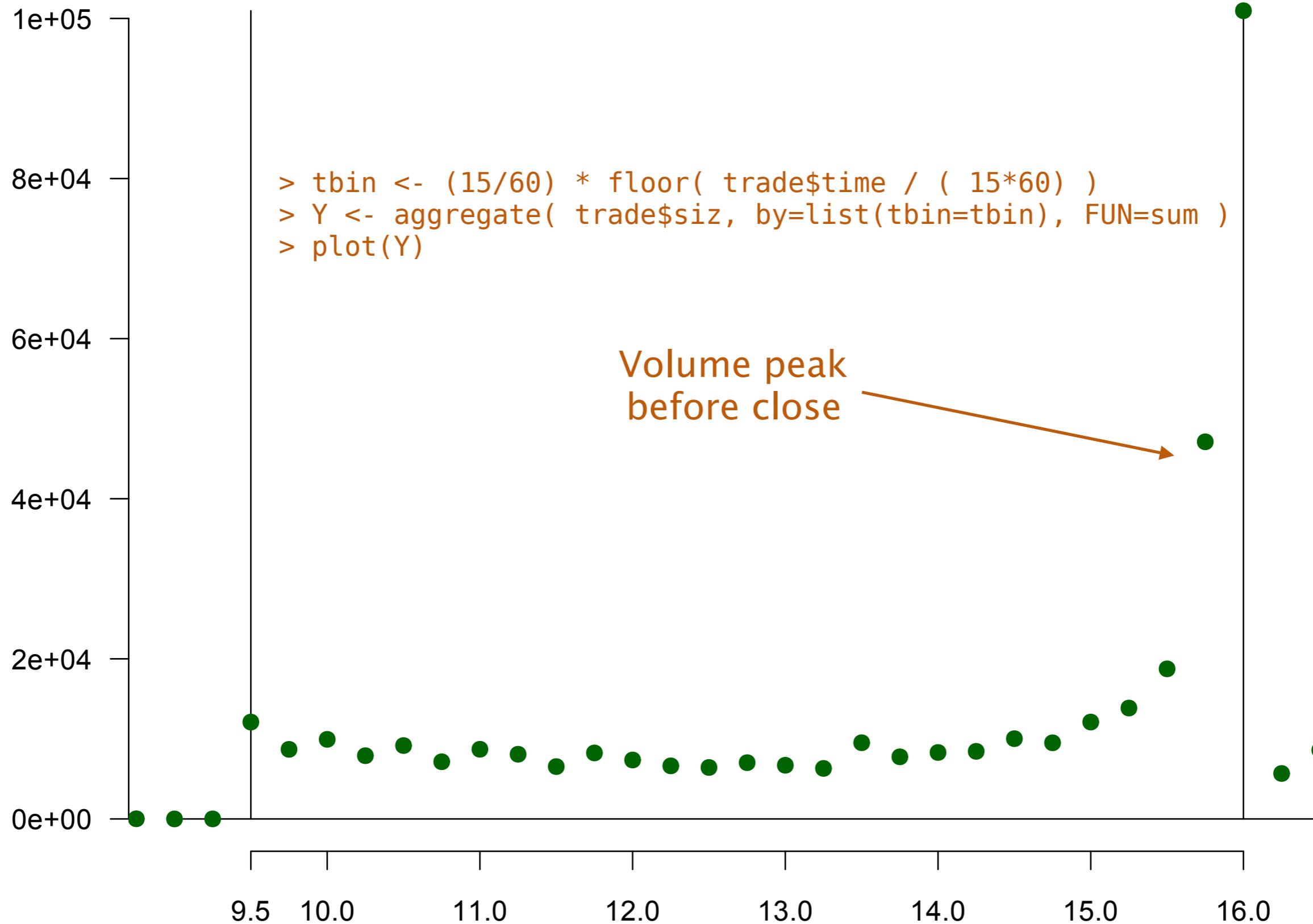
[Market Data](#)

[Trade Reporting Frequently Asked Questions](#)

[FINRA Automated Data Delivery System \(FINRA ADDS\)](#)

Intraday volume profile (15-minute bins)

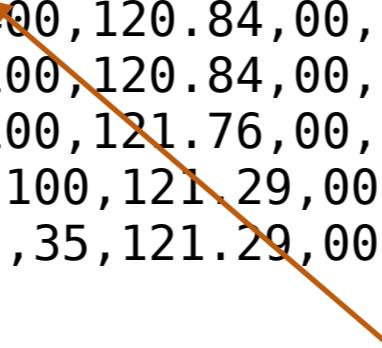
LECO



IBM in Nov 2018

CSV trade file as downloaded from WRDS (.csv.gz)

europa: gunzip -c trade-IBM-20181114.csv.gz | head

DATE,TIME_M,EX,SYM_ROOT,SYM_SUFFIX,TR_SCOND,SIZE,PRICE,TR_CORR,TR_SEQNUM,TR_SOURCE,TR_RF
20181114,6:21:43.139325000,T,IBM,,FTI,1,121.37,00,62201,C,
20181114,6:21:43.139356000,T,IBM,,FTI,2,121.49,00,62301,C,
20181114,6:56:23.149498000,T,IBM,,FTI,1,120.63,00,65701,C,
20181114,6:56:23.149521000,T,IBM,,FTI,89,120.58,00,65801,C,
20181114,8:00:05.497908000,D,IBM,,T,400,120.84,00,159301,C,N

20181114,8:00:05.498517000,D,IBM,,T,100,120.84,00,159401,C,N
20181114,8:00:05.529900000,D,IBM,,T,100,121.76,00,160801,C,N
20181114,8:08:43.932471000,P,IBM,,FT,100,121.29,00,175501,C,
20181114,8:12:30.265043000,P,IBM,,FTI,35,121.29,00,178601,C,

