Internet Appendix for

"Liquidity Measurement Problems in Fast, Competitive Markets:

Expensive and Cheap Solutions*"

Summary

This Internet Appendix discusses our SAS code to implement our MTAQ recommendations, provides more detail on the Errors-In-Variables model, and provides a variety of additional results. It is organized as follows. Section I provides a short description as well as the online location of of our SAS code to implement our MTAQ recommendations. Section II provides a more detailed description of the Errors-In-Variables model. Section III discusses the performance of the Duration Limited Control (DLC) technique relative to the techniques recommended in the paper. Section IV presents the empirical results when an alternative lag time is used to match trades to NBBO quotes. Section V examines how excluding regional exchanges compares to the MTAQ solutions recommended in the paper. In Section VI, we compare exchange Percent Effective Spread rankings under alternative MTAQ methods versus under Rule 605 data.

I.   SAS Program to Implement Our MTAQ Recommendations

A complete, commented SAS file to implement our MTAQ recommendations is available on our individual web sites (Holden: www.kelley.iu.edu/cholden and Jacobsen: www.cox.smu.edu/web/sjacobsen). We appreciate any feedback on this program and plan to update it over time. The July 2013 version of this program is provided at the end of this Internet Appendix. The SAS code does the following: downloads MTAQ data from WRDS, applies our recommended techniques (adjust for withdrawn quotes, use interpolated time, and eliminate NBBO quotes and trades when the market is crossed or locked), and computes standard liquidity measures (time-weighted *Percent Quoted Spread*, volume-weighted *Percent Effective Spread*, *Percent Realized Spread*, and *Percent Price Impact*).

II. More Details on the Simple Errors-In-Variables Model

This section provides enough additional information about the simple error-in-variables model to enable the interested reader to replicate our results. The specific parameter values used in the simulation were found by an iterative process of trial-and-error in order to match the simulated liquidity measures to the actual liquidity measures.

Trades are generated relative to the DTAQ NBBO with no errors. Specifically, let L_k be the trade location of the kth trade with values $\{$*Above Offer*, *At Offer*, *Inside*, *At Bid*, *Below Bid*$\}$ and probabilities $\{1.5\%, 35\%, 27\%, 35\%, 1.5\%\}$. Let P_k be the price of the kth trade as given by

$$P_k = \begin{cases} NBO_k^D + .01 & \text{if } L_k = \textit{Above Offer} \\ NBO_k^D & \text{if } L_k = \textit{At Offer} \\ \left(NBO_k^D + NBB_k^D\right)\big/2 & \text{if } L_k = \textit{Inside} \\ NBB_k^D & \text{if } L_k = \textit{At Bid} \\ NBB_k^D - .01 & \text{if } L_k = \textit{Below Bid}. \end{cases} \tag{IA.1}$$

Let L_{k-1} be the trade location of the $k\text{-}1^{st}$ trade with the same distribution at L_k. Let T_k be the tick test for the kth trade as follows:

$$T_k = \begin{cases} Up & \begin{aligned} &\text{if } L_k = Above\ Ask \\ &\quad \text{Or } \left(L_k = At\ Ask \text{ and } L_{k-1} \neq Above\ Ask \right) \\ &\quad \text{Or } \left(L_k = Inside,\ L_{k-1} = Inside, \text{ and } TieBreak = Buy \right) \\ &\quad \text{Or } \left(L_k = Inside,\ L_{k-1} = At\ Bid, \text{ and } TieBreak = Buy \right) \\ &\quad \text{Or } \left(L_k = At\ Bid,\ L_{k-1} = Below\ Bid, \text{ and } TieBreak = Buy \right) \end{aligned} \\ Down & \text{otherwise,} \end{cases}$$

(IA.2)

where $TieBreak = \{Buy,\ Sell\}$ with probabilities $\{0.5,\ 0.5\}$. This structure for the tick test ensures that the tick test and trade location are strongly correlated, which by reverse engineering is critical to maintaining similar results using the three trade typing conventions that put different degrees of emphasis on the tick test versus trade location in typing trades.

We need the DTAQ and MTAQ midpoint five minutes into the future in to compute the realized spread and price impact. Let M_{k+5} be the true midpoint five minutes after the kth trade given by $M_{k+5} = M_k + Z_k Q_k$, where $Z_k = \{-\$0.02, -\$0.01, \$0.00, \$0.01, \$0.02, \$0.03, \$0.04\}$ with probability $\{5\%,\ 10\%,\ 10\%,\ 15\%,\ 20\%,\ 20\%,\ 20\%\}$ and $Q_k = \{1 \text{ if } L_k = Above\ Ask$ or $L_k = At\ Ask$ or $(L_k = Inside$ and $T_k = Up)$, 0 otherwise$\}$.

Let NBO^i_{k+5} and NBB^i_{k+5} be the National Best Offer and National Best Bid, respectively, that are in-force five minutes after the kth trade as computed from i^{th} data set, where $i = M$ means MTAQ and $i = D$ means DTAQ. As before, the DTAQ NBBO is assumed to be measured with no error based on the midpoint plus or minus the half-spread,

$$NBO^D_{k+5} = M_{k+5} + .5 * S_{k+5}$$

(IA.3)

$$NBB^D_{k+5} = M_{k+5} - .5 * S_{k+5},$$

(IA.4)

3

where S_{k+5} is a random spread in-force five minutes after the kth trade; S_{k+5} is assumed to have

the same distribution as S_k, which is $\{-\$0.02, \$0.00, \$0.02, \$0.04, \$0.06, \$0.08, \$0.10\}$ with

probability $\{0.5\%, 1.7\%, 7.8\%, 10\%, 10\%, 30\%, 40\%\}$. We assume that the MTAQ NBBO is the

DTAQ NBBO plus errors as given by

$$NBO^{M}_{k+5} = NBO^{D}_{k+5} + \varepsilon^{O}_{k+5}, \tag{IA.5}$$

$$NBB^{M}_{k+5} = NBB^{D}_{k+5} + \varepsilon^{B}_{k+5}, \tag{IA.6}$$

where ε^{O}_{k+5} and ε^{B}_{k+5} are assumed to have the same distribution as ε^{O}_{k} and ε^{B}_{k}. Specifically, the

offer error ε^{O}_{k} has the following distribution:

- with probability of 71.215%: $0.00

- with probability of 1.515% for each value: -$0.80, -$0.70, -$0.60, -$0.50, -$0.40, -$0.30, -$0.20, -$0.10, -$0.05, -$0.03, $0.03, $0.05, $0.10, $0.20, $0.25, $0.30, $0.35, $0.40, and $0.50.

