Liquidity Measures Accounting for Withdrawn and Cancelled Quotes: A Comparative Analysis Between Thomson Reuters Time of Sales and NYSE Daily Trade and Quote

by

Han Wei Chee

An essay submitted to the Department of Economics in partial fulfillment of the requirements for the

degree of Master of Arts

Queen's University

Kingston, Ontario, Canada

September 2015

Acknowledgements

I would like to express my great appreciation to Dr. Ryan Riordan for his role as my supervisor and for suggesting such an interesting research topic. His willingness to spare some of his much of limited free time has been very much appreciated.

I would also like to thank Saad Khan for all the helps he had provided. His assistance in data retrieval, concepts clarifications and programming will never be forgotten.

Finally, I wish to extend my thanks to all the support staff at Thomson Reuters for their efficient and helpful services.

Table of Contents

I. Introduction

Electronic trading has long been part of exchanges around the world. Whether it is something as simple as automation from the beginning of the century, or something as sophisticated as algorithm trading to the more advanced high frequency trading (hereafter HFT and will be used interchangeably to represent high frequency trader). Depending on the source, between 2009 and 2012, HFT accounted for somewhere between 40% of NASDAQ volume and 70% of the volume in U.S. equity markets.[1] Regardless of the exact figure, HFT's presence in the U.S. market is no doubt prevalent. However, the debate as to whether HFT is a good or bad innovation to the overall social welfare is still ongoing.

In general, the proponents for HFT argue that HFT increases liquidity of the market, improves the price discovery process and further postulate that prices are more informative and finally argue that HFT reduces the intraday volatility. On the other hand, those who are against it claimed the HFT may compete for liquidity when rebalancing their positions, the price impact of liquidity trades is higher in the presence of HFT[2] , price volatility would actually increase, and the arm race in speed has no social benefit.

In order to better evaluate the true impact of HFT on social welfare, one needs to consider the various strategies employed. Though numerous strategies are known to be at

[1] In 2009, HFT accounted for 70% of the volume in U.S. equity market, and 50% of the volume in futures markets (Iati and CFTC as cited in Easley et al., 2012).
In 2009, HFT accounted for 40% of NASDAQ volume (Hirschey, 2013).
Tabb Group estimated in 2010 HFT made up 56% of equity in the U.S. (Biais and Woolley, 2011).
According to MacIntosh (2013), 50% of equities trading in the U.S. came from HFT.

[2] HFT reduces (increases) the prices that liquidity traders receive when selling (buying) their equity holdings (Cartea and Penalva, 2011).

play, for the purpose of this paper, the focus will be on *automated liquidity provision, quote stuffing* and *quote spoofing*.

Automated liquidity provision requires traders to hold both long and short position in the market in order to smooth out temporary imbalance in the supply and demand. Through rapid order placements and cancellations of bid and offer limit orders, traders are able to derive profits from the bid-ask spreads. As can be seen in this case, the market makers' profit is positively related to the size of the spread.[3]

By continuously updating their position, the order cancelation and submission to transaction ratio is high (Jones as cited by Goldstein et al. 2014).[4] Given an effective and efficient algorithm and low latency connection, this strategy has fairly low risk. This probably explains why the majority of the HFT volume and over 80% of HFT limit order submission on the NOMX-St are associated with market making strategies (Hagströmer and Nordén 2012).

Quote spoofing however, refers to an act of placing sequences of orders against the true intention. An increase in limit order to buy could lead other naïve traders to begin buying. The upward pressure pushes the price of the target asset up whereby the HFT will then sell at an over-inflated price while simultaneously cancelling all previous *spoofed* limit orders before they get executed (Biais and Woolley, 2011). Nevertheless, market manipulation is illegal and the act of *spoofing* is considered a deliberate act to manipulate

[3] As an illustration, if stock XYZ has a bid to ask of $100 to $100.05. The market maker can purchase 1,000 units of the stock at $100 and immediately sell all of them at $100.04 (if there are sufficient demands for all of them) and make a profit of $40 per transaction. Now imagine repeating the same process 10,000 times each day. However, the stock volatility must be low and transactions carried out fast enough that the value of the stock does not fall below $100.

[4] SEC (2010) notes that passive market making results in cancelation rate of 90% or more.

the market. Gai et al. (2012) tested the data they obtained from NASDAQ for the trading day of May 24, 2010 on their three models; a factor regression model, a discontinuity test and diff-in-diff test. All three models produce significant results supporting the presence of *quote stuffing*.

On the other hand, *quote stuffing* is carried out with the intention to congest the market so that it hinders or delays slower traders' ability to understand the market situation (Biais and Woolley, 2011). Just as *spoofing*, this technique requires traders to place a large amount of limit order while cancelling or withdrawing them before they get executed. However, the difference with this techniques is that the limit order could be a combination of buys and sells.

Based on 118 randomly selected stocks listed on NASDAQ and the NYSE between March 19, 2010 and June 7, 2010, Gai, Yao and Ye (2012) found evidence consistent with *quote stuffing*. Furthermore, all the stocks within their sample had a cancellation/execution ratio of above 90% and of all the quotes that were cancelled within one second, 30% of them were cancelled within 5 milliseconds. These figures should not be taken lightly. By not accounting for withdrawn/cancelled orders, one may introduce an upward bias to their measures of market liquidity.

Holden and Jacobsen (2014, HJ) highlighted various error in measurements when comparing several liquidity measures obtained from the NYSE *Monthly* Trade and Quote (MTAQ) database and the NYSE *Daily* Trade and Quote (DTAQ) database. They further postulate these errors are due to failure to account for withdrawn quotes, the limitation of the MTAQ second time stamps against the DTAQ millisecond time stamps, and several other causes including cancelled quotes. They concluded that the DTAQ is the first-best when

attempting to evaluate market quality while MTAQ the second-best – after adjusting for withdrawn quotes, deleting nonsensical states, and applying their Interpolated Time technique.