Loaded into R as a data.table

T = "extended hours trade"

```
> head(T, 9)
    date        time exch sym   prc siz cond corr     seq src rf
1: 2018-11-14 22903.1393249    T IBM 121.37   1 FTI    0 62201 C
2: 2018-11-14 22903.1393559    T IBM 121.49   2 FTI    0 62301 C
3: 2018-11-14 24983.1494980    T IBM 120.63   1 FTI    0 65701 C
4: 2018-11-14 24983.1495211    T IBM 120.58  89 FTI    0 65801 C
5: 2018-11-14 28805.4979081    D IBM 120.84 400    T    0 159301 C N
6: 2018-11-14 28805.4985170    D IBM 120.84 100    T    0 159401 C N
7: 2018-11-14 28805.5299001    D IBM 121.76 100    T    0 160801 C N
8: 2018-11-14 29323.9324710    P IBM 121.29 100    FT   0 175501 C
9: 2018-11-14 29550.2650430    P IBM 121.29  35 FTI   0 178601 C
```

Volume by exchange (IBM, NYSE stock, Nov 2018)

```
> X <- aggregate( T$siz,  
+                 by=list(T$exch), FUN=sum )  
> X[ order(X$x, decreasing=TRUE)  
  Group.1      x  
4       D 56285418 FINRA ADF  
8       N 48380216 NYSE  
10      T 14007009 Nasdaq  
9       P 12757813 NYSE Arca  
14      Z 8632275 Bats  
7       M 5206733 CSX  
6       K 5093790 Bats  
11      V 3822729 IEX  
13      Y 2667438 Bats  
2       B 2639380 Nasdaq  
5       J 1552678 Bats  
3       C 720671 NSX  
12      X 276414 Nasdaq  
1       A 187664 NYSE
```

15. APPENDIX F: PARTICIPANT IDs

Table 16 The Participant ID is 1 text, Alphabetic character field, which identifies the Exchange or Processor that initiated the message:

CODE	Description
A	NYSE American, LLC (NYSE American)
B	NASDAQ OMX BX, Inc. (NASDAQ OMX BX)
C	NYSE National, Inc (NYSE National)
D	FINRA Alternative Display Facility (ADF)
I	International Securities Exchange, LLC (ISE)
J	Cboe EDGA Exchange (Cboe EDGA)
K	Cboe EDGX Exchange (Cboe EDGX)
M	Chicago Stock Exchange, Inc. (CHX)
N	New York Stock Exchange, LLC (NYSE)
P	NYSE Arca, Inc. (NYSE Arca)
S	Consolidated Tape System (CTS)
T	NASDAQ Stock Market, LLC (in Tape A, B securities) (NASDAQ)
Q	NASDAQ Stock Exchange, LLC (in Tape C securities)
V	The Investors' Exchange, LLC (IEX)
W	Chicago Broad Options Exchange, Inc.(CBSX)
X	NASDAQ OMX PSX, Inc. (NASDAQ OMX PSX)
Y	Cboe BYX Exchange, Inc (Cboe BYX)
Z	Cboe BZX Exchange, Inc (Cboe BZX)

Volume by condition code

(IBM, NYSE stock, Nov 2018)

```
> X <- aggregate( T$siz,
  by=list(T$cond), FUN=sum )
> X[ order(X$x, decreasing=TRUE), ]
  Group.1      x
1       66471609 Regular
16      F 32813676 Sweep
23      M 12621562 Close
6       6 12605931 Close
21      I 8001507 Odd lot
47      Z 5607740 Out of seq
17      F I 5332622 ...
7       7 B 5020336
40      T 3514085
9       7 V 2620008
30      0 2296896
3       4 B 2239658
43      TP 970432
41      TB 693841
26      N T 663177
18      FT 203101
46      V 160000
...
...
```

3. Daily TAQ Trades File

Every trade reported to the consolidated tape by all CTA and UTP participants.

The Daily TAQ Trades File is sorted by symbol, time, and message sequence number.

	FIELD NAME	MAX SIZE	FORMAT	DESCRIPTION
4	Sale Condition	4	Text	<p>A special condition associated with the trade. Up to four codes are displayed per trade.</p> <p>CTA issues</p> <ul style="list-style-type: none"> blank = Regular Trade (no associated conditions) B = Average Price Trade C = Cash Trade (Same Day Clearing) E = Automatic Execution F = Intermarket Sweep Order H = Price Variation Trade I = Odd Lot Trade K = Rule 127 (NYSE only) or Rule 155 Trade (NYSE American only) L = Sold Last (Late Reporting) M = Market Center Official Close N = Next Day Trade (Next Day Clearing) O = Market Center Opening Trade P = Prior Reference Price Q = Market Center Official Open R = Seller T= Extended Hours Trade U = Extended Hours (Sold Out of Sequence) V = Contingent Trade X = Cross Trade Z = Sold (Out of Sequence) 4 = Derivatively Priced 5 = Market Center Re-Opening Trade 6 = Market Center Closing Trade 7 = Qualified Contingent Trade (from August 3, 2015) 8 = Reserved 9 = Corrected Consolidated Close Price per the Listing Market

IBM around open and close

```
> head(T[ T$time >= 9.5*60*60, ], 10 )
  date      time exch sym   prc   siz cond corr   seq src rf
1: 2018-11-14 34200.3737619    T IBM 121.26     3 FTI    0 517201 C
2: 2018-11-14 34202.2676430    Z IBM 121.12     2 I    0 586201 C
3: 2018-11-14 34202.2677050    Z IBM 121.12     1 I    0 586301 C
4: 2018-11-14 34207.5961990    Z IBM 121.12     1 I    0 644001 C
5: 2018-11-14 34210.0651469    N IBM 121.11 90759     0       0 660201 C
6: 2018-11-14 34210.0675740    N IBM 121.14    400       0 660401 C
7: 2018-11-14 34210.0680850    N IBM 121.13     10 I    0 660501 C
8: 2018-11-14 34210.0681040    N IBM 121.12     90 I    0 660601 C
9: 2018-11-14 34210.0718701    T IBM 121.18    100 F    0 660701 C
10: 2018-11-14 34210.0718930   T IBM 121.18    100 Q    0 660801 C
```

NYSE closing auction:
550,000 shares, \$66M
17 seconds after 16:00

3. Daily TAQ Trades File

Every trade reported to the consolidated tape by all CTA and UTP participants.