The bid error ε^{B}_{k} distribution is the mirror image of the offer error distribution (i.e., multiply each

value above by -1). The offer and bid error distributions have following three properties:

(1) By reverse engineering, the vast majority of the probability is on zero. It must be this way,

as otherwise Prob(*Outside the NBBO*) and Prob(*Crossed NBBO*) would become too high.

(2) $E\left[\varepsilon^{O}_{k}\right] = -\$0.023$ and $E\left[\varepsilon^{O}_{k}\right] = +\$0.023$. By reverse engineering, these nonzero means

are important because they are what causes E[*MTAQ Quoted Spread*] < E[*DTAQ Quoted

Spread*]. It is easy to prove the connection. Take expectations of both sides of the MTAQ

NBO and MTAQ NBB equations and then subtract one from the other to get

$$E\left[MTAQ\ Quoted\ Spread\right] = E\left[DTAQ\ Quoted\ Spread\right] + E\left[\varepsilon^{O}_{k}\right] - E\left[\varepsilon^{B}_{k}\right]. \tag{IA.7}$$

(3) The nonzero values of the offer and bid errors are relatively large in absolute value. By reverse engineering it must be this way, because what drives the E[*MTAQ Effective Spread*] to be larger than the E[*DTAQ Effective Spread*] is essentially the Variance (Offer Error + Bid Error). Given that the MTAQ *Percent Effective Spread* is 53.9% larger than its counterpart and the MTAQ *Dollar Effective Spread* is 107.3% larger than its counterpart, and given that only a small amount of the probability can be nonzero (see point #1), then the absolute size of the nonzero values must be large. The absolute size of these errors is consistent with what we know about the size and frequency of cancelled quotes, withdrawn quotes, and millisecond timestamps. The size of these errors is consistent with Figure 1 in the main article, where a cancelled error leads to more than a negative one dollar quoted spread. The 29.5% probability of a nonzero error fits with our one-day subsample, which shows that withdraw quotes alone enter the NBBO for a time-weighted median of 10.7% of the trading day.

Allocating a trade to a particular trade location can be tricky when markets are crossed or locked, because sometimes a given trade can have two different but equally legitimate interpretations. For example, suppose the market is crossed with the National Best Bid = $20.10 and the National Best Offer = $20.00, and a trade happens in the middle at $20.05. The trade is both above the offer and below the bid. In such situations, we allocate half-credit for the trade to both trade locations to avoid double-counting. The listing below shows how both the simulation and the empirical analysis allocate full credit for a trade to a single trade location or half-credit for a trade to two trade locations under normal, locked, and crossed markets. The listing applies to both MTAQ $(i = M)$ and DTAQ $(i = D)$.

A. *Trade Location Allocation under Normal, Locked, and Crossed Markets*

In a Normal Market $\left(NBO_k^i > NBB_k^i\right)$:

- If $P_k > NBO_k^i$, then full credit to Above Ask
- If $P_k = NBO_k^i$, then full credit to At Ask
- If $NBB_k^i < P_k < NBO_k^i$, then full credit to Inside
- If $P_k = NBB_k^i$, then full credit to At Bid
- If $P_k < NBB_k^i$, then full credit to Below Bid

In a Locked Market $\left(NBO_k^i = NBB_k^i\right)$:

- If $P_k > NBO_k^i$, then full credit to Above Ask
- If $P_k = NBO_k^i = NBB_k^i$, then half credit to At Ask and half credit to At Bid
- If $P_k < NBB_k^i$, then full credit to Below Bid

In a Crossed Market $\left(NBO_k^i < NBB_k^i\right)$:

- If $P_k > NBB_k^i > NBO_k^i$, then full credit to Above Ask
- If $NBB_k^i = P_k > NBO_k^i$, then half credit to At Bid and half credit to Above Ask
- If $NBB_k^i > P_k > NBO_k^i$, then half credit to Below Bid and half credit to Above Ask
- If $NBB_k^i > P_k = NBO_k^i$, then half credit to Below Bid and half credit to At Ask
- If $NBB_k^i > NBO_k^i > P_k$, then full credit to Below Bid

III.   A Test of Duration Limited Control

A recent paper by Jain, Upson, and Wood (2008) addresses the problem of cancelled quotes and proposes Duration Limited Control (DLC) to mitigate this issue. DLC drops all quotes older than a one-minute duration when computing the NBBO. The idea is that recent quotes are very likely to still be in-force. By contrast, older quotes run an increasing risk of cancellation and are more likely to be further away from current values. This technique admittedly throws away some older quotes that are still valid, but it limits the likelihood and potential size of cancelled quote contamination.

Table IA.I reports trade locations, cost of trading measures, and depths shown with and without Duration Limited Control (DLC) using six methods that match trades with MTAQ Quotes

and use the Prior Second quote timing technique. For columns (1) to (6), absolute differences are discussed relative to the reference benchmark in column (7), which uses the DTAQ Complete Official NBBO.

The first comparison is between column (2) with DLC and column (1) without DLC, where both methods have made no adjustments. Compared to column (1), column (2) has less absolute difference in trade location and cost of trading measures. Thus, DLC *helps* reduce absolute differences across-the-board when no adjustments have been made.

The next comparison is between column (4) with DLC and column (3) without DLC, when withdrawn quotes have been accounted for. Compared to column (3), column (4) has less absolute differences in trade location measures and many cost of trading measures (with the exception of *Percent Effective Spread* and *Percent Price Impact*). Thus, for many measures, DLC *helps* reduce absolute differences when withdrawn quotes have been accounted for.

The next comparison is between column (6) with DLC and column (5) without DLC, when withdrawn quotes have been accounted for and NBBO crossed and locked are excluded. Compared to column (5), column (6) has slightly less absolute difference in *Outside the NBBO*, by construction the same absolute difference in *Crossed NBBO*, much more absolute difference in *Percent (Dollar) Quoted Spread*, *Percent (Dollar) Effective Spread*, *Percent Realized Spread,* and *Percent Price Impact* using all three conventions (LR, EMO, and CLNV), less absolute difference in percent of time that *Percent Effective Spread* is greater than *Percent Quoted Spread*, and very little absolute difference in the dollar and share depth measures. Thus, DLC meaningfully *increases* absolute differences in the *Quoted Spread*, *Effective Spread*, *Realized Spread*, and *Price Impact* and has modest or no benefit on other liquidity measures when withdrawn quotes have been accounted for and NBBO crossed and locked are excluded.