Theoretically, the methods and steps used by HJ will account for all cancelled, withdrawn and erroneous quotes and trades. As such, if one were to follow their guidelines accordingly, regardless of data set, the final result should be a *cleaned* data set – free from cancelled, withdrawn and erroneous quotes and trades.

Therefore, it is the attempt of this paper to replicate the steps outlined by HJ on a different sample obtained through DTAQ and compares the results obtained to those from another sample obtained from the Thomson Reuters Tick History (TRTH) file obtained through SIRCA. Thomson Reuters populates the Tick History file based on feeds obtained from NYSE. On that note, if one was to undergo the exact *cleansing* procedures[5] on both data set, the end results should be identical. Once both the samples have been rid of cancelled, withdrawn and erroneous quotes and trades, various standard measures of market liquidity will be used to determine if they both yield the same results, with measures obtained from DTAQ acting as the base points.

The paper is arranged as follows. The next section provides detailed descriptions on the different measures used. Section III describes the data used, Section IV describes the methodology employed. Section V presents the initial results obtained while the subsequent section analyzes the data used in detail. The following section then concludes.

[5] The SAS programme provided on the authors' website was modified and used to clean the DTAQ millisecond data set used for this paper. However, much of the fields found in DTAQ are not applicable to TRTH. Subsequently, several steps were omitted and modified. These steps will be discussed in details in the methodology section of the paper.

II. The Different Measures

The first measure begins with analyzing trade locations which are divided into trades that are at, inside, and outside the NBBO. As described by HJ, the k^{th} trade is considered *At the NBBO* if $P_k = A_k$ or $P_k = B_k$, where P_k is the price the k^{th} trade is traded at and both A_k and B_k are the National Best Ask and National Best Bid respectively; assigned to the k^{th} trade based on a particular technique to be explained in the methodology section of the paper. Following similar notation, a trade is considered to be *Inside the NBBO* when $A_k > P_k > B_k$ while a trade is considered to be *Outside the NBBO* when $P_k > A_k$ or $P_k < B_k$. Since it would be illogical for any particular buyer (seller) to bid (ask) at a price higher (lower) than the National Best Ask (Bid), both the *At the NBBO* and *Inside the NBBO* should not be too much of a concern. On the other hand, it would be rather disturbing to observe trades that are *Outside the NBBO*. Hence, the technique used to match trades and quotes must be one that minimizes the percentage of *Outside the NBBO*[6].

Aside from the conditions mentioned above, situations may arise to reflect improbable market conditions if withdrawn or cancelled quotes are not accounted for. If a trade occurs when $A_k < B_k$, the trade happens when the market displays a *Crossed NBBO*. Whereas if the trade occurs when $A_k = B_k$, the trade is said to take place when the market displays a *Locked NBBO.* Since the SEC regulations require national exchanges to suppress quotes that indicates a locked market, trades which take place under such condition can be viewed as an implication of unaccounted withdrawn or cancelled quotes. However, a locked

[6] This paper finds the percentage of trades *Outside the NBBO* when matched to quotes observed a millisecond before and the percentage of those matched to the last observed quotes to be not statistically different. The rest of the analysis proceeds by using trades matched with the last observed quotes.

market is surely not as dire as a crossed market since the latter represents a clear arbitrage opportunity.[7] Thus, the percentage of *Crossed NBBO* should be given much consideration.

The analysis proceeds to evaluate the quoted and effective spreads. Quotes are evaluated at every 5 minutes interval. The dollar and percent quoted spread are defined as

$$Dollar\ Quoted\ Spread_i = A_i - B_i \tag{1}$$

$$Percent\ Quoted\ Spread_i = \frac{A_i - B_i}{M_i} \tag{2}$$

$$where\ A_i = National\ Best\ Ask\ at\ interval\ i$$

$$B_i = National\ Best\ Bid\ at\ interval\ i$$

$$M_i = \frac{A_i + B_i}{2}$$

The *Dollar (Percent) Quoted Spread* is the mean of *Dollar (Percent) Quoted Spread*$_i$ over all the time intervals.

The dollar and percent effective spread on the k^{th} trade are defined as

$$Dollar\ Effective\ Spread_k = 2D_k(P_k - M_{k)} \tag{3}$$

$$Percent\ Effecetive\ Spread_k = \frac{2D_k(P_k - M_k)}{M_k} \tag{4}$$

$$where\ D_k = +1\ if\ the\ k^{th}\ trade\ is\ a\ buy\ and$$

$$-1\ if\ the\ k^{th}\ trade\ is\ a\ sell$$

$$P_k = price\ of\ the\ k^{th}\ trade$$

$$M_k = midpoint\ on\ the\ NBBO\ quotes\ matched\ to\ the\ k^{th}\ trade$$

[7] Imagine a scenario where stock XYZ has an Ask Price of $1.00 in Exchange A and a Bid Price of $1.10 in Exchange B. An arbitrageur could take a short position on stock XYZ in Exchange B and uses the proceeds to purchase the stock from Exchange A to cover the short position in Exchange B. The arbitrageur profits $0.10 from each unit of stock he shorts with no risk.

The *Dollar (Percent) Effective Spread* is the dollar-volume-weighted average of *Dollar (Percent) Effective Spread*$_k$ over all trades.

In order to compute the dollar and percent realized spread, a trade needs to be identified as a buyer initiated or seller initiated. Similar to HJ, this paper uses the three trade-typing conventions mentioned in their paper: Lee and Ready (1991, LR) convention, Ellis, Michaely, and O'Hara (2000, EMO) convention, and Chakrabarty et al. (2006, CLNV) convention.