The Daily TAQ Trades File is sorted by symbol, time, and message sequence number.

	FIELD NAME	MAX SIZE	FORMAT	DESCRIPTION
4	Sale Condition	4	Text	A special condition associated with the trade. Up to four codes are displayed per trade. CTA issues
				M = Market Center Official Close

```
> T[ cond=='M' ]
  date      time exch sym   prc   siz cond corr   seq src rf
1: 2018-11-14 57600.1926169    P IBM 120.2 1800 M    0 196930301 C
2: 2018-11-14 57600.4111030    T IBM 120.2    12 M    0 196940701 C
3: 2018-11-14 57777.6774750    N IBM 120.2 549748 M    0 197121401 C
```

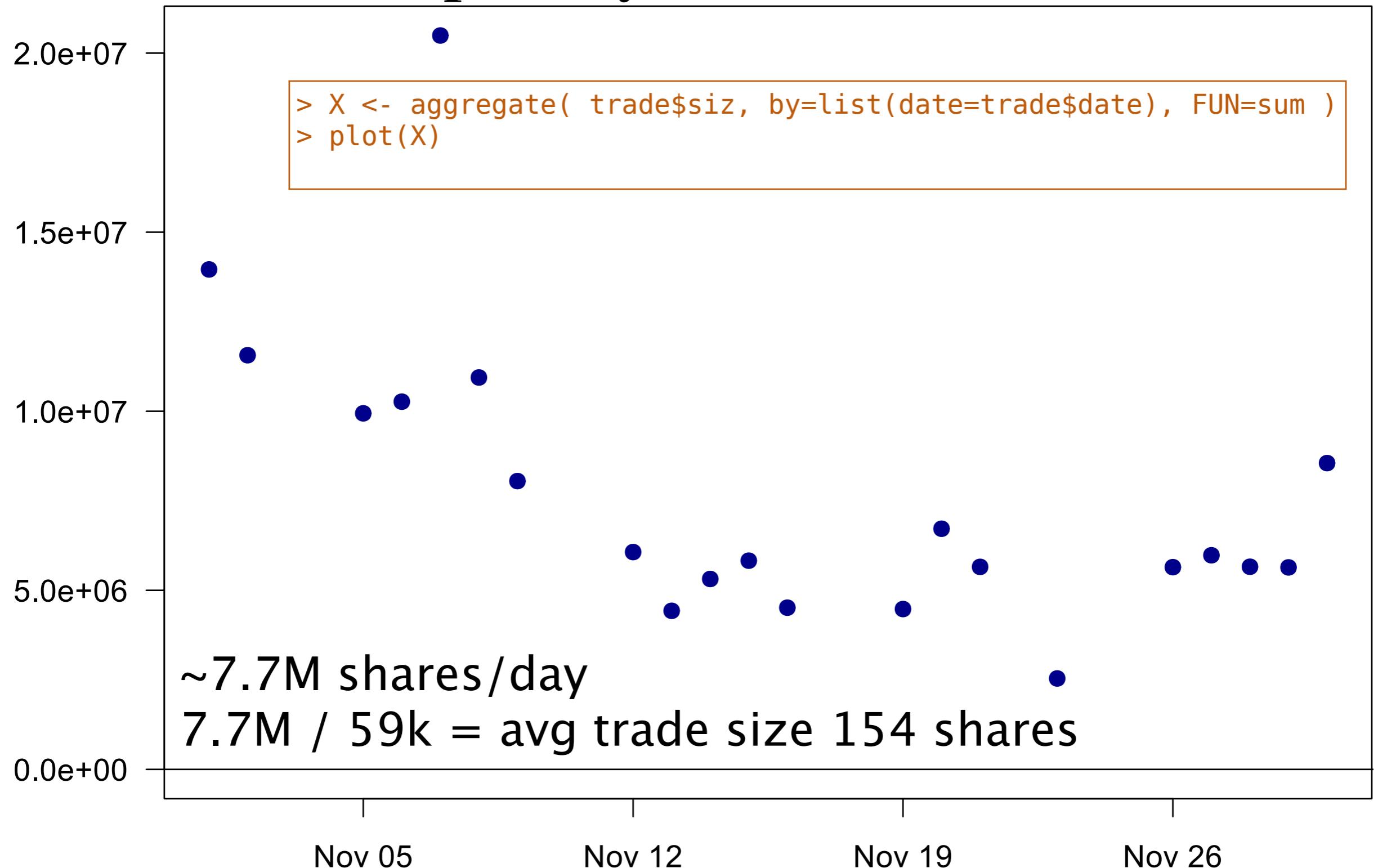
Larger number of data records per day

```
> aggregate( T$siz, by=list(date=T$date), FUN=length )  
    date      x  
1 2018-11-01 96415  
2 2018-11-02 73261  
3 2018-11-05 75840  
4 2018-11-06 85962  
5 2018-11-07 56747  
6 2018-11-08 47164  
7 2018-11-09 47119  
8 2018-11-12 45055  
9 2018-11-13 34693  
10 2018-11-14 41587  
11 2018-11-15 46415  
12 2018-11-16 33220  
13 2018-11-19 37970  
14 2018-11-20 53968  
15 2018-11-21 44521  
16 2018-11-23 20495  
17 2018-11-26 38033  
18 2018-11-27 37152  
19 2018-11-28 44926  
20 2018-11-29 46360  
21 2018-11-30 47259
```