The final comparison is across all MTAQ columns (1) to (6). In Panel A, column (6) has the lowest absolute difference in *Outside the NBBO*, yet column (5) is only slightly higher. In Panel B, column (5) has the lowest absolute difference in *Quoted Spread* and *Effective Spread*, and column (6) has the lowest absolute difference in percent of time that *Percent Effective Spread* is greater than *Percent Quoted Spread*. In Panel C, column (5) has the lowest absolute difference in *Percent Realized Spread* under EMO and the lowest absolute difference in *Percent Price Impact* under all three trade-typing conventions. In Panel D, all six columns have only minimal absolute differences in the dollar and share depth measures. Thus, the best method across the six MTAQ columns does *not* include DLC in the cases of *Quoted Spread*, *Effective Spread*, *Realized Spread*, and *Price Impact*.

To summarize Table IA.I, we find that DLC does *better* than no adjustments and adding DLC when accounting for withdrawn quotes does *better* than accounting for withdrawn quotes alone. However, adding DLC to accounting for withdrawn quotes and excluding NBBO crossed and locked does *worse* in most instances than accounting for withdrawn quotes and excluding NBBO crossed and locked without DLC. We conclude that accounting for withdrawn quotes and excluding NBBO crossed and locked should be used and DLC should not.

IV.   Alternative DTAQ Lags

Table IA.II examines how a longer DTAQ lag of 100 milliseconds compares to the minimum DTAQ lag of one millisecond in terms of its impact on liquidity measures. The selection of a 100 millisecond lag is a somewhat arbitrary choice of a point that is between one millisecond and one full second (1,000 milliseconds).

Table IA.II is identical to Table III in the paper, but with the 100 millisecond lag column added. In Panel A, *Outside the NBBO* is significantly higher with a 100 millisecond lag than a one millisecond lag (5.1% versus 3.3%), *Locked NBBO* is significantly lower (1.3% versus 1.7%), and *Crossed NBBO* is significantly lower (0.3% versus 0.5%). In Panel B, *Percent (Dollar) Quoted Spread* are the same for both lags, *Percent Effective Spread* is insignificantly higher with a 100 millisecond lag than one millisecond lag (0.420% versus 0.377%), *Dollar Effective Spread* is insignificantly higher (7.93¢ versus 6.46¢), and percent of the time that *Percent Effective Spread* is greater than *Percent Quoted Spread* is significantly higher (42.8% versus 34.5%). In Panel C, *Percent Realized Spread* (*Percent Price Impact*) is lower (higher) with a 100 millisecond lag than a one millisecond lag using all three conventions. In Panel D, the dollar and share depth measures are the same for both lags.

Overall, the differences between a 100 millisecond lag and one millisecond lag are relatively small. Each lag has some advantages. We recommend using the minimum value of a one millisecond lag due to the smaller *Outside the NBBO*. Further, we think that if the choice is close, then a "tie" should be broken in favor of the minimum lag value of one millisecond, because an arbitrarily longer lag may not be stable over time as trading speeds accelerate and system latencies decline over time.

V. Excluding Regional Exchanges

Table IA.III examines how excluding regional exchanges affect liquidity measures compared to other alternative MTAQ solutions. Deciding which exchanges to exclude requires the researcher to use judgment while applying a reasonable criterion, such as market share. In recent years, most regional exchanges have been acquired by national exchanges. We decide to exclude:

NASDAQ OMX BX (formerly Boston Stock Exchange), NASDAQ OMX PHLX (formerly Philadelphia Stock Exchange), National Stock Exchange (formerly Cincinnati Stock Exchange), Chicago Stock Exchange (formerly Midwest Stock Exchange), International Securities Exchange (primarily an options exchange), and Chicago Board Options Exchanges (primarily an options exchange). We decided to keep NYSE-ARCA, even though it integrated the former regional Pacific Stock Exchange, because it currently has a sizable market share of 10%. Thus, we keep the following: NYSE, NYSE-American, NYSE-ARCA, NASDAQ T & Q, BATS Y & Z, Direct Edge J & K, Alternative Display Facility D, Market Independent E, Consolidated Tape System S, and the market makers.

Table IA.III is identical to Table III in the paper, but with the excluding regional exchanges column added. We focus on comparison between Prior Second and excluding regional exchanges (the 2nd column) and Interpolated Time and excluding NBBO crossed and locked (the 5th column). In Panel A, *Outside the NBBO* is higher with Prior Second and excluding regional exchanges than Interpolated Time and excluding NBBO crossed and locked (11.8% versus 5.6%), *Locked NBBO* is closer to DTAQ (0.9% versus 0.0%), and *Crossed NBBO* is further away from DTAQ (3.1% versus 0.0%). In Panel B, *Percent Quoted Spread* is nearly the same, *Dollar Quoted Spread* is lower with Prior Second and excluding regional exchanges than Interpolated Time and excluding NBBO crossed and locked (7.10¢ versus 7.56¢), *Percent Effective Spread* is higher (0.506% versus 0.396%), *Dollar Effective Spread* is higher (9.61¢ versus 6.86¢), and percent of the time that *Percent Effective Spread* is greater than *Percent Quoted Spread* is higher (45.8% versus 37.0%). In Panel C, *Percent Realized Spread* (*Percent Price Impact*) is lower (higher) using all three conventions. In Panel D, the dollar and share depth measures are relatively close for both methods.

Overall, each method has comparative advantages. We recommend Interpolated Time and excluding NBBO crossed and locked given that this method outperforms on the most critical measures: *Percent Effective Spread*, *Dollar Effective Spread*, *Crossed NBBO*, and *Outside the NBBO*. A further drawback to excluding regional exchanges is the difficulty in choosing the appropriate exchanges to exclude in different time periods.

VI.   Comparing Exchange Rankings Under Alternative MTAQ Methods versus Rule 605 Data

This section compares exchange *Percent Effective Spread* rankings under alternative MTAQ methods versus under Rule 605 data. This analysis is based on Boehmer, Jennings, and Wei (2007, BJW), who analyze a sample of 258 NYSE-listed stocks from June 2001 to April 2004, a period of time during which the NYSE still had a dominant market share. DTAQ did not exist during most of this period, so we compare unadjusted MTAQ and second-best MTAQ to the Rule 605 data.

The Securities and Exchange Commission adopted Regulation NMS, which includes Rule 605 (formerly Regulation 11Ac1-5). Rule 605 requires that all market centers disclose monthly performance statistics by stock, order type, and order size. Under Rule 605:

- the percent effective spread is computed from order data (not trade and quote data),
- trades are matched to the NBBO at the time of order arrival (not at the time of order execution),
- buy/sell trades types are known (not estimated), and
- trades of more than 10,000 shares are excluded.

