PRINCETON UNIVERSITY

ORF 474: High Frequency Trading
Spring 2020
Robert Almgren

Lecture 3b

Feb 19, 2020

Today

- Matching trades and quotes
- Lee-Ready algorithm for trade direction
- Roll model empirical (from last week)
- Roll model theoretical (on board)
- Glosten-Milgrom model (on board)

Matching trades and quotes

- Interleave trade file and quote file
- Look at one exchange at a time
- Expect to see
  - trades at bid or ask
  - quote size decreasing by trade size
- Do not generally see this in data
- Matching trades and quotes in TAQ is frustrating
- Other exchanges give trade sign directly

0.5 sec MSFT on Nasdaq only

468066	2020-01-31	36001.1028766632	Q	3439759	MSFT	2	171.68			171.71	10		R
468068	2020-01-31	36001.2539756298	Q	3439887	MSFT	2	171.69			171.71	10		R
468069	20	4857	Q	3439941	MSFT	2	171.69			171.71	9		R
468073	20	5488	Q	3440114	MSFT	2	171.69			171.71	10		R
468074	20	6204	Q	3440115	MSFT	2	171.69			171.70	1		R
468076	20	3590	Q	297354	MSFT			171.69	10			@	I
468077	20	1105	Q	3440120	MSFT	1	171.68			171.70	1		R
468078	2020-01-31	36001.5656019250	Q	297355	MSFT			171.69	60			@	I
468079	2020	60	Q	297356	MSFT			171.69	200				@
468080	2020	92	Q	3440121	MSFT	1	171.68			171.70	2		R
468081	2020	80	Q	3440122	MSFT	1	171.68			171.70	3		R
468083	2020	85	Q	3440124	MSFT	1	171.68			171.69	1		R
468084	2020-01-31	36001.5656540394	Q	3440125	MSFT	1	171.68			171.69	2		R
468102	2020-01-31	36001.5662071040	Q	297367	MSFT			171.68	21			@F	I
468103	2020-01-31	36001.5662078857	Q	3440201	MSFT	3	171.67			171.69	2		R
468109	2020-01-31	36001.5667243004	Q	3440202	MSFT	3	171.67			171.69	1		R
468110	2020-01-31	36001.5668914800	Q	297368	MSFT			171.68	79			@	I
468117	2020-01-31	36001.5673024654	Q	3440203	MSFT	3	171.67			171.68	1		R
468120	2020-01-31	36001.5673406124	Q	3440204	MSFT	2	171.67			171.68	1		R
468121	2020-01-31	36001.5673694611	Q	3440205	MSFT	2	171.67			171.68	2		R
468128	2020-01-31	36001.5696887970	Q	3440329	MSFT	1	171.67			171.68	2		R
468135	2020-01-31	36001.5713284016	Q	3440337	MSFT	2	171.67			171.68	2		R
468136	2020-01-31	36001.5736122131	Q	3440345	MSFT	1	171.67			171.68	2		R
468137	2020-01-31	36001.5736186150	Q	297379	MSFT			171.67	4			@	I
468138	2020-01-31	36001.5736223220	Q	297380	MSFT			171.67	96			@	I
468139	2020-01-31	36001.5736865997	Q	3440347	MSFT	1	171.66			171.68	2		R
468140	2020-01-31	36001.5736866400	Q	297381	MSFT			171.67	4			@F	I
468141	2020-01-31	36001.5736897950	Q	297382	MSFT			171.67	1			@F	I
468142	2020-01-31	36001.5736926770	Q	297383	MSFT			171.67	15			@F	I
468143	2020-01-31	36001.5737819570	Q	297384	MSFT			171.67	44			@F	I
468144	2020-01-31	36001.5738180790	Q	297385	MSFT			171.67	17			@F	I
468147	2020-01-31	36001.5738563810	Q	297387	MSFT			171.67	24			@	I
468148	2020-01-31	36001.5739059448	Q	3440350	MSFT	1	171.66			171.68	3		R
468149	2020-01-31	36001.5739252567	Q	3440351	MSFT	1	171.65			171.68	3		R
468151	2020-01-31	36001.5740256310	Q	3440352	MSFT	1	171.65			171.68	4		R
468163	2020-01-31	36001.5744607449	Q	3440353	MSFT	1	171.65			171.67	1		R
468164	2020-01-31	36001.5744729042	Q	3440354	MSFT	1	171.65			171.67	2		R
468171	2020-01-31	36001.5745480190	Q	297411	MSFT			171.67	100				@
468181	2020-01-31	36001.5757246017	Q	3440418	MSFT	5	171.65			171.67	1		R
468185	2020-01-31	36001.5759181976	Q	3440421	MSFT	4	171.65			171.67	1		R

Quote sizes in round lots (100 shares)

Trade sizes in shares

Trades at bid
Take out bid
(timing error)

4μsec

Trades within spread

Are these separate trades?