~50,000 rows/day

Traded volume per day

IBM

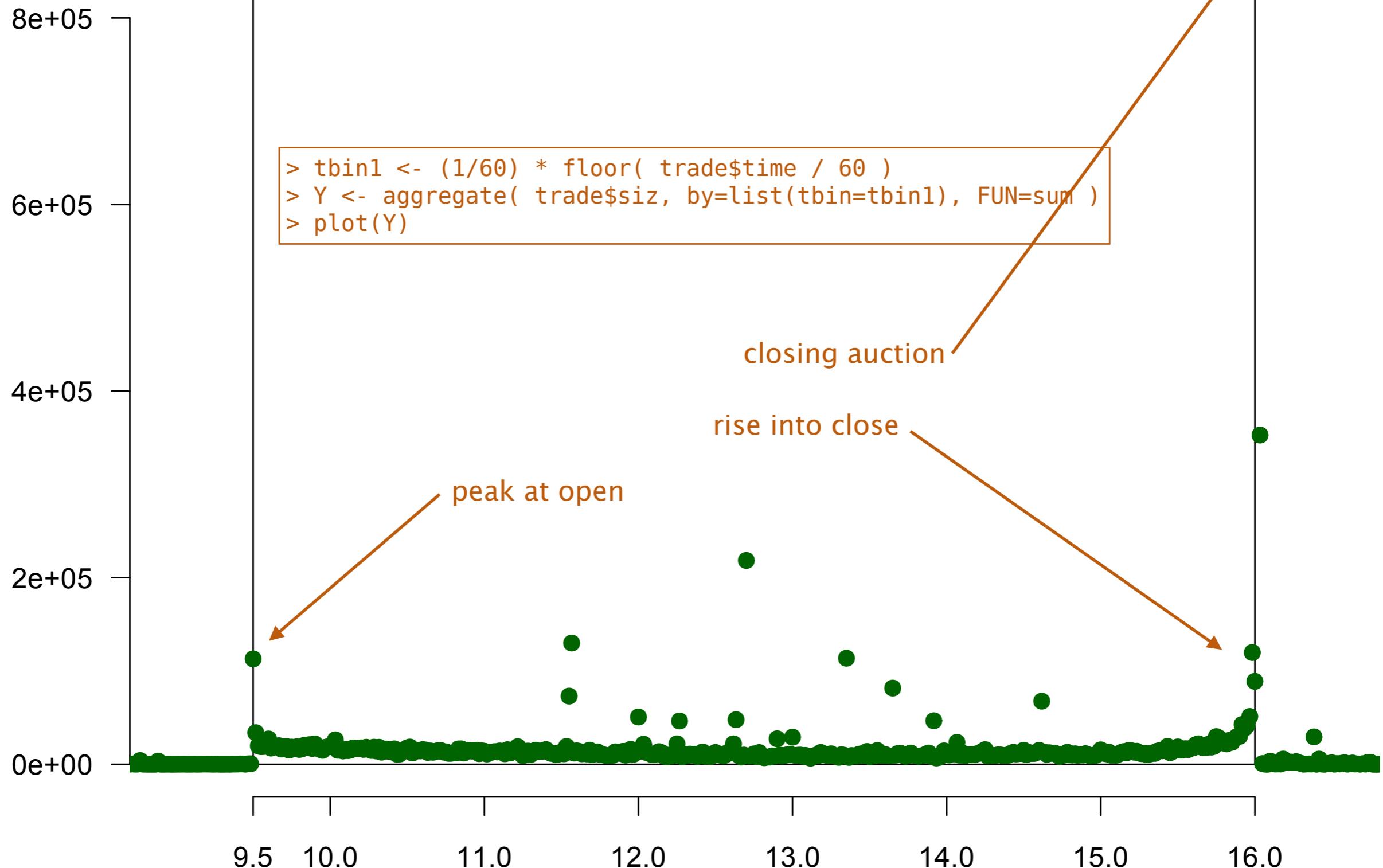


~7.7M shares/day

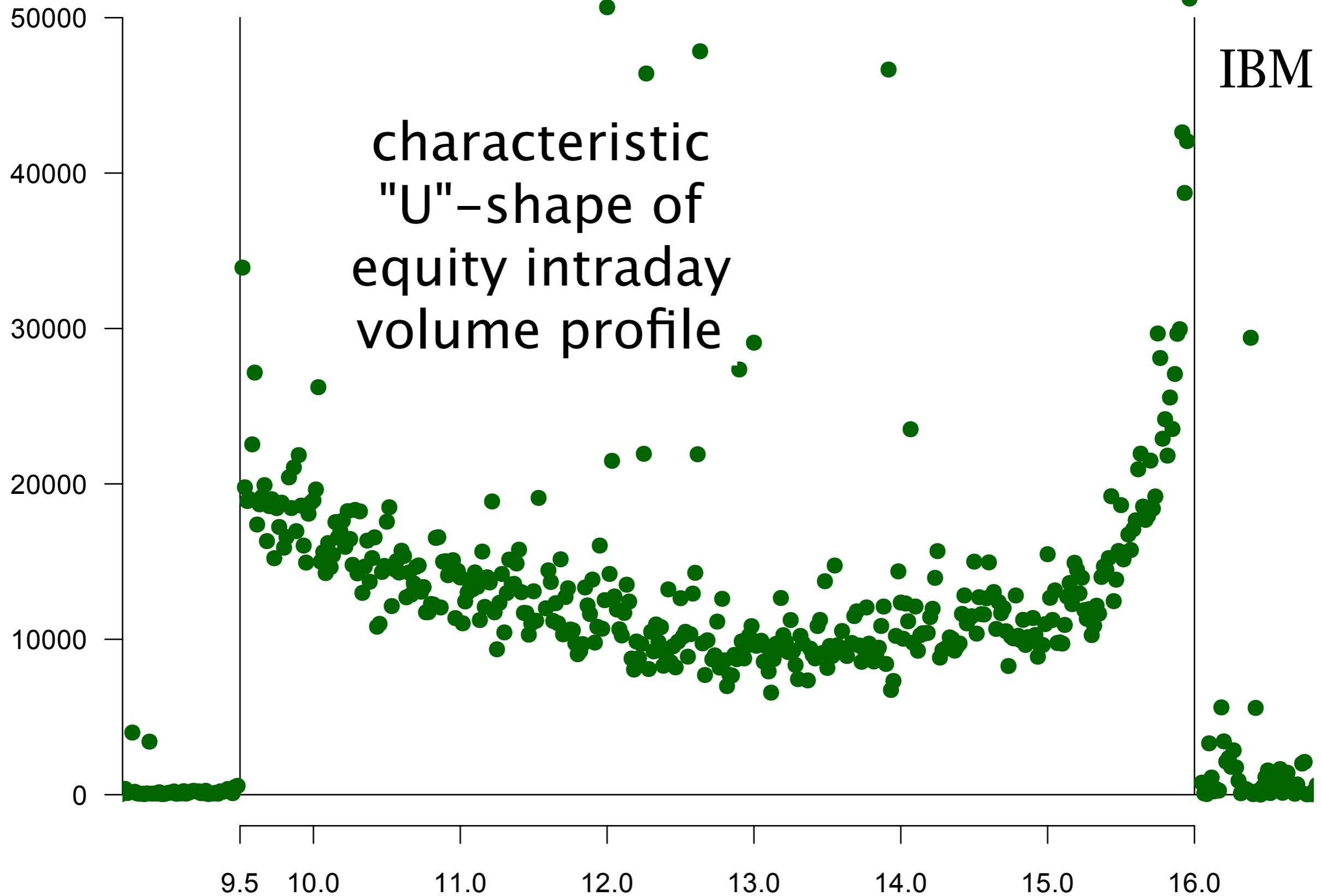
7.7M / 59k = avg trade size 154 shares