So, in principle, there should be differences between Rule 605 percent effective spread and MTAQ percent effective spread.

Table IA.IV reports the frequency distribution of the difference in exchange rankings under alternative MTAQ methods versus under Rule 605 data. Panel A compares MTAQ with no adjustments versus Rule 605. In the All market centers column, the rankings agree (the difference is zero) 30% of the time and the difference is 1 or less 71% of the time. Panel B compares MTAQ with the second-best solution versus Rule 605. In the All market centers column, the rankings agreement increases 33% of the time and the difference is one or less 74% of the time. Panel C reproduce the BJW results in which the rankings agree 34% of the time and the difference is one or less 76% of the time. So it appears that the second-best solution and BJW are approximately equally effective in increasing the degree of agreement between MTAQ and Rule 605. This result is consistent with the data treatment reported by BJW. BJW is one of only two studies that explicitly state that they delete all NBBO crossed and locked markets out of the 28 papers that use MTAQ data in our survey of JF-JFE-RFS publications from 2006 to 2011.

REFERENCES

Boehmer, Ekkehart, Robert Jennings, and Li Wei, 2007, Public disclosure and private decisions: Equity market execution quality and order routing, *Review of Financial Studies* 20, 315-358.

Jain, Pankaj, James Upson, and Robert Wood, 2008, Post-Reg NMS transaction costs: An efficient NBBO estimation method, Working paper, University of Memphis.

Table IA.I
The Impact of Adding Duration Limited Control To Alternative MTAQ Solutions

In this table, we present liquidity measures shown with and without Duration Limited Control (DLC). The six methods to calculate the NBBO use the MTAQ Quotes file and the Prior Second quote timing technique. Two methods are based on MTAQ with no adjustments. Two methods are based on MTAQ with withdrawn quotes accounted for. Two methods are based on MTAQ with withdrawn quotes accounted for and Excluding NBBO Crossed and Locked. Everything is compared to DTAQ NBBO with the NBBO file. The sample spans April to June 2008 inclusive and consists of 100 randomly selected stocks, resulting in 33,754,779 trades. Liquidity measures are calculated using firm-day averages. * indicates statistically different from the last column (DTAQ) at the 1% level. Standard errors are robust to clustering along both the firm and the day dimensions.

	MTAQ in seconds, with no adjustments		MTAQ in seconds, with withdrawn quotes accounted for		MTAQ in seconds, with withdrawn quotes accounted for and Excluding NBBO Crossed and Locked		DTAQ in milliseconds, with NBBO file
	No DLC (1)	Plus DLC (2)	No DLC (3)	Plus DLC (4)	No DLC (5)	Plus DLC (6)	(7)
Panel A: Trade Location							
At the NBBO	61.0%*	66.8%*	67.8%*	68.5%*	73.0%*	71.5%	71.0%
Inside the NBBO	12.7%*	17.1%*	14.6%*	17.9%*	16.5%*	19.2%*	25.7%
Outside the NBBO	26.3%*	16.1%*	17.6%*	13.6%*	10.5%*	9.3%*	3.3%
Locked NBBO	3.0%*	1.7%	1.7%	1.7%	0.0%*	0.0%*	1.7%
Crossed NBBO	18.3%*	7.4%*	9.9%*	5.1%*	0.0%*	0.0%*	0.5%
Panel B: Quoted and Effective Spreads							
Percent Quoted Spread	0.213%*	0.502%	0.287%*	0.512%*	0.401%	0.592%*	0.399%
Dollar Quoted Spread	2.03¢*	7.80¢	4.76¢*	8.17¢	7.56¢	9.89¢*	7.47¢
Percent Effective Spread	0.580%*	0.555%*	0.506%*	0.515%*	0.445%*	0.493%*	0.377%
Dollar Effective Spread	13.39¢*	11.42¢*	9.62¢*	9.15¢*	7.76¢*	8.26¢*	6.46¢
Time %: % Eff Spd > % Quo Spd	49.4%*	47.9%*	52.1%*	47.9%*	54.9%*	46.0%*	34.5%
Panel C: Realized Spread and Permanent Price Impact							
Percent Realized Spread: LR	0.159%	0.135%	0.142%*	0.123%*	0.138%*	0.122%*	0.154%
Percent Realized Spread: EMO	0.105%*	0.061%*	0.089%	0.066%*	0.077%	0.060%*	0.083%
Percent Realized Spread: CLNV	0.147%	0.116%*	0.132%*	0.116%*	0.124%*	0.115%*	0.140%
Percent Price Impact: LR	0.516%*	0.504%*	0.365%*	0.396%*	0.307%*	0.378%*	0.223%
Percent Price Impact: EMO	0.396%*	0.391%*	0.351%*	0.387%*	0.290%*	0.367%*	0.235%
Percent Price Impact: CLNV	0.507%*	0.449%*	0.357%*	0.386%*	0.300%*	0.365%*	0.224%
Panel D: Depth Measures							
Dollar Ask Depth (000's)	$13.5	$13.5*	$14.1	$13.8*	$14.1	$13.8*	$14.1
Dollar Bid Depth (000's)	$13.2*	$13.6*	$13.5*	$13.5*	$13.8*	$13.7*	$14.0
Share Ask Depth	531*	530.5*	561	542.5*	558	544.3*	561.0
Share Bid Depth	529*	542.4*	535*	532.8*	553*	545.1*	567.0
# of Trades (000's)	33.8	33.8	33.8	33.8	29.1	31.6	33.8
# of Trades Dropped (Mil.)	0.0	0.0	0.0	0.0	4.7	2.2	0.0

Table IA.II
Alternative DTAQ Lags Compared to Alternative MTAQ Solutions

In this table we present liquidity measures using alternative lags (100 ms versus 1 ms) of the DTAQ data set in milliseconds with the NBBO file are compared to alternative MTAQ solutions. Trade locations, cost of trading measures, and depths, are shown under alternative treatments. In the base MTAQ case, withdrawn quotes are already accounted for. Three quote timing rules are compared: Prior Second, Same Second, and Interpolated Time. The Exclusion of NBBO Crossed and Locked is also considered. The sample spans April to June 2008 inclusive and consists of 100 randomly selected stocks, resulting in 33,754,779 trades. Liquidity measures are calculated using firm-day averages. * indicates statistically different from the last column (DTAQ, one millisecond lag) at the 1% level. Standard errors are robust to clustering along both the firm and the day dimensions.