According to the Lee and Ready (1991) identification criteria, when $P_k > M_k$ the trade is classified as a buy order whereas when $P_k < M_k$ the trade is classified as a sell order. In the event that $P_k = M_k$, the tick test is used. Based on the test, a trade direction falls into one of the four groups: an uptick, a downtick, a zero-uptick and a zero-downtick. When the price of a trade is higher (lower) than the last observed trade price, it is considered as an uptick (downtick). However, in the event that the trade price is the same as the last observed trade price, then the trade is categorized as a zero tick. Subsequent to this, the last price change will need to be considered. If it was previously an uptick (downtick), then the trade is categorized as a zero-uptick (zero-downtick). A trade is then classifieds as a buy if it happens on an uptick or a zero-uptick – a sell otherwise (Lee and Ready, 1991).

The trade-typing convention proposed by EMO indicates that, "All trades executed at the ask quote are categorized as buys. All trades executed at the bid quote are categorized as sells. All other trades are categorized by the tick rule." The effectiveness of the EMO convention is similar to the LR convention when identifying trades that are *At the NBBO*, but are more effective when identifying trades that are *Outside the NBBO* and *Inside the NBBO* (EMO, 2000).

The CLNV convention suggests that, "a trade is a buy when $P_k \in [0.3B_k + 0.7A_k, A_k]$, a sell when $P_k \in [B_k, 0.7B_k + 0.3A_k]$, and the tick test is used otherwise" (HJ, 2014). CLNV shows that their trade-typing convention provides a more accurate measure of effective spreads and price impacts, compared to both LR and EMO.

Following the same trade classification criteria as explained above, the dollar and percent realized spread for the k^{th} trade are defined as

$$Dollar\ Realized\ Spread_k = 2D_k(P_k - M_{k+5}) \tag{5}$$

$$Percent\ Realized\ Spread_k = \frac{2D_k(P_k - M_{k+5})}{M_k} \tag{6}$$

$$where\ M_{k+5} = the\ midpoint\ 5\ minutes\ after\ the\ the\ midpoint\ M_k$$

The *Dollar (Percent) Realized Spread* is the dollar-volume-weighted average of the *Dollar (Percent) Realized Spread*$_k$ over all trades.

Using similar notations as described above, the dollar and percent price impact on the k^{th} trade are defined as

$$Dollar\ Price\ Impact_k = 2D_k(M_{k+5} - M_k) \tag{7}$$

$$Percent\ Price\ Impact_k = \frac{2D_k(M_{k+5} - M_k)}{M_k} \tag{8}$$

The *Dollar (Percent) Price Impact* is the dollar-volume-weighted average of the *Dollar (Percent) Price Impact*$_k$ over all trades.

The final measure to be considered is *returns* and is defined as

$$Returns_t = \ln(P_t) - ln(P_{t-5})$$

$$where\ P_t = trade\ price\ at\ time\ t$$

$$P_{t-5} = trade\ price\ 5\ minutes\ earlier$$

Returns is the average of *Returns*$_t$ over all time interval within each trading day.

III. Data

The data are obtained from DTAQ and TRTH for the period between 2 Jan 2008 to 14 Jun 2015. Since the purpose of this paper is to compare the results obtained from the two data sets, the sample period should not have any significant impact on the outcome,[8] while the analysis on one stock is sufficient.

Though GE, General Electric Co, is randomly selected for this analysis, it meets all the criteria similar to those used by HJ. Namely, (i) it is a common stock; (ii) it is present on the TAQ and TRTH master file for the first and last date of the sample period; (iii) its primary listing is on the NYSE; and (iv)it did not change primary exchange, ticker symbol, or CUSIP code during the sample period.

Three files are downloaded, Daily Trades File, Daily Quotes File and Daily NBBO Files, from DTAQ through Wharton Research Data Services (WRDS). The three millisecond files are downloaded under the following filter: (i) stock symbol equals to 'GE'; and (ii) the time period is between 2 Jan 2008 and 14 Jun 2015.

On the other hand, the Time of Sales file from TRTH database is downloaded through SIRCA. Aside from the two filters applied above, it must be noted that the option to display correction/cancellation is provided when retrieving data from the site. This is the similar option as provided by Thomson Reuters, the "Apply Corrections and Cancellations"[9] option, when assessing the data from their database.

[8] Any event that would impact the performance of a stock and subsequently its market value will be identically recorded in any data set. Hence, the impact should be identical.

[9] Thomson Reuters Tick History – Cancellation and Correction User Guide issued on 06-Dec-2012, provided a detailed description on how the "Apply Corrections and Cancellations" option works. With the option set off (as per default setting) any correction will creates a new observation with the *Type*

IV. Methodology

A complete SAS code used by HJ to obtain the liquidity measures they used in the paper for the *Monthly* Trade and Quote database is provided in both of the authors' websites. This paper goes through the steps highlighted in their code, while making necessary adjustments, to compute the various measures. The complete steps for each data sets (Time of Sales, Daily Quotes File, Daily NBBO File, and Daily Trades File) and the construction on the Complete Official NBBO are as follow:

Time of Sales

1. Split the data set into two, a data set consisting of only trades and another consisting of only quotes.
2. **Quotes data set**
    a. Remove quotes that are outside normal market hours – prior to 9:00 a.m. and after 4:00 p.m.
    b. Quotes that have been halted is to be removed.[10]

field displaying 'Correction' and the *New Price* and/or *New Volume* fields displaying the correct value. However, with the option opted, both the *New Price* and *New Volume* fields are replaced with *Original Price* and *Original Volume*. In the event there is correction, the corrected value will be filled inserted directly to the *Price* and/or *Volume* fields. The *Original Price* and *Original Volume* fields will then be filled with the initial erroneous value.