Volume profile, 1-minute resolution

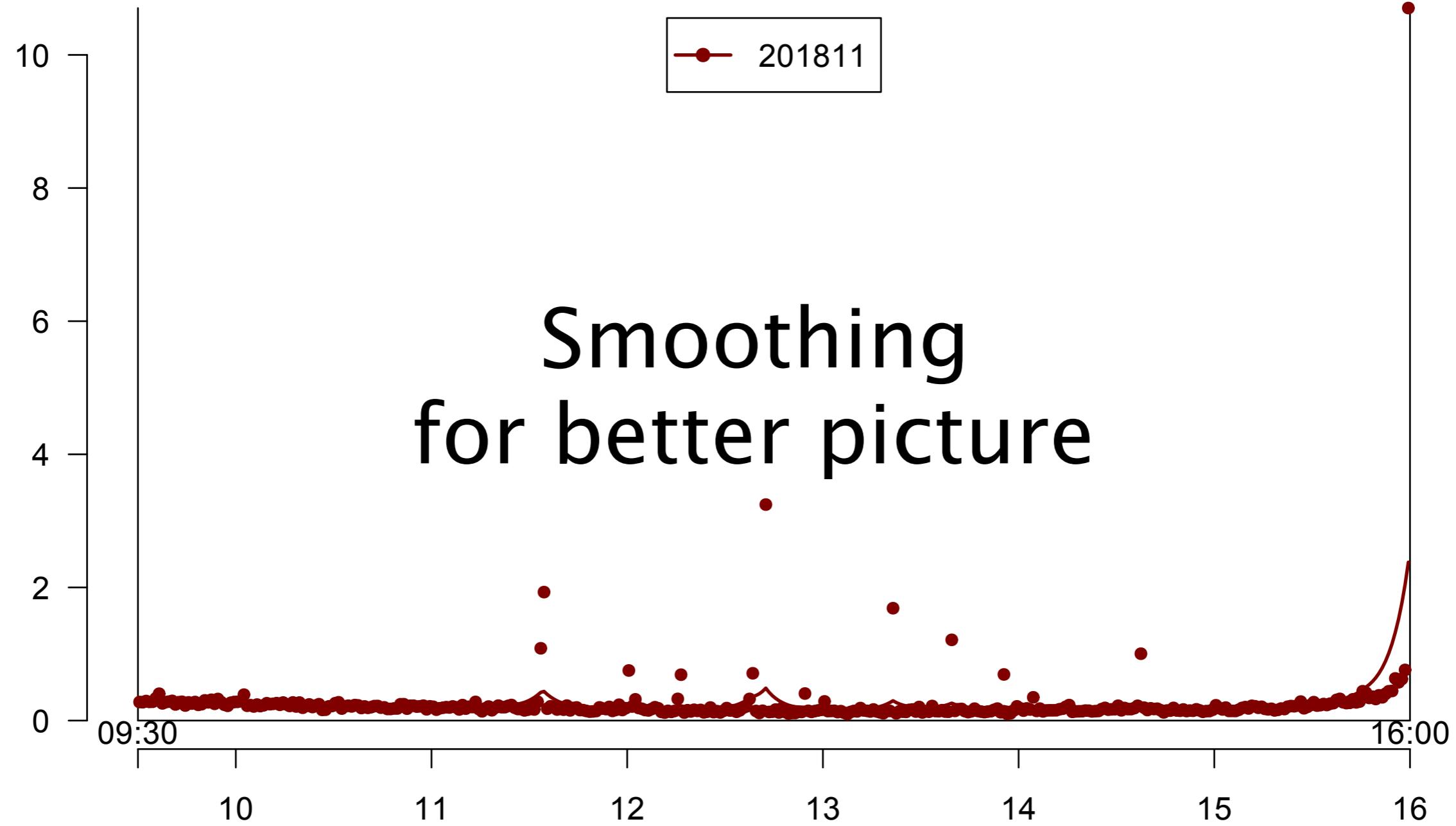
IBM

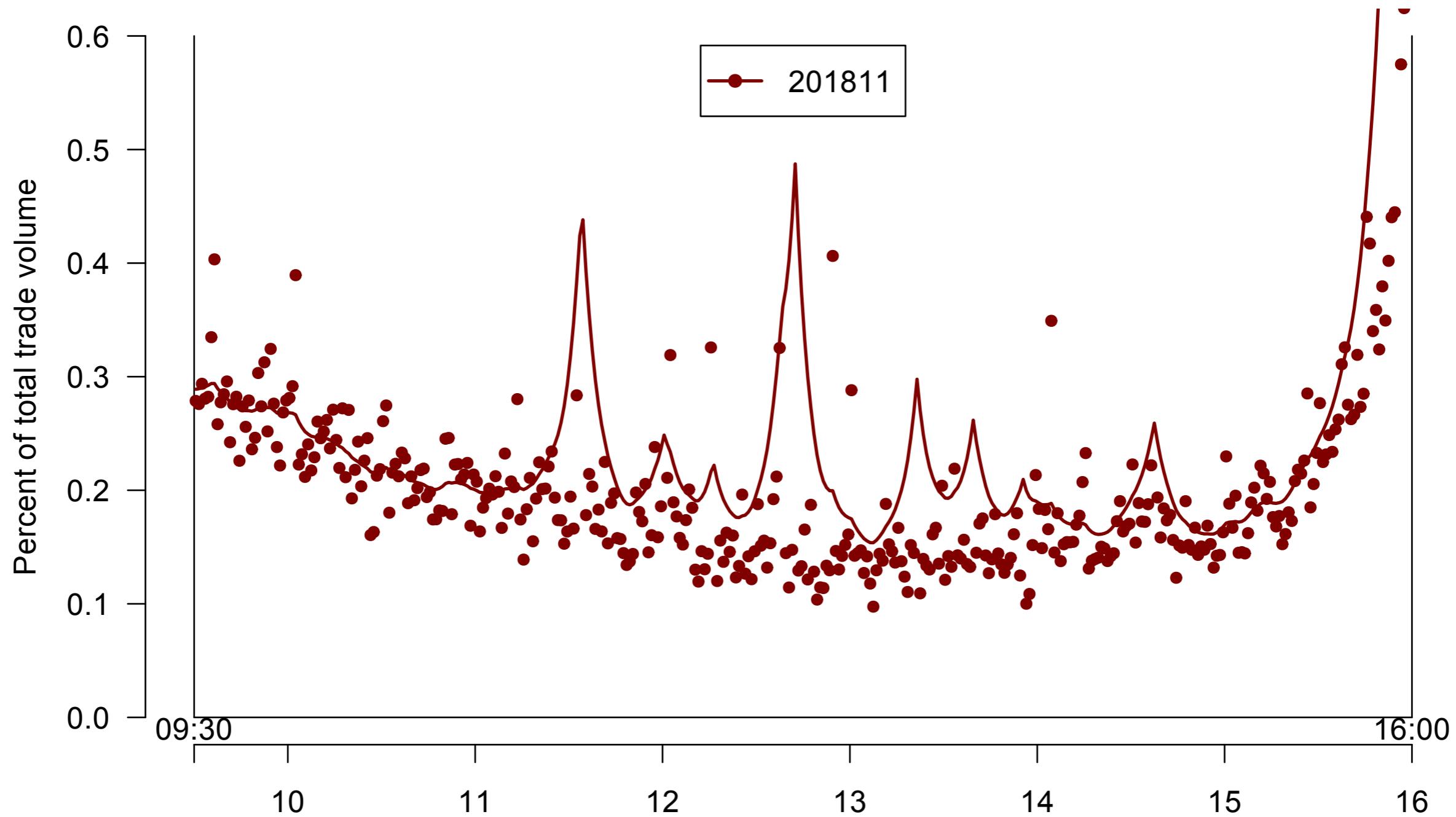


Expanded scale



Percent of total trade volume

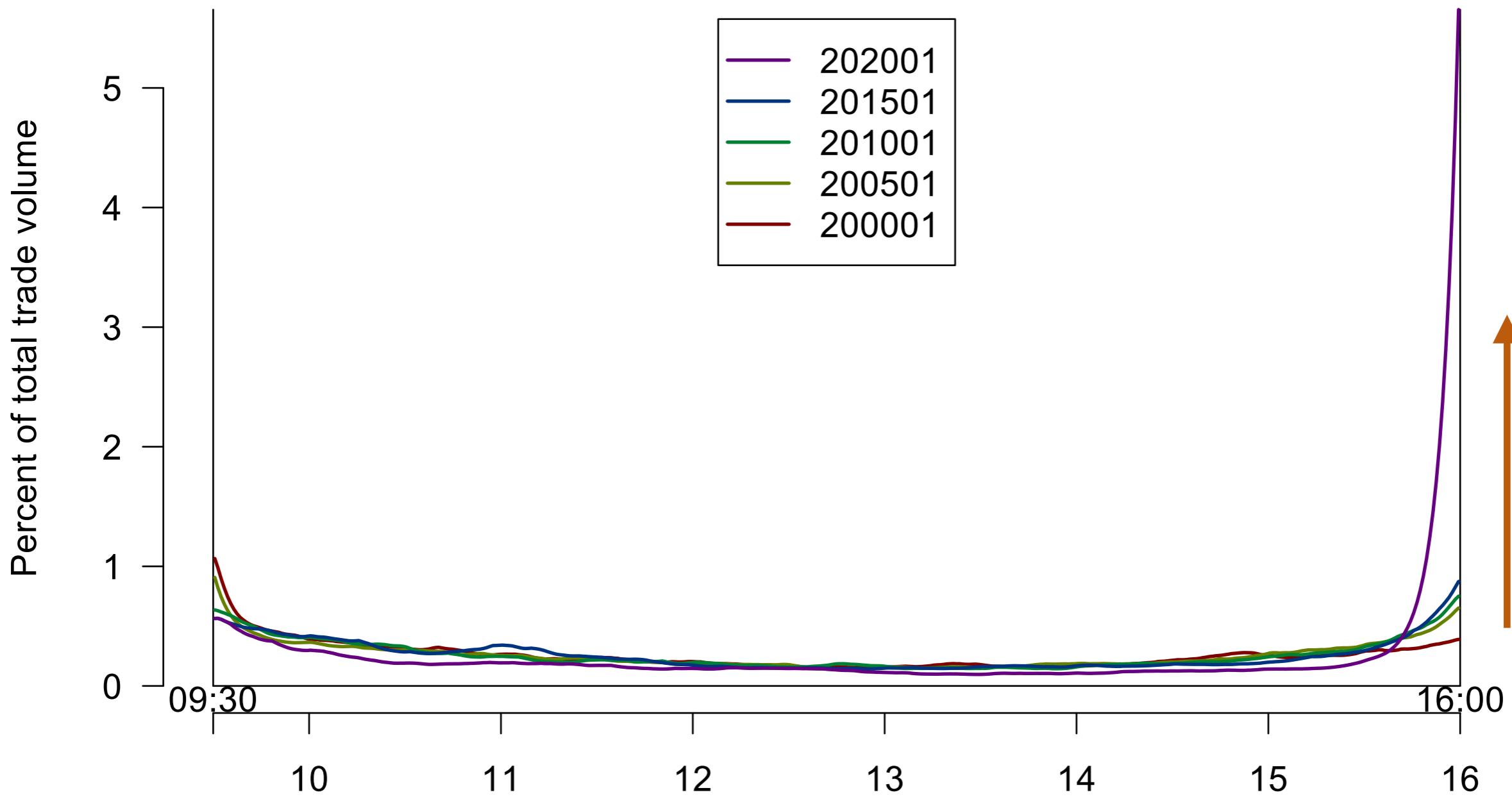




Reasons for U-shaped intraday profile

- Traders are less active in middle of day (lunch?)
- Volume at end since indices are marked to close
- News accumulates overnight