	MTAQ in seconds with withdrawn quotes accounted for					
	Prior Second	Same Second	Interpolated Time	Interpolated Time, Exclude NBBO Crossed and Locked	DTAQ in milliseconds, with NBBO file, 100 ms lag	DTAQ in milliseconds, with NBBO file, 1 ms lag
Panel A: Trade Location						
At the NBBO	67.8%*	51.4%*	59.5%*	63.9%*	79.6%*	71.0%
Inside the NBBO	14.6%*	33.3%*	26.7%*	30.5%*	15.3%*	25.7%
Outside the NBBO	17.6%*	15.4%*	13.7%*	5.6%*	5.1%*	3.3%
Locked NBBO	1.7%	2.1%*	2.5%*	0.0%*	1.3%*	1.7%
Crossed NBBO	9.9%*	9.9%*	10.2%*	0.0%*	0.3%*	0.5%
Panel B: Quoted and Effective Spreads						
Percent Quoted Spread	0.287%*	0.287%*	0.285%*	0.401%	0.399%	0.399%
Dollar Quoted Spread	4.76¢*	4.76¢*	4.56¢*	7.56¢*	7.47¢	7.47¢
Percent Effective Spread	0.506%*	0.450%*	0.467%*	0.396%*	0.420%	0.377%
Dollar Effective Spread	9.62¢*	8.66¢*	9.07¢*	6.86¢*	7.93¢	6.46¢
Time %: % Eff Spd > % Quo Spd	52.1%*	37.6%*	38.4%*	37.0%*	42.8%*	34.5%
Panel C: Realized Spread and Permanent Price Impact						
Percent Realized Spread: LR	0.142%*	0.238%*	0.197%*	0.196%*	0.125%*	0.154%
Percent Realized Spread: EMO	0.089%	0.135%*	0.113%*	0.101%*	0.069%*	0.083%
Percent Realized Spread: CLNV	0.132%*	0.200%*	0.172%*	0.166%*	0.106%*	0.140%
Percent Price Impact: LR	0.365%*	0.212%	0.271%*	0.205%*	0.356%	0.223%
Percent Price Impact: EMO	0.351%*	0.277%*	0.302%*	0.238%	0.251%*	0.235%
Percent Price Impact: CLNV	0.357%*	0.241%*	0.285%*	0.221%	0.236%*	0.224%
Panel D: Depth Measures						
Dollar Ask Depth (000's)	$14.1	$14.1	$14.0*	$14.0	$14.1	$14.1
Dollar Bid Depth (000's)	$13.5*	$13.5*	$13.4*	$13.8*	$14.0	$14.0
Share Ask Depth	561	561	560	558	561	561
Share Bid Depth	535*	535*	531*	551*	567	567
# of Trades (Millions)	33.8	33.8	33.8	29.1	33.8	33.8
# of Trades Dropped (Mil.)	0.0	0.0	0.0	4.7	0.0	0.0

Table IA.III
Excluding Regional Exchanges Compared to Alternative MTAQ Solutions

In this table we present analysis of how excluding regional exchanges affect liquidity measures compared to other alternative MTAQ solutions. Trade locations, cost of trading measures, and depths, are shown under alternative treatments. In the base case, withdrawn quotes are already accounted for. Three quote timing rules are compared: Prior Second, Same Second, and Interpolated Time. The Exclusion of Regional Exchanges and the Exclusion of NBBO Crossed and Locked are also considered. Everything is compared to the DTAQ dataset with millisecond timestamps using the prior millisecond quote timing. The sample spans April to June 2008 inclusive and consists of 100 randomly selected stocks, resulting in 33,754,779 trades. Liquidity measures are calculated using firm-day averages. * indicates statistically different from the last column (DTAQ) at the 1% level. Standard errors are robust to clustering along both the firm and the day dimensions.

	MTAQ in seconds with withdrawn quotes accounted for					DTAQ in milliseconds, with NBBO file
	Prior Second	Prior Second, Exclude Regional Exchanges	Same Second	Interpolated Time	Interpolated Time and Exclude NBBO Crossed and Locked	
Panel A: Trade Location						
At the NBBO	67.8%*	67.9%*	51.4%*	59.5%*	63.9%*	71.0%
Inside the NBBO	14.6%*	20.3%*	33.3%*	26.7%*	30.5%*	25.7%
Outside the NBBO	17.6%*	11.8%*	15.4%*	13.7%*	5.6%*	3.3%
Locked NBBO	1.7%	0.9%*	2.1%*	2.5%*	0.0%*	1.7%
Crossed NBBO	9.9%*	3.1%*	9.9%*	10.2%*	0.0%*	0.5%
Panel B: Quoted and Effective Spreads						
Percent Quoted Spread	0.287%*	0.398%	0.287%*	0.285%*	0.401%	0.399%
Dollar Quoted Spread	4.76¢*	7.10¢	4.76¢*	4.56¢*	7.56¢*	7.47¢
Percent Effective Spread	0.506%*	0.506%	0.450%*	0.467%*	0.396%*	0.377%
Dollar Effective Spread	9.62¢*	9.61¢	8.66¢*	9.07¢*	6.86¢*	6.46¢
Time %: % Eff Spd > % Quo Spd	52.1%*	45.8%*	37.6%*	38.4%*	37.0%*	34.5%
Panel C: Realized Spread and Permanent Price Impact						
Percent Realized Spread: LR	0.142%*	0.172%	0.238%*	0.197%*	0.196%*	0.154%
Percent Realized Spread: EMO	0.089%	0.073%*	0.135%*	0.113%*	0.101%*	0.083%
Percent Realized Spread: CLNV	0.132%*	0.118%*	0.200%*	0.172%*	0.166%*	0.140%
Percent Price Impact: LR	0.365%*	0.377%*	0.212%	0.271%*	0.205%*	0.223%
Percent Price Impact: EMO	0.351%*	0.333%*	0.277%*	0.302%*	0.238%	0.235%
Percent Price Impact: CLNV	0.357%*	0.310%*	0.241%*	0.285%*	0.221%	0.224%
Panel D: Depth Measures						
Dollar Ask Depth (000's)	$14.1	$13.6*	$14.1	$14.0*	$14.0	$14.1
Dollar Bid Depth (000's)	$13.5*	$13.6	$13.5*	$13.4*	$13.8*	$14.0
Share Ask Depth	561	541	561	560	558	561
Share Bid Depth	535*	552*	535*	531*	551*	567
# of Trades (Millions)	33.8	33.8	33.8	33.8	29.1	33.8
# of Trades Dropped (Mil.)	0.0	0.0	0.0	0.0	4.7	0.0

Table IA.IV

Comparing Exchange Rankings Under Alternative MTAQ Methods vs. Rule 605

In this table we present the frequency distribution of the differences in exchange Percent Effective Spread rankings under alternative MTAQ treatments versus under Rule 605 data. Panel A compares MTAQ with no adjustments versus Rule 605 data. Panel B compares the MTAQ second-best solution (accounting for withdrawn quotes, using Interpolated Time, and Excluding NBBO Crossed and Locked) versus Rule 605 data. Panel C reproduces BJW results. We use the same sample as BJW, which is 258 stocks over the period June 2001 to April 2004.