[10] For identifying halted quotes refer to footnote 12. While the condition of a quote is identified under the *quote condition* field in DTAQ, such information is stored under the *qualifiers* field in TRTH. Attempts have been made to obtain detailed descriptions of the *qualifiers* field from Thomson Reuters but was told the information pertaining to the request is limited to registered users only. Based upon the format of how the field is populated, the code "where Qualifiers not contains ';A'" has been used in SAS to identify quotes with an "A" condition. The letter "A" is replaced accordingly to represent the other values highlighted in footnote 12.

c. Remove cases of *Crossed Market* (within the same market); delete if both bid and ask come from the same exchange and the bid price is higher than ask price.

d. When an observation consists of an ask price of 0 and a bid price greater than 0, such observation is assumed to be an update indicating the seller has withdrawn the order. Such observation is to be deleted.

e. Compute the spread[11] and midpoint[12] for each observations.

f. Delete quotes where the spread is greater than $5.00 and an ask price greater than the previous midpoint plus $2.50.

g. Delete quotes where the spread is greater than $5.00 and a bid price less than the previous midpoint minus $2.50.

h. Observation with a spread greater than $5.00 and both ask and bid price greater than zero is to be deleted as well.

i. Observations with the ask price, ask size, bid price or bid size being 0 or missing are to be removed.

j. If multiple observations exist within a given millisecond, only the last observation will considered.

k. Retain only *new* quotes. A quote is considered *new* if one of the following variables is different from those of quote observed immediately before it: date, ask price, ask size, bid price, and bid size.

[11] $Spread = Ask\ Price - Bid\ Price$

[12] $Midpoint = \frac{Ask\ Price + Bid\ Price}{2}$

l. The quotes data set is clear of cancelled/withdrawn or erroneous quotes at this point.

3. **Trades data set**
    a. Remove trades that occurred outside of normal market hours – prior to 9:30 a.m. and after 4:00 p.m.
    b. Delete observations where the price and/or volume equal to 0 or missing.
    c. Assuming the Time of Sales file downloaded through SIRCA has already adjusted for all correction and cancellation, the trades data set should now be clear of cancelled/withdrawn or erroneous trades at this point.
4. Both trades and quotes data sets are now ready to be merged.
5. Once merged, sort the data set by date them time.
6. Each trade observed are then matched with the quote immediately before it.
7. The Time of Sales file is now *complete* and is ready to be matched to the *cleaned* quotes data set for computing the measures described in section II.

Daily Quotes File

1. Remove quotes that are outside normal market hours – prior to 9:00 a.m. and after 4:00 p.m.
2. Maintain only quotes with normal quote conditions, where trading has not been halted.[13]

[13] Quotes with nonnormal conditions are those with *quote conditions* field taking up the value A, B, H, O, or W. Descriptions of this values can be found in page 19 of the version 2 of the Daily TAQ Client Specification manual. HJ excluded quote condition with a value 'R', representing 'regular, two-sided open quotes'. However, excluding such quotes would leave the sample with less than 1% of its original number of observations. Hence, such quotes are not removed instead.

3. Delete quotes where its *Quote Cancel/Correction* field equals to the value 'B'.[14]
4. Repeat step (b) through step (h) as outlined for cleaning the time of sales – quotes data set.
5. Retain only quotes where the *National BBO Indicator* is equal to "1" or the *NASDAQ BBO Indicator* is equal to "4".[15]
6. Following step 5, quotes remaining are those representing the NBBO. As such, the initial *Ask Price, Ask Size, Bid Price* and *Bid Size* should naturally be renamed as *Best Ask Price, Best Ask Size, Best Bid Price* and *Best Bid Size*, respectively. The variables must be renamed to match the name given to similar variables found in the NBBO file.
7. Retain only *new* quotes. A quote is considered *new* if one of the following variables is different from those of quote observed immediately before it: date, ask price, ask size, bid price, and bid size.
8. The Daily Quotes file is clear of cancelled/withdrawn or erroneous quotes at this point and is ready to be merged with the *cleaned* Daily NBBO file to construct the Complete Official NBBO.

[14] While HJ find the *Quote Cancel/Correction* field to be always blank in their sample, the sample used for this paper consists of instances where the field is equal to "A" – where "A" means "Not a Cancel Quote". Though the field can take up a value of "A", "B", or "C", the majority of the field are blanks. A value of "B" represents "Cancel Quote/Cancel Price Indication/Cancel Trading Range Indication." Further clarification on the reason behind these unpopulated fields will need to be addressed. Unfortunately, the support team at WRDS was not able to provide an answer, while attempts to reach NYSE Market Data support team were futile.

[15] *National BBO Indicator* equals to "1" means "Quote Contains all National BBO Information – Current quote is itself the new National BBO. No National appendage is required."
NASDAQ BBO Indicator equals to "4" means "Quote Contains all NASD BBO Information – Current quote is itself the new NASD BBO. No NASD appendage is required. This filter is not necessary for the sample since GE is a NYSE stock and not NASDAQ's. Nevertheless it is added as a generalization.

Daily NBBO File

1. Remove quotes that are outside normal market hours – prior to 9:00 a.m. and after 4:00 p.m.
2. Delete quotes where its *Quote Cancel/Correction* field equals to the value 'B'.
3. For instances where Best Ask (Bid) price or size is equal to 0 or missing, set the Best Ask (Bid) size and price to missing. The Ask (Bid) order is assumed to have been withdrawn while the Bid (Ask) order is assumed to be valid.
4. Retain only *new* quotes. A quote is considered *new* if one of the following variables is different from those of quote observed immediately before it: date, ask price, ask size, bid price, and bid size.
5. The Daily NBBO file is clear of cancelled/withdrawn or erroneous quotes at this point and is ready to be merged with the *cleaned* Daily Quotes file to construct the Complete Official NBBO.

Daily Trades File

1. Remove trades that are outside normal market hours – prior to 9:30 a.m. and after 4:00 p.m.
2. Retain only trades with a *Trade Correction Indicator* equals to "00"[16], trade price of greater than 0 and a *Sale Condition* not equals to "T"[17].

[16] *Trade Correction Indicator* equals to "00" means "Regular trade which was not corrected, changed or signified as cancel or error."
[17] *Sale Condition* equals to "T" means "Extended Hours Trade". This condition was not indicated by HJ, however it is the opinion of this paper that it is a necessary condition to ensure trades outside of the normal market hours are filtered out.