Evolution of close volume over time



Challenges in building real volume curves

- What time period to average over?
Balance statistical accuracy against changes
one month, or longer?
Exponential or window weighting?
- What bin size to use?
1 minute, 15 minutes, smoothing function?
- Combine data from different products?
5000 US equities, not all independent
- How to handle information events?
less important for equities because after hours

Intraday curves for volume and volatility (futures)

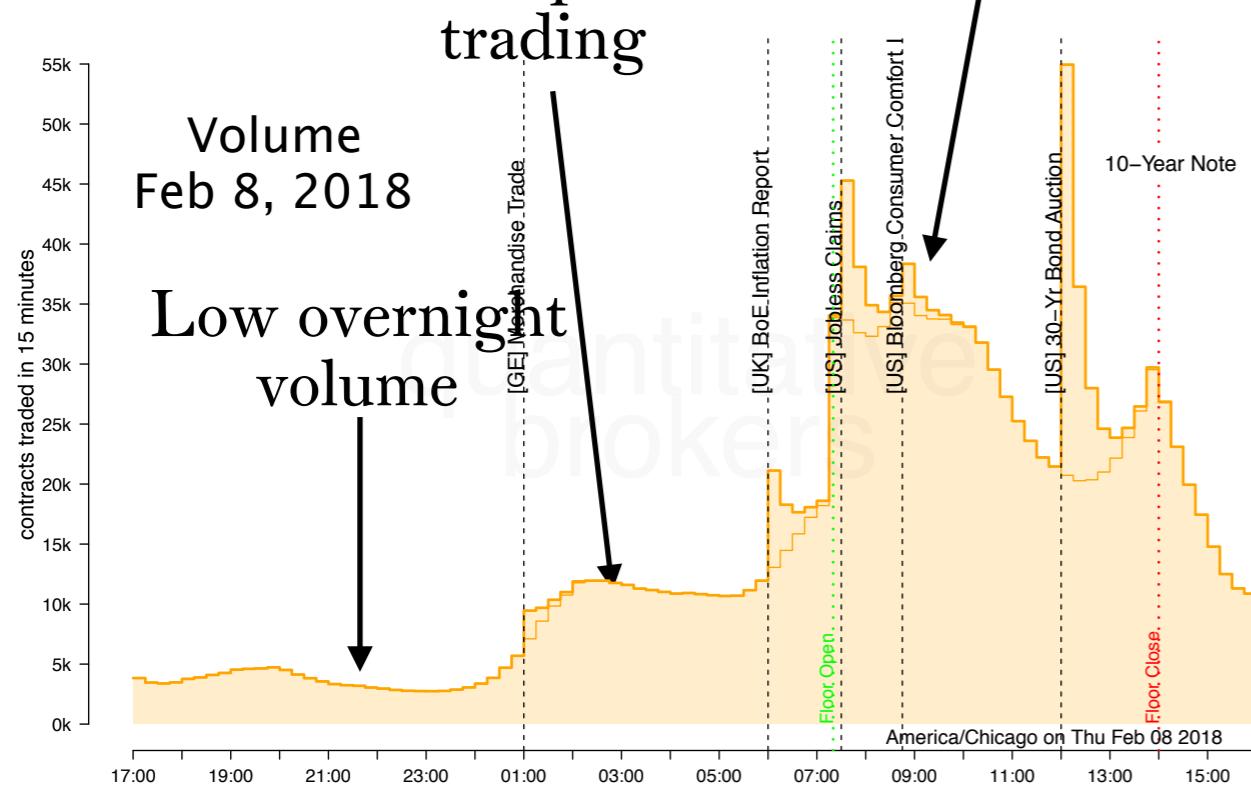
10 -year Treasury note futures

Information events

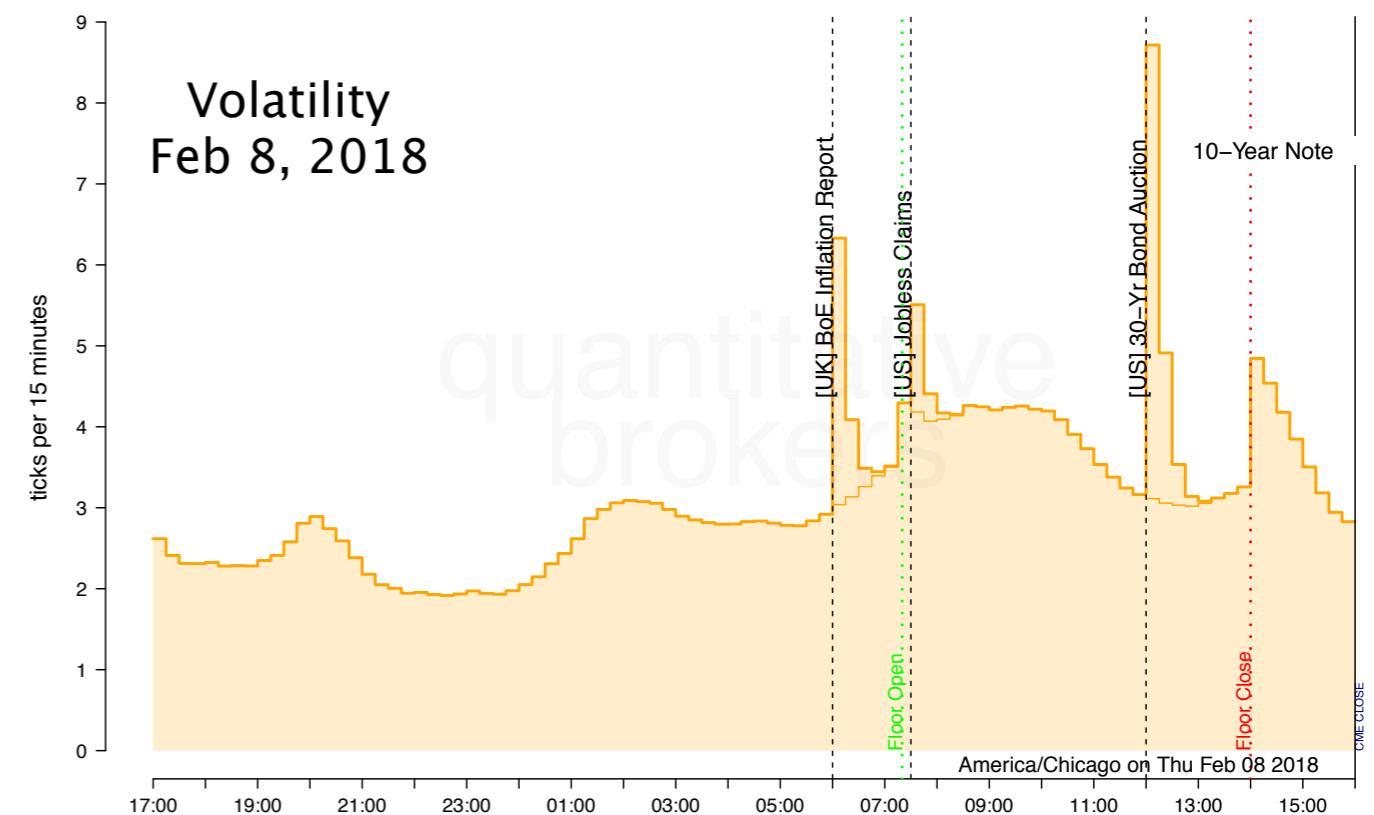
European
trading

Volume
Feb 8, 2018

Low overnight
volume



Volatility
Feb 8, 2018



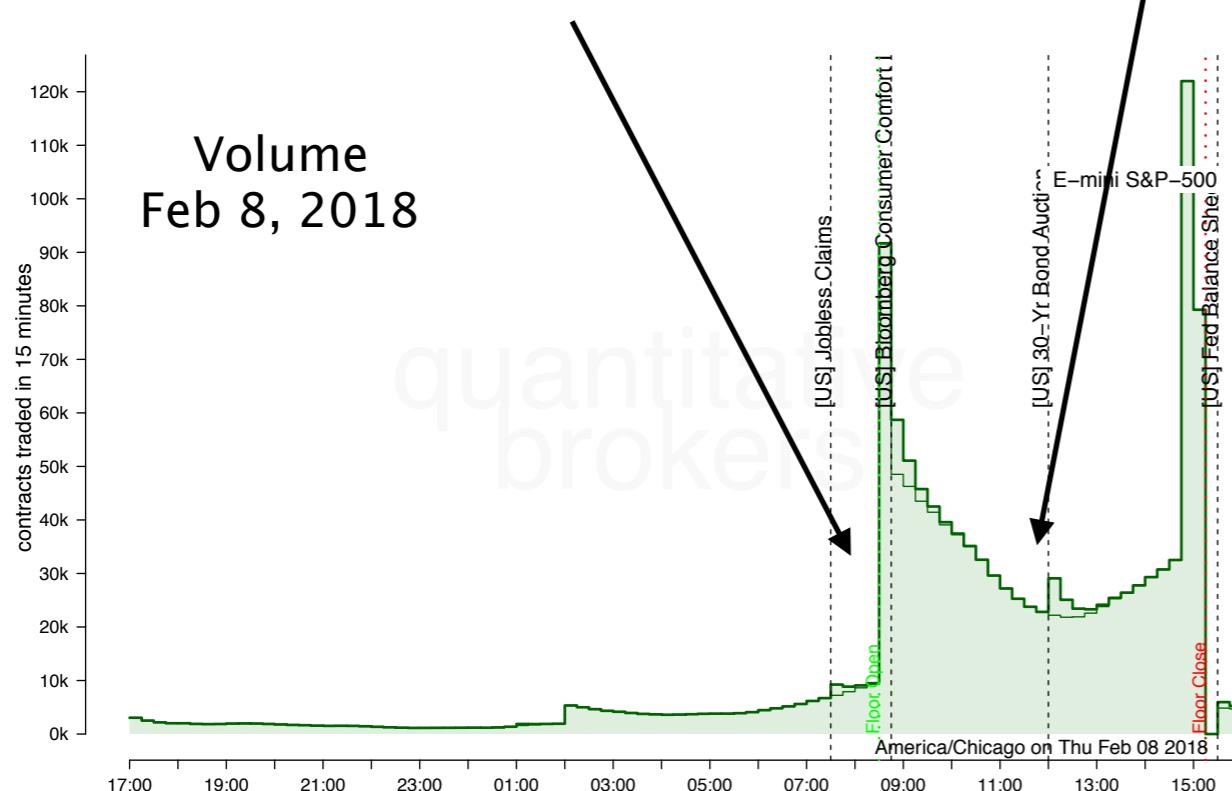
Interest rate futures have extended trading hours
trade through events

Produced by QB from CME market data

Intraday curves for volume and volatility (futures)

S&P 500 equity index futures

Jumps when cash markets open/close
(9:30 to 16:00 NY time)



"U" profile characteristic of equity markets



Produced by QB from CME market data

Challenges in TAQ data

- Practical
 - WRDS interface, .csv.gz files
 - R technology: data.table, fasttime
- Market fragmentation
 - Mix of trades and quotes from different exchanges
- Timestamps
 - Trades and quotes do not match exactly
- No aggressiveness tag
 - Hard to determine "sign" of trades
- Complexity of equity markets
 - What do all the condition codes mean?