Rule 605 rank - MTAQ rank	NYSE (%)	Boston (%)	Chicago (%)	Nasdaq (%)	All market centers
Panel A: Frequency distribution of differences between effective spread rankings under MTAQ with no adjustments versus under Rule 605 data					
-3	0%	4%	4%	8%	4%
-2	2%	8%	9%	25%	11%
-1	8%	18%	20%	29%	19%
0	36%	30%	33%	22%	30%
1	34%	22%	20%	11%	22%
2	14%	13%	12%	4%	11%
3	6%	4%	2%	0%	3%
Off by 1 or less	78%	71%	72%	62%	71%
Panel B: Frequency distribution of differences between effective spread rankings under the MTAQ second-best solution versus under Rule 605 data					
-3	0%	3%	5%	8%	4%
-2	0%	6%	10%	22%	10%
-1	4%	18%	22%	31%	19%
0	38%	33%	37%	22%	33%
1	37%	25%	19%	11%	23%
2	14%	13%	7%	4%	10%
3	7%	2%	0%	0%	2%
Off by 1 or less	78%	76%	78%	65%	74%
Panel C: Frequency distribution of differences between effective spread rankings under MTAQ as treated by Boehmer, Jennings, and Wei versus under Rule 605 data					
-3	0%	2%	3%	10%	4%
-2	0%	5%	6%	24%	9%
-1	2%	18%	18%	34%	18%
0	38%	35%	42%	23%	34%
1	40%	27%	23%	7%	24%
2	16%	12%	8%	2%	9%
3	3%	1%	0%	0%	1%
Off by 1 or less	80%	80%	83%	64%	76%

```
/* HOLDEN AND JACOBSEN SAS CODE FOR NATIONAL BEST BID AND OFFER (NBBO)
   AND LIQUIDITY MEASURE COMPUTATION
   Updated: September 10, 2013

   Based on: Holden, C. and S. Jacobsen, 2013, Liquidity Measurement
   Problems in Fast, Competitive Markets: Expensive and Cheap Solutions,
   Forthcoming in the Journal of Finance

   We welcome any comments or suggestions. We can be reached at:
   Holden: cholden@indiana.edu
   Jacobsen: staceyj@cox.smu.edu

   We thank Charles Collver, Ruslan Goyenko, and Zhong Zhang who have
   provided helpful suggestions and corrections. We are solely responsible
   for any remaining errors. Of course, you use this code at your own risk.
   */

/* Global settings */
options errors=50;
libname project 'C:\project\';

/* STEP ONE: RETRIEVE MONTHLY TRADE AND QUOTE (MTAQ) FILES FROM WRDS
             AND DOWNLOAD TO PC */
/* Connect to WRDS */
%let wrds = wrds.wharton.upenn.edu 4016;
options comamid=TCP remote=WRDS;
signon username=_prompt_;

/* Submit SAS statements to WRDS */
rsubmit;
    libname taq '/wrds/taq/sasdata';
    option msglevel=i mprint source;

    /* Retrieve quote data */
    data monthlyquote (drop=QSEQ);

        /* Enter MTAQ dates in YYYYMMDD format for Consolidated Quote
            files ("cq")*/
        set taq.cq_20030401 taq.cq_20030402 taq.cq_20080401 taq.cq_20080402;
            by symbol date time;

        /* Enter company tickers */
        where symbol in ('A','ABII','IBM') and

        /* Quotes are retrieved prior to market open time to ensure NBBO
           Quotes are available for beginning of the day trades*/
        (("9:00:00"t) <= time <= ("16:00:00"t));
        format date date9.;
    run;

     /* Retrieve trade data */
    data monthlytrade (drop=G127 COND TSEQ);

        /* Enter the same MTAQ dates as above in YYYYMMDD format for
           Consolidated Trade files ("ct")*/
        set taq.ct_20030401 taq.ct_20030402 taq.ct_20080401 taq.ct_20080402;
            by symbol date time;
```

```
        /* Enter same company tickers as above */
        where symbol in ('A','ABII','IBM') and

        /* Retrieve trades during normal market hours */
        (("9:30:00"t) <= time <= ("16:00:00"t));
        type='T';
        format date date9.;
    run;

    /* Download to PC */
    proc download data=monthlyquote out=project.monthlyquote;
    run;
    proc download data=monthlytrade out=project.monthlytrade;
    run;
endrsubmit;


/* STEP TWO: CLEAN QUOTE DATA */
data project.monthlyquote;
    set project.monthlyquote;

    /* Quote Filter 1: Abnormal Modes. Quotes with abnormal modes (i.e.,
       abnormal quote conditions) are set to extreme values so that they
       will not enter the NBBO */
    if mode in (4,7,9,11,13,14,15,19,20,27,28) then do; OFR=9999999; BID=0;
end;

    /* Quote Filter 2: Crossed Quotes on the Same Exchange. Quotes from a
       given exchange with positive values in which the Bid is greater than
       the Ask (i.e., crossed quotes) are set to extreme values so that they
       will not enter the NBBO */
    If BID>OFR and BID>0 and OFR>0 then do; OFR=9999999; BID=0; end;

    /* Quote Filter 3: One-Sided Bid Quotes. One-sided bid quotes (i.e.,
       quotes in which the Bid is a positive value and the Ask is set to '0')
       are allowed to enter the NBB; the Ask is set to an extreme value so
       that it will not enter the NBO. One-sided ask quotes are also allowed
       to enter the NBO (i.e., quotes in which the Ask is a positive value and
       the Bid is set to '0'). In these cases, the bid is already the extreme
       value 0; as a result, no adjustment is necessary to ensure it does not
       enter the NBB. */
    If BID>0 and OFR=0 then OFR=9999999;

    /* Quote Filter 4: Abnormally Large Spreads. Quotes with positive values
       and large spreads (i.e., spreads greater than $5.00) are set to extreme
       values so that they will not enter the NBBO */
    spr=OFR-BID;
    If spr>5 and BID>0 and OFR>0 and OFR ne 9999999 then do;
        BID=0; OFR=9999999; end;

    /* Quote Filter 5: Withdrawn Quotes. This is when an exchange temporarily
       has no quote, as indicated by quotes with price or depth fields
       containing values less than or equal to 0 or equal to '.'. See