3. The Daily Trades File is clear of cancelled/withdrawn or erroneous trades at this point and is ready to be matched with the quotes from the Complete Official NBBO.

Constructing the Complete Official NBBO[18]

1. Merge both the cleaned DTAQ Quotes file and DTAQ NBBO file.
2. Sort the merged data set based on date, time and quote sequence number.
3. Aside from having quotes that are not found in the DTAQ NBBO file, the DTAQ Quotes file is deemed to include quotes that are found in the DTAQ NBBO file. Remove these duplicates observation based on step 2 sorting criteria.
4. The complete official NBBO data set is ready to be used to be matched with the trades from the Daily Trades file.

[18] In instances when the best bid and best offer are found within the same exchange, the official SIP NBBO will record such quotes in the DTAQ Quotes file instead of the DTAQ NBBO file. As such, the DTAQ NBBO file is missing such quotes. Hence, the DTAQ Quotes file is merged with the DTAQ NBBO file to construct a complete official NBBO data set.

V. Initial Results

Table 1: Overall Measures Difference between TRTH and DTAQ

The following table compares the various measures computed from data sets obtained from both the TRTH and DTAQ file after undergoing the steps suggested by HJ for the period between 2nd January 2008 and 14th June 2015. * indicates non-statistically different from DTAQ at 5% level. t indicates the value is non-statistically different due to large difference in standard deviation between the two values. Statistical difference is based on the mean comparison test.

	TRTH in milliseconds		DTAQ in milliseconds	
Panel A: Trade Location				
At the NBBO	93.59%		80.72%	
Inside the NBBO	4.10%		16.63%	
Outside the NBBO	2.34%		2.66%	
Crossed NBBO	0.39%*		0.33%	
Locked NBBO	11.69%		11.13%	
Panel B: Quoted and Effective Spreads				
Dollar Quoted Spread	0.01*t		1.92	
Percent Quoted Spread	0.0005%		0.0007%	
Dollar Effective Spread	0.01		0.08	
Percent Effective Spread	0.0006%		0.0033%	
Panel C: Realized Spread and Permanent Price Impact				
Percent Realized Spread: LR	0.0003%		0.0006%	
Percent Realized Spread: EMO	0.0003%		0.0005%	
Percent Realized Spread: CLNV	0.0003%		0.0005%	
Percent Price Impact: LR	0.0003%		0.0041%	
Percent Price Impact: EMO	0.0003%*		0.0004%	
Percent Price Impact: CLNV	0.0003%*		0.0005%	
Panel D: Returns				
	Mean	Std. Dev	Mean	Std. Dev
Returns (compared to 5 minutes prior)	-0.0002%*	0.00028	-0.00020%	0.00042
Returns (compared to 10 minutes prior)	-0.00041%*	0.00055	-0.00041%	0.00071

Aside from having the percentage of trades happening at *Crossed NBBO*, *Dollar Quoted Spread*, *Percent Price Impact* (based on EMO and CLNV tick direction identification) and *Returns* being non-statistically different between the two data sets, all the other measures are statistically different at the 5% level. Furthermore, aside from measures of trade location, all the liquidity measures obtained from TRTH that are statistically different and are lower as well.

While the effective spread is calculated based on the LR trade-typing convention, both percent realized spread and percent price impact are calculated based on the three trade-typing conventions mentioned earlier. Of these, only the percent price impact calculated using the EMO and CLNV trade-typing convention are not statistically different at the 5% level for both the sample. While CLNV found evidence that their algorithm provides a much accurate results when computing price impact, the results obtained by this paper are not able to conclusively support their finding.

The results are most disappointing at this point. In order to better understand why the results are not similar for both data sets, this paper proceeds to compare the various measures on yearly basis. The results are presented in table 2.

Table 2: Overall Measures Difference between TRTH and DTAQ Evaluated on a Year-by-Year Basis

* indicates non-statistically different from DTAQ at 5% level.

	2008		2009		2010		2011	
Panel A: Trade Location								
	TRTH in milliseconds	DTAQ in milliseconds	TRTH in milliseconds	DTAQ in milliseconds	TRTH in milliseconds	DTAQ in milliseconds	TRTH in milliseconds	DTAQ in milliseconds
At the NBBO	91.77%	89.07%	90.83%	88.11%	97.85%	82.75%	97.49%	81.66%
Inside the NBBO	2.52%	5.12%	6.35%	8.67%	0.04%	14.30%	0.09%	16.21%
Outside the NBBO	5.71%*	5.80%	2.84%	3.22%	2.11%	2.95%	2.43%	2.13%
Crossed NBBO	1.01%*	0.96%	0.44%*	0.39%	0.29%*	0.27%	0.11%*	0.11%
Locked NBBO	11.52%*	10.86%	11.94%	10.69%	16.48%	14.96%	16.54%*	16.33%
Panel B: Quoted and Effective Spreads								
Dollar Quoted Spread	0.013*[t]	14.18	0.010	0.014	0.010	0.016	0.010	0.013
Percent Quoted Spread	0.0005%*	0.0006%	0.0008%	0.0009%	0.0006%	0.0009%	0.0006%	0.0006%
Dollar Effective Spread	0.015	0.031	0.011	0.082	0.011	0.213	0.011	0.105
Percent Effective Spread	0.0006%	0.0008%	0.0009%	0.0041%	0.0007%	0.0085%	0.0006%	0.0042%
Panel C: Realized Spread and Permanent Price Impact								
Percent Realized Spread: LR	0.0002%	0.0004%	0.0003%	0.0010%	0.0003%	0.0012%	0.0003%	0.0009%
Percent Realized Spread: EMO	0.0002%	0.0004%	0.0003%	0.0009%	0.0003%*	0.0006%	0.0003%	0.0006%
Percent Realized Spread: CLNV	0.0002%	0.0004%	0.0003%	0.0009%	0.0003%	0.0007%	0.0003%	0.0006%
Percent Price Impact: LR	0.0003%*	0.0011%	0.0006%	0.0054%	0.0004%	0.0120%	0.0003%	0.0052%
Percent Price Impact: EMO	0.0003%*	0.0002%	0.0005%*	0.0004%	0.0004%*	0.0005%	0.0003%*	0.0005%
Percent Price Impact: CLNV	0.0003%*	0.0008%	0.0006%*	0.0005%	0.0004%*	0.0005%	0.0003%*	0.0005%

Table 2: Overall Measures Difference between TRTH and DTAQ Evaluated on a Year-by-Year Basis (Continue)

* indicates non-statistically different from DTAQ at 5% level.