Trade at ask
Takes out ask

6.6 sec CUZ on Nasdaq only

Trade at ask
Ask increases?

Trade at ask
Ask decreases

Trade within spread

12874	2020-01-31	37008.0331091881	T	376882701	CUZ	2	40.98		41.10	2	R
12875	2020-01-31	37008.0334284306	T	376882801	CUZ	2	40.98		41.10	4	R
12876	2020-01-31	37008.0340299606	T	376882901	CUZ	2	40.98		41.10	3	R
12879	2020-01-31	37009.0104875565	T	376988001	CUZ	2	40.98		41.10	2	R
12880	2020-01-31	37010.0022010803	T	377114301	CUZ	2	40.98		41.10	3	R
12881	2020-01-31	37010.0259323120	T	377119101	CUZ	2	40.98		41.10	2	R
12882	2020-01-31	37010.0259364880	T	23534401	CUZ			41.10 10			I
12898	2020-01-31	37014.5996139050	T	377658201	CUZ	2	40.98		41.10	3	R
12899	2020-01-31	37014.5997741222	T	377658301	CUZ	2	40.98		41.09	1	R
12900	2020-01-31	37014.5998470783	T	377658401	CUZ	2	40.98		41.09	2	R
12903	2020-01-31	37014.6001243591	T	377658602	CUZ	2	40.98		41.09	3	R
12904	2020-01-31	37014.6002879143	T	377658701	CUZ	2	40.98		41.09	5	R
12910	2020-01-31	37014.7090132236				2	40.98		41.09	4	R
12911	2020-01-31	37014.7090247230						41.09 100			
12912	2020-01-31	37014.7090327770						41.09 100			
12914	2020-01-31	37014.7090515000						41.09 100			
12915	2020-01-31	37014.7090640068	T	377669201	CUZ	2	40.98		41.09	3	R
12916	2020-01-31	37014.7090640068	T	377669202	CUZ	2	40.98		41.09	2	R
12917	2020-01-31	37014.7090640068	T	377669203	CUZ	2	40.98		41.14	1	R
12918	2020-01-31	37014.7090640068	T	377669204	CUZ	1	41.01		41.14	1	R
12919	2020-01-31	37014.7090640068	T	377669205	CUZ	1	41.03		41.14	1	R
12922	2020-01-31	3701			UZ			41.09 200			
12925	2020-01-31	3701			UZ	1	41.08		41.14	1	R
12926	2020-01-31	37014.7090885639	T	377669602	CUZ	1	41.08		41.10	1	R
12927	2020-01-31	37014.7091057301	T	377669701	CUZ	2	41.01		41.10	1	R
12928	2020-01-31	37014.7091057301	T	377669702	CUZ	2	41.01		41.14	1	R
12929	2020-01-31	37014.7091627121	T	377669801	CUZ	1	41.02		41.14	1	R
12931	2020-01-31	37014.7093636990	T	377670001	CUZ	1	41.02		41.14	2	R

Lee-Ready algorithm for sign identification

The intraday trade and quote data do not identify whether a trade was triggered by a market buy or sell order, so this information must be inferred from the data. Two general approaches have been used to infer the direction of a trade: 1) compare the trade price to adjacent trades (techniques commonly known as "tick tests") or 2) compare the trade price to the bid/ask prices of the prevailing quote.

The tick test is a technique which infers the direction of a trade by comparing its price to the price of the *preceding* trade(s). The test classifies each trade into four categories: an uptick, a downtick, a zero-uptick, and a zero-downtick. A trade is an uptick (downtick) if the price is higher (lower) than the price of the previous trade. When the price is the same as the previous trade (a zero tick), if the last price change was an uptick, then the trade is a zero-uptick. Similarly, if the last price change was a downtick, then the trade is a zero-downtick. A trade is classified as a buy if it occurs on an uptick or a zero-uptick; otherwise it is classified as a sell.

We present evidence that trading inside the spread is due largely to "standing orders" that cause the effective spread to be narrower than the quoted spread. As a result, in a given time span it is not generally true that trades between the spread are equally likely to be buys and sells. In the context of a simple model, we demonstrate that the tick test will correctly classify at least 85 percent of all trades at the midpoint of a spread. For trades closer to the bid or ask we show that the tick test continues to perform well, although a simple assignment of trades as buys (sells), if they are closer to the bid (ask), will also perform well.

Lee-Ready algorithm:

1. Classify trade direction based on quotes
   - either NBBO or each exchange
   - 5 sec delay on quotes, in 1991
2. Where trade price is within spread, use tick test from previous trade