       discussion in Holden and Jacobsen (2013), page 11. They are set to
       extreme values so that they will not enter the NBBO. They are NOT
```

```
          deleted, because that would incorrectly allow the prior quote from that
          exchange to enter the NBBO. NOTE: Quote Filter 5 must come last */
     if OFR le 0 then OFR=9999999;
     if OFR =. then OFR=9999999;
     if OFRSIZ le 0 then OFR=9999999;
     if OFRSIZ =. then OFR=9999999;
     if BID le 0 then BID=0;
     if BID =. then BID=0;
     if BIDSIZ le 0 then BID=0;
     if BIDSIZ =. then BID=0;
 run;


/* STEP THREE: CLEAN TRADE DATA */
data project.monthlytrade;
     set project.monthlytrade;

     /* Trade Filter: Keep only trades in which the Correction field
        contains '00' and the Price field contains a value greater than
        zero */
     where corr eq 0 and price gt 0;
 run;


/* STEP FOUR: CREATE INTERPOLATED TIME VARIABLES
   Based on: Holden and Jacobsen (2013), pages 23-25 */

/* Create Interpolated Quote Time for quote dataset */
/* 'J' indexes the order of quotes within a given second */
data project.monthlyquote (drop=MODE spr);
     set project.monthlyquote;
     retain J;
     by symbol date time;
     if first.symbol or first.date or first.time then J=1; else J=J+1;
 run;
/* 'N' is the total number of quotes within a given second */
proc sql;
       create table project.monthlyquote2
       as select a.*,  max(J) as N
       from project.monthlyquote as a
       group by symbol,date,time
       order by symbol,date,time,J;
quit;
data project.monthlyquote2 (drop=J N);
     set project.monthlyquote2;
     InterpolatedTime=time+((2*J-1)/(2*N));
     format InterpolatedTime best15.;
 run;
data project.monthlyquote;
     set _null_;
 run;

/* Create Interpolated Trade Time for trade dataset */
/* 'I' indexes the order of trades within a given second */
data project.monthlytrade;
     set project.monthlytrade;
     retain I;by symbol date time;
```

```
        if first.symbol or first.date or first.time then I=1; else I=I+1;
run;
/* 'N' is the total number of quotes within a given second */
proc sql;
        create table project.monthlytrade2
        as select a.*,  max(I) as N
        from project.monthlytrade as a
        group by symbol,date,time
        order by symbol,date,time,I;
quit;
data project.monthlytrade2 (drop=I N);
    set project.monthlytrade2;
    InterpolatedTime=time+((2*I-1)/(2*N));
    tradetime=time;
    format InterpolatedTime best15.;
    format tradetime time.;
run;
data project.monthlytrade;
    set _null_;
run;


/* STEP FIVE: NATIONAL BEST BID AND OFFER (NBB0) CALCULATION */
/* Assign ID to Each Unique Exchange or Market Maker and Find
   The Maximum Number of Exchanges*/
proc sort data=project.monthlyquote2;
    by ex mmid;
run;
data project.monthlyquote2;
    set project.monthlyquote2;
    retain ExchangeID;
    if _N_=1 then ExchangeID=0;
    if first.ex or first.mmid then ExchangeID=ExchangeID+1;
    by ex mmid;
run;
data _null_;
        set project.monthlyquote2 end=eof;
        retain MaxExchangeID;
        if ExchangeID gt MaxExchangeID then MaxExchangeID=ExchangeID;
        if eof then call symput('MaxExchangeID',MaxExchangeID);
run;
%put &MaxExchangeID;
proc sort data=project.monthlyquote2;
    by symbol date time InterpolatedTime;
run;

%macro BBO;
/* Create Dataset that has a Column for Each Exchange ID's Bid and Offer
   Quote for All Interpolated Times*/
data project.monthlyquote2;set project.monthlyquote2;
        by symbol date;
        array exbid(&MaxExchangeID);exbid(ExchangeID)=bid;
        array exofr(&MaxExchangeID);exofr(ExchangeID)=ofr;
        array exbidsz(&MaxExchangeID);exbidsz(ExchangeID)=bidsiz;
        array exofrsz(&MaxExchangeID);exofrsz(ExchangeID)=ofrsiz;
/* For Interpolated Times with No Quote Update, Retain Previous Quote
   Outstanding*/
```

```
%do i=1 %to &MaxExchangeID;
      retain exbidR&i exofrR&i exbidszR&i exofrszR&i;
      if first.symbol or first.date then exbidR&i=exbid&i;
        if exbid&i ge 0 then exbidR&i=exbid&i; else exbidR&i=exbidR&i+0;
      if first.symbol or first.date then exofrR&i=exofr&i;
      if exofr&i ge 0 then exofrR&i=exofr&i; else exofrR&i=exofrR&i+0;
      if first.symbol or first.date then exbidszR&i=exbidsz&i;
      if exbidsz&i ge 0 then exbidszR&i=exbidsz&i; else
      exbidszR&i=exbidszR&i+0;
      if first.symbol or first.date then exofrszR&i=exofrsz&i;
      if exofrsz&i ge 0 then exofrszR&i=exofrsz&i; else
      exofrszR&i=exofrszR&i+0;
%end;
/* Find Best Bid and Offer Across All Exchanges and Market Makers*/
%do i=&MaxExchangeID %to &MaxExchangeID;
      BestBid = max(of exbidR1-exbidR&i);
      BestOfr = min(of exofrR1-exofrR&i);
%end;
/* Find Best and Total Depth Across All Exchanges and Market Makers that
   are at the NBBO*/
%do i=1 %to &MaxExchangeID;
      if exbidR&i=BestBid then MaxBidDepth=max(MaxBidDepth,exbidszR&i);
      if exofrR&i=BestOfr then MaxOfrDepth=max(MaxOfrDepth,exofrszR&i);
      if exbidR&i=BestBid then TotalBidDepth=sum(TotalBidDepth,exbidszR&i);
      if exofrR&i=BestOfr then TotalOfrDepth=sum(TotalOfrDepth,exofrszR&i);
%end;
run;
%mend BBO;
%BBO;



data FinalNBBO (keep=symbol date time InterpolatedTime
    BestBid BestOfr MaxBidDepth MaxOfrDepth TotalBidDepth TotalOfrDepth);
    set project.monthlyquote2;
/* Only Output Changes in NBBO Records (e.g., changes in quotes or depth)*/
    if symbol eq lag(symbol)
        and date eq lag(date)
        and BestOfr eq lag(BestOfr)
        and BestBid eq lag(BestBid)
        and MaxOfrDepth eq lag(MaxOfrDepth)
        and MaxBidDepth eq lag(MaxBidDepth)
        and TotalOfrDepth eq lag(TotalOfrDepth)
        and TotalBidDepth eq lag(TotalBidDepth) then delete;
/* If Abnormal Quotes Enter the NBBO Then Set To ".". There Will Be
   NO NBBO */
    if BestBid < .00001 then
        do;
        BestBid=.;
        BestOfr=.;
        MaxOfrDepth=.;
        MaxBidDepth=.;
        TotalOfrDepth=.;
        TotalBidDepth=.;
        end;
    else if BestOfr > 9999998 then
        do;
        BestBid=.;
```

```
        BestOfr=.;
        MaxOfrDepth=.;
        MaxBidDepth=.;
        TotalOfrDepth=.;
        TotalBidDepth=.;
        end;
run;


/* STEP 6: INTERWEAVE TRADES WITH QUOTES: TRADES AT INTERPOLATED TIME TMMM
   ARE MATCHED WITH QUOTES IN FORCE AT INTERPOLATED TIME TMM(M-1)
   To Do This, Increase Interpolated Quote Time in Quotes Dataset by One
   Millisecond = .001*/
data FinalNBBO;
    set FinalNBBO;
    type='Q';
    InterpolatedTime+.001;
run;

/* Stack Quotes and Trades Datasets */
data Tradesquotes;
    set project.monthlytrade2 FinalNBBO;
    by symbol date InterpolatedTime type;
run;

/* For Each Trade, Identify the Outstanding NBBO, Best Depth and Total
   Depth */
data Tradesquotesv2 (drop=time BestOfr BestBid MaxOfrDepth MaxBidDepth
    TotalOfrDepth TotalBidDepth corr);
    set Tradesquotes;
    by symbol date;
    retain quotetime BestOfr2 BestBid2 MaxOfrDepth2 MaxBidDepth2
        TotalOfrDepth2 TotalBidDepth2;
    if type='Q' then quotetime=time;else quotetime=quotetime;
    if type='Q' then BestOfr2=BestOfr;else BestOfr2=BestOfr2;
    if type='Q' then BestBid2=BestBid;else BestBid2=BestBid2;
    if type='Q' then MaxOfrDepth2=MaxOfrDepth;else MaxOfrDepth2=MaxOfrDepth2;
    if type='Q' then MaxBidDepth2=MaxBidDepth;else MaxBidDepth2=MaxBidDepth2;
    if type='Q' then TotalOfrDepth2=TotalOfrDepth;
        else TotalOfrDepth2=TotalOfrDepth2;
    if type='Q' then TotalBidDepth2=TotalBidDepth;
        else TotalBidDepth2=TotalBidDepth2;
    format quotetime time.;
run;
data tradesquotes;
    set _null_;
run;


/* STEP 7: Classify Trades as "Buys" or "Sells" Using Three Measures: Lee &
   Ready (LR), Ellis, Michaely and O'Hara (EOH), and Chakrabarty, Li, Nguyen,
   and Van Ness (CLNV); Determine NBBO Midpoint and Locked and Crossed NBBOs
*/
data TSpread1;
    set Tradesquotesv2;
    where type='T';
    midpoint=(BestOfr2+BestBid2)/2;
```

```
    if BestOfr2=BestBid2 then lock=1;else lock=0;
    if BestOfr2<BestBid2 then cross=1;else cross=0;
run;


/* Determine Whether Trade Price is Higher or Lower than Previous Trade
   Price, or "Trade Direction" */
data TSpread1;
    set TSpread1;
    by symbol date;
    direction=dif(price);
    if first.symbol then direction=.;
    if first.date then direction=.;
run;
data TSpread1;
    set TSpread1;
    retain direction2;
    if direction ne 0 then direction2=direction;
    else direction2=direction2;
run;


/* First Classification Step: Classify Trades Using Tick Test */
data TSpread1 (drop=direction);
    set TSpread1;
    if direction2>0 then tradeLR='buy';
    if direction2<0 then tradeLR='sell';
    if direction2=. then tradeLR=.;
    if direction2>0 then tradeEOH='buy';
    if direction2<0 then tradeEOH='sell';
    if direction2=. then tradeEOH=.;
    if direction2>0 then tradeCLNV='buy';
    if direction2<0 then tradeCLNV='sell';
    if direction2=. then tradeCLNV=.;
run;


/* Second Classification Step: Update Trade Classification When
   Conditions are Met as Specified by LR, EOH, and CLNV */
data TSpread1;
    set TSpread1;
    if lock=0 and cross=0 and price gt midpoint then tradeLR='buy';
    if lock=0 and cross=0 and price lt midpoint then tradeLR='sell';
    if lock=0 and cross=0 and price=BestOfr2 then tradeEOH='buy';
    if lock=0 and cross=0 and price=BestBid2 then tradeEOH='sell';
    ofr30=BestOfr2-.3*(BestOfr2-BestBid2);
    bid30=BestBid2+.3*(BestOfr2-BestBid2);
    if lock=0 and cross=0 and price le BestOfr2 and price ge ofr30
        then tradeCLNV='buy';
    if lock=0 and cross=0 and price le bid30 and price ge BestBid2
        then tradeCLNV='sell';
run;



/* STEP 8: CALCULATE LIQUIDITY MEASURES - QUOTE STATISTICS */
/* Use Quotes During Normal Market Hours */;
data QSpread1;
    set FinalNBBO;
    if time lt ("9:30:00"t) then delete;
run;
```

```sas
/* Determine Time Each Quote is In Force Based on Interpolated Time */
proc sort data=QSpread1;
    by symbol date descending InterpolatedTime;
run;
data QSpread1;
    set QSpread1;
    by symbol date;
    inforce=abs(dif(InterpolatedTime));
run;

/* If Last Quote of Day, then Quote is Inforce Until 4:00 pm */
data QSpread1;
    set QSpread1;
    by symbol date;
    end=57600;   /* 4:00 pm = 57,600 seconds after midnight  */
    if first.symbol or first.date then inforce=max((end-InterpolatedTime),0);
run;
proc sort data=QSpread1 (drop=end);
    by symbol date InterpolatedTime;
run;

/* Find Midpoint */
data QSpread1;
    set QSpread1;
    midpoint=(BestOfr+BestBid)/2;
run;


data QSpread2;
    set QSpread1;
/* Delete Locked and Crossed Quotes */
    if BestOfr=BestBid or BestOfr<BestBid then delete;
/* Multiply Dollar Quoