Depth and Volume

Teddy Cho

August 17, 2015

1 Data

We use historical data from the NYSE Trade and Quote database[1], made available by Wharton Research Data Services. The dataset "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"[3].

Our data cover all trades for a sample of symbols[4] over the period of Mar 1, 2014 - Mar 31, 2014. No time filtering is used, and we are left with 20 trading days total.

1.1 Trades Data

The trading data is organized as a sequential collection of trades, where a trade is defined as a transaction with a unique trade sequence number. A trade must be between no more than two parties, in one symbol, at one price, on one exchange, at one time.

1.1.1 Alice Trades Against One Counterparty

Alice buys 50 shares of GOOG for $700 on NASDAQ at 11:00 against a corresponding quote of at least 50 shares. The trades data would show the following:

[1] We note that this data is known to have inaccuracies regarding quotes' timestamps, and such issues may carry over to trade timestamps. These issues stem from timestamps originating from a secondary source (i.e., not the exchanges' order matching engines) and being derived after an exchange aggregation method which introduces further latency[2].

[3] See https://wrds-web.wharton.upenn.edu/wrds/ds/taq/index.cfm.

[4] Our sample of symbols consists of AMD, BAC, C, GOOG, GRPN, JBLU, MSFT, and RAD. We would like to include BRKA and BRKB, but millisecond TAQ data does not exist for them. Strange.

Time	Symbol	Volume	Price	Exchange
11:00:00	GOOG	50	$700.00	NASDAQ

1.1.2 Alice Trades Against Multiple Counterparties

Alice sells 50 shares of GOOG for $700 on NASDAQ at 11:00, but the first-in-line quote is only big enough for 10 shares. The next two quotes, however, are of size 6 an 80, respectively, and so Alica can be accommodated by these three quotes. The trades data would show the following:

Time	Symbol	Volume	Price	Exchange
11:00:00	GOOG	15	$700.00	NASDAQ
11:00:00	GOOG	6	$700.00	NASDAQ
11:00:00	GOOG	29	$700.00	NASDAQ

1.1.3 Alice Trades On Multiple Exchanges

Alice sends a buy order 20 shares of GOOG for $700 on NASDAQ and a buy order 30 shares of GOOG for $700 on NYSE. Furthermore, the first-in-line quote on NYSE is only big enough for 1 share, but the next quote in the NYSE queue is for 100 shares. The trades data would show the following:

Time	Symbol	Volume	Price	Exchange
11:00:00	GOOG	20	$700.00	NASDAQ
11:00:00	GOOG	1	$700.00	NYSE
11:00:00	GOOG	29	$700.00	NYSE

1.1.4 Alice Trades At Multiple Prices

Alice sends a market buy order for 50 shares of GOOG on BATS BYX. The first-in-line quote is to sell 3 shares at $700, and the next quote in the queue is for 16 shares at $700.01. The next quote is for 19 shares at $700.01, and the following quote is for 60 shares at $700.03. The trades data would show the following:

Time	Symbol	Volume	Price	Exchange
11:00:00	GOOG	3	$700.00	BATS BYX
11:00:00	GOOG	16	$700.01	BATS BYX
11:00:00	GOOG	19	$700.01	BATS BYX
11:00:00	GOOG	12	$700.03	BATS BYX

1.1.5 Alice's Order Gets Routed

Alice sends a market sell order for 50 shares of GOOG on BATS BYX. The first-in-line quote is to buy 5 shares at $700, and the next quote in the queue is for 60 shares at $699.98. However, NASDAQ is the national best bid with a size of 1 share at $699.99. By Reg NMS, Alice's sell order gets routed to NASDAQ and she sells 1 share for $699.99. Now, the national best bid is the 10 shares at

$699.98 on BATS BYX. Alice's order transacts her remaining 44 shares against the 60 share bid on BATS BYX. The trades data would show the following:

Time	Symbol	Volume	Price	Exchange
11:00:00	GOOG	5	$700.00	BATS BYX
11:00:00	GOOG	1	$699.99	NASDAQ
11:00:00	GOOG	44	$699.98	BATS BYX

1.1.6 Trades At Different Times

Alice buys 40 shares of GOOG for $700 on BATS BYX at 11:00:00 against two orders of size 13 and 90, respectively. Bob then sells 10 shares of GOOG for $699.97 on NASDAQ at 11:00:04. The trades data would show the following:

Time	Symbol	Volume	Price	Exchange
11:00:00	GOOG	13	$700.00	BATS BYX
11:00:00	GOOG	27	$700.00	BATS BYX
11:00:04	GOOG	10	$699.99	NASDAQ

Since we filter out some special cases (trade corrections, duplicates, etc.), we will not go into that here. [5]

2 Transforming Data

Each data point concerned a specific symbol i's activity on a specific exchange j for a specific date t.

2.1 Volume

Volume Q_{ijt} is computed as the sum of the volumes q of trades occurring between 10:00 and 15:30 in symbol i on exchange j on date t:

$$Q_{ijt} = \sum_{q \in ijt} q$$

2.2 Depth

We begin by computing the depth of each quote update, indexed by u. This is done by taking the size of the best bid q_{bu} and the size of the best ask q_{au} and taking their average $\overline{q_u}$:

$$\overline{q_u} = \frac{q_{bu} + q_{au}}{2}.$$

For each quote update u, we compute its duration ΔT_u–that is, how long this

[5] www.nyxdata.com/doc/224904

quote update u represents the actual state of the book. We do this by subtracting the timestamp T_u of quote update u from the timestamp T_{u+1} of the next quote update $u+1$:

$$\Delta T_u = T_{u+1} - T_u$$

.
To summarize the depth D_{ijt} between 10:00 and 15:30 in symbol i on exchange j on date t, we take the sum of $\overline{q_u}$ weighted by their corresponding ΔT_u.

$$D_{ijt} = \sum_{u \in ijt} q_u \cdot \Delta T_u$$

John mentioned how NASDAQ quotes effectively treat its different DMMs as different exchanges; make sure to do special logic for NASDAQ.

3 Results

Due to wide heterogeneity in volumes between symbols, we take the logarithm of both depth and volume. We try
Chicago, National, NASDAQ PSX OMX had negative values for $\log D_{ijt}$, which is because. ???????????

Graphs are located in Dropbox[6].

[6] https://www.dropbox.com/sh/9m485w2yvkczkle/AACP6PtrgDxGnU7Vd8I2ofYla?dl=0

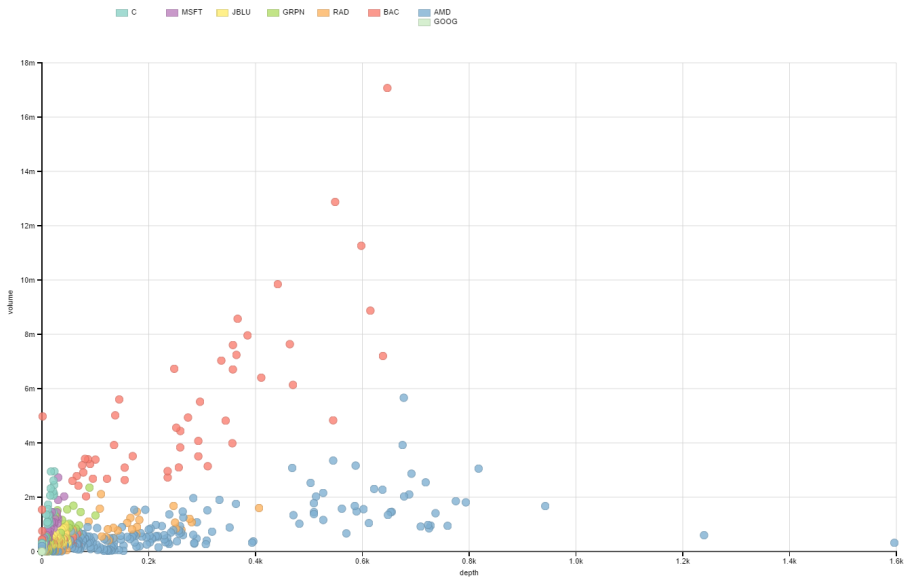

Figure 1: The values of daily depth and volume are plotted against each other across symbols. The colors correspond to different exchanges.

Figure 2: The above graph, logged.