	2012		2013		2014		2015	
Panel A: Trade Location								
	TRTH in milliseconds	DTAQ in milliseconds	TRTH in milliseconds	DTAQ in milliseconds	TRTH in milliseconds	DTAQ in milliseconds	TRTH in milliseconds	DTAQ in milliseconds
At the NBBO	98.36%	77.31%	89.63%	74.42%	86.95%	74.09%	98.77%	75.19%
Inside the NBBO	0.18%	20.61%	9.12%	23.85%	11.89%	24.60%	0.32%	23.61%
Outside the NBBO	1.46%	2.08%	1.25%	1.74%	1.20%*	1.31%	0.91%	1.20%
Crossed NBBO	0.08%*	0.07%	0.07%*	0.05%	0.13%*	0.08%	0.06%*	0.04%
Locked NBBO	11.00%	9.13%	7.23%*	7.27%	9.19%	10.36%	7.00%*	7.38%
Panel B: Quoted and Effective Spreads								
Dollar Quoted Spread	0.010	0.015	0.011	0.013	0.011*	0.011	0.011*	0.011
Percent Quoted Spread	0.0005%	0.0007%	0.0004%	0.0005%	0.0004%*	0.0004%	0.0004%*	0.0004%
Dollar Effective Spread	0.011	0.118	0.011	0.038	0.010*	0.010	0.011	0.010
Percent Effective Spread	0.0006%	0.0045%	0.0004%	0.0015%	0.0004%*	0.0004%	0.0004%	0.0004%
Panel C: Realized Spread and Permanent Price Impact								
Percent Realized Spread: LR	0.0003%	0.0006%	0.0003%*	0.0002%	0.0003%*	0.0002%	0.0003%	0.0002%
Percent Realized Spread: EMO	0.0003%*	0.0004%	0.0003%*	0.0003%	0.0002%*	0.0002%	0.0003%	0.0002%
Percent Realized Spread: CLNV	0.0003%*	0.0004%	0.0003%*	0.0003%	0.0003%*	0.0002%	0.0003%	0.0002%
Percent Price Impact: LR	0.0002%	0.0053%	0.0002%	0.0014%	0.0002%*	0.0002%	0.0002%*	0.0002%
Percent Price Impact: EMO	0.0002%*	0.0005%	0.0002%*	0.0003%	0.0001%*	0.0002%	0.0002%*	0.0002%
Percent Price Impact: CLNV	0.0002%*	0.0005%	0.0002%*	0.0003%	0.0002%*	0.0002%	0.0002%*	0.0002%

Table 2: Overall Measures Difference between TRTH and DTAQ Evaluated on a Year-by-Year Basis (Continue)

* indicates non-statistically different from DTAQ at 5% level.

	2008				2009			
Panel D: Returns								
	Mean	Std. Dev	Mean	Std. Dev	Mean	Std. Dev	Mean	Std. Dev
Returns (compared to 5 minutes prior)	0.0000%*	0.00044	0.0000%	0.00044	-0.0000%*	0.00045	0.00000%	0.00082
Returns (compared to 10 minutes prior)	-0.0001%*	0.00086	-0.0001%	0.00086	-0.00001%*	0.00091	-0.00001%	0.00140

	2010				2011			
Panel D: Returns								
	Mean	Std. Dev	Mean	Std. Dev	Mean	Std. Dev	Mean	Std. Dev
Returns (compared to 5 minutes prior)	0.0000%*	0.00022	0.00001%	0.00022	0.0000%*	0.00025	0.00000%	0.00025
Returns (compared to 10 minutes prior)	0.00002%*	0.00045	0.00002%	0.00045	0.0000%*	0.00048	0.00000%	0.00048

	2012				2013			
Panel D: Returns								
	Mean	Std. Dev	Mean	Std. Dev	Mean	Std. Dev	Mean	Std. Dev
Returns (compared to 5 minutes prior)	0.00001%*	0.00015	0.00001%	0.00048	0.00001%*	0.00014	0.00001%	0.00033
Returns (compared to 10 minutes prior)	0.00001%*	0.00030	0.00001%	0.00055	0.00003%*	0.00027	0.00003%	0.00039

Table 2: Overall Measures Difference between TRTH and DTAQ Evaluated on a Year-by-Year Basis (Continue)

* indicates non-statistically different from DTAQ at 5% level.

	2014				2015			
Panel D: Returns								
	Mean	Std. Dev	Mean	Std. Dev	Mean	Std. Dev	Mean	Std. Dev
Returns (compared to 5 minutes prior)	0.0000%*	0.00012	0.00000%	0.00012	0.00001%*	0.00018	0.00001%	0.00018
Returns (compared to 10 minutes prior)	-0.00001%*	0.00024	-0.00001%	0.00024	0.00002%*	0.00036	0.00002%	0.00036

2014 appears to be the only year where nearly all the measures are not significantly different – except for the percentage of trades that occur *At the NBBO* and *Inside the* NBBO, all the other trade location measures are not significantly different. Another trend observed is as the time period goes further back to the beginning of the sample period, there are less number of measures that are not significantly different. Nevertheless, there does not appear to be any not significantly different from returns calculated using DTAQ, regardless if it is based on trading price 5 minutes prior or 10 minutes prior.

Another observation made is that most of the measures that are statistically different, excluding trade locations, when computed using TRTH are smaller than those computed using DTAQ. This is similar to the observation obtained when analyzing the results for the entire sample period as a whole. This should not be surprising since if the measures have been consistently smaller for each year, evaluation on the entire sample should yield similar result.

None of the results are meaningful for the purpose of this paper up to this point. The next question that must be addressed is then naturally, "What went wrong?" The following section will attempt to compare the two data sets at each of the cleansing steps and try identify the cause of the discrepancies.

VI. Data Analysis

For the purpose of this section of this paper, a larger sample size would not provide much added benefit.[19] Therefore, this paper has decided to pick the first 5 trading days from each year to observe the changes the data sets go through after applying the steps explained in the methodology section. The logic behind this exercise is if similar steps and procedures are applied analogously on two identical data sets, any error made on one data set will occur on the other data set as well. The incorrect results obtained from both would still be identical.

Evaluation begins with the original data sets as obtained from WRDS and SIRCA. Comparisons of number of trades recorded by both the data sets are fairly straightforward. The DTAQ Daily Trade file against all the observations within the TRTH Time of Sales with the *type* field labelled 'Trade.' Both the DTAQ Daily Quote file and Daily NBBO file are merged together and duplicates based on similar *quote sequence number* are removed. The results are summarized in table 3.

The paper acknowledges that the sample size of five days may be small, however for illustration purposes without making further inferences, it is sufficient. Number of trades recorded by both data sets appears to be the same during the first five trading days of each year. However, the number of quotes reported by DTAQ have been consistently higher than the amount recorded by TRTH.

[19] Both the TRTH and DTAQ data sets come from the same source, NYSE. As such, they should be identical at the time of download, except for certain discrepancies that might occur as a result of difference in user settings or options when retrieving these data from SIRCA and WRDS. If the data sets are identical, they should be identical at every time interval – daily, weekly, monthly or quarterly.

Table 3: Average Number of Trades and Quotes per Day Recorded on TRTH and DTAQ for the First Five Trading Days from Each Year.

The data sets used to generate table 3 are of those obtained directly from SIRCA and WRDS without applying any manipulations or filters. * indicates non-statistically different from DTAQ at 5% level.

	Average Number of Quotes per Day Recorded				Average Number of Trades per Day Recorded			
	TRTH		DTAQ		TRTH		DTAQ	
	Mean	**Std Dev**	**Mean**	**Std Dev**	**Mean**	**Std Dev**	**Mean**	**Std Dev**
2008	171,827.80	31,612.52	758,937.60	143,576.90	90,217.8*	14,976.31	90,221.20	14,978.86
2009	196,495.80	28,407.41	1,039,547.00	138,081.04	169,210.4*	23,368.61	169,299.20	23,386.77
2010	186,219.40	53,306.24	1,099,730.40	377,712.25	157,256.2*	74,604.94	157,332.60	74,596.88
2011	141,898.40	12,372.66	984,483.00	141,493.69	104,888.8*	14,321.30	104,955.20	14,298.64
2012	128,043.00	15,209.42	934,875.20	88,410.70	90,694.6*	8,701.12	90,753.60	8,692.92
2013	141,431.80	24,340.41	926,912.80	142,265.43	82,619.2*	12,548.66	82,678.00	12,549.79
2014	269,859.80	32,365.77	1,139,018.40	182,818.17	81,917.4*	10,613.96	81,921.20	10,614.08
2015	256,397.40	48,573.57	1,755,405.80	385,757.78	122,540.8*	28,567.58	122,546.60	28,569.52

Table 4: Average Number of Quotes per Day Recorded on TRTH and DTAQ after Accounting for Crossed Markets.

	Average Number of Quotes per Day Recorded			
	TRTH		DTAQ	
	Mean	**Std Dev**	**Mean**	**Std Dev**
2008	171,484.20	31,446.33	644,786.40	115,608.10
2009	196,247.40	28,337.94	887,545.20	109,908.87
2010	186,025.60	53,114.43	978,567.20	341,376.93
2011	141,766.40	12,322.65	877,427.60	128,270.35
2012	128,025.20	15,221.07	831,446.60	76,663.91
2013	141,423.20	24,326.74	810,262.40	120,936.00
2014	269,836.80	32,361.40	997,155.00	148,373.88
2015	256,386.60	48,564.43	1,576,623.40	346,847.94

The subsequent step removes quotes which create a *Crossed Markets* situation – where the bid price is greater than the ask price, regardless of the exchange. The steps described in the methodology section first evaluates such quotes within the same market before reconsider them again when evaluating them at the NBBO level among different markets. The step proposed here combined both steps and attempt to identify such situations at one go.

From table 4, it is observed that between 10 to 15% of the quotes recorded in DTAQ are identified as *Crossed Markets* whereas the highest among recorded in TRTH is 0.2%. Though this process lowers the difference of quotes recorded between the two files, the difference are still large.

The following results presented in table 5 are obtained after accounting for cases when the bid price or ask price is equal to zero, and after applying the conditions highlighted in steps *f* to *g* for cleaning the quote data set. Few quotes fall within the conditions set in steps *f* to *g*. The $5.00 spread might probably be too large and at any given time period for a liquid stock, both ask and bid price should not vary by too much. The additional $2.50 condition might not have been small enough to account for cancelled quotes.

Up to this point, the paper has yet to account for halted quotes or quotes with nonnormal market conditions. Applying the conditions as outlined by HJ would include the removal of quotes with quote condition labeled 'R', meaning 'regular, two-sided open quotes.' Excluding quotes labeled 'R' will reduce the number of quotes for the entire 40 days to a total of only 88 quotes. As a result, quotes labeled 'R' are left among the data set. Furthermore, a two-sided open quotes does seem normal enough as a market liquidity provider. Nevertheless, only a few quotes are identified with nonnormal market conditions

Table 5: Average Number of Quotes per Day Recorded by TRTH and DTAQ after Accounting for Withdrawn Quotes.

	Average Number of Quotes per Day Recorded			
	TRTH		DTAQ	
	Mean	**Std Dev**	**Mean**	**Std Dev**
2008	171,484.20	31,446.33	644,590.40	115,573.30
2009	196,247.40	28,337.94	887,087.60	109,897.24
2010	186,025.60	53,114.43	978,390.40	341,368.13
2011	141,766.40	12,322.65	876,785.20	128,259.27
2012	128,025.20	15,221.07	831,294.40	76,609.57
2013	141,423.20	24,326.74	810,169.20	120,920.17
2014	269,836.80	32,361.40	996,527.00	148,272.73
2015	256,386.60	48,564.43	1,576,046.40	346,788.96

(excluding quotes condition labeled 'R').

As suggested earlier, the $5 spread condition[20] might have been too wide for a liquid stock in a HFT environment. This paper evaluated this criteria further by narrowing the spread further from $5 to $4 and then again to $3 while its accompanying previous midpoint condition adjusted from plus/minus $2.50 to $2.00 and then again to $1.50 respectively. This adjusted criteria has little to no impact to the number of average quotes per day recorded by TRTH. However, though the number of average quotes per day recorded by DTAQ decreases with the narrowing of the spread, the number of average quotes per day remains high and significantly different.

Further analysis indicates that in cases when trades locations are statistically different, data set from TRTH consistently yields more *Crossed NBBO* trades than the data set from DTAQ. Additionally, DTAQ out performs TRTH by recording smaller percentage of

[20] Quotes with spread greater than $5.00 and ask(bid) price greater(less) than the previous midpoint plus(minus) $2.50.

instances when the percent effective spread is greater than the percent quoted spread. It needs to be noted as well that the number of times the *Percent Quoted Spread* goes below zero are significantly smaller for DTAQ than TRTH.

VII. Conclusion

Over the entire sample period, replicating the steps and procedures, as suggested by HJ in both their paper and SAS code, on the Time of Sales file from TRTH, and Daily Quote file, Daily Trade file, and Daily NBBO file from DTAQ, this paper fails to conclude that data sets obtained from the two separate databases yield the results. Though they build their database with information from the same source, the amount of information presented are significantly different at the time of data retrieval.

DTAQ presents significantly higher amount of quotes than TRTH, and the steps suggested by HJ are not able to reduce the amount of quotes to a comparable level. Nevertheless, the size of a data set does not necessarily have a significant impact on the quality of the analysis. However, based on the size of the data set, this paper concludes that the TRTH and DTAQ are not identical. As HJ demonstrated the effectiveness of their methodology, they were only able to *improve* the MTAQ liquidity measures – *improve* in the sense they were able to bring values of the measures closer to those obtained using DTAQ. The majority of the measures were significantly different.

The smaller amount of reported quotes in TRTH is not due to Thomson Reuters' ability to identify quotes place by a single agent and merge them as a single quote. The total bid size and ask size has been evaluated and were found to be significantly different. In addition, the support team at Thomson Reuters was contacted and they have clarified that Thomson Reuters receive all feeds from NYSE as it is and these feeds are updated with each update they receive from NYSE. It must be emphasized that no manipulation has been applied.

Finally, this paper finds sufficient evidence to support HJ claims that the DTAQ is in fact the *first-best* based on the following;

i. DTAQ consistently yields less *Crossed NBBO* trades

ii. DTAQ has less *Percent Quoted Spread* that are less than zero

iii. DTAQ recorded lower percentage of *Percent Effective Spread* that are greater than the *Percent Quoted Spread*

References

Biais, B. & Woolley, P. (March, 2011). High Frequency Trading. Retrieved from http://idei.fr/doc/conf/pwri/biais_pwri_0311.pdf

Cartea, A. & Penalva, J. (December, 2011). Where is the Value on High Frequency Trading? Retrieved from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1712765

Chakrabarty, B., Li, B., Nguyen, V. & Van Ness, R (March, 2007). Trade classification algorithms for electronic communications network trades. *Journal of Banking & Finance 31, 3806-3821.*

Easley[a], S., de Prado, M. M. L. & O'Hara M. (February, 2012). Flow Toxicity and Liquidity in a High Frequency World. *Review of Financial Studies, Forthcoming.*

Easley[b], S., de Prado, M. M. L. & O'Hara M. (May, 2012). The Volume Clock: Insights Into The High Frequency Paradigm. *Journal of Portfolio Management, Forthcoming.*

Ellis, K., Michaely, R. & O'Hara, M. (December, 2000). The Accuracy of Trade Classification Rules: Evidence from Nasdaq. *The Journal of Financial and Quantitative Analysis, Vol. 35, No. 4.*

Gai, J., Yao, C. & Ye, M. (March, 2012). The Externalities of High-Frequency Trading. Retrieved from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2066839

Goldstein, M. A., Kumar, P. & Graves, F. C. (May, 2014). Computerized and High-Frequency Trading. *The Financial Review, Vol. 49, No. 2.*

Hagströmer, B. & Nordén, L. (September, 2012). The diversity of high frequency traders. Retrieved from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2153272

Hirschey, N. H. (April, 2013). Do High-Frequency Trades Anticipate Buying and Selling Pressure? Retrieved from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2238516

Holden, C. & Jacobsen, S. (August, 2014). Liquidity Measurement Problems in Fast, Competitive Markets: Expensive and Cheap Solutions. *The Journal of Finance, Vol. 69, No. 4.*

Lee, C. & Ready, M. (June, 1991). Inferring Trade Direction from Intraday Data. *Journal of Finance, Vol. 46, No. 2.*

MacIntosh J. G. (October, 2013). High Frequency Traders: Angels or Devils? *C.D. Howe Institute, Commentary No. 391.* Retrieved from http://www.cdhowe.org/pdf/Commentary_391.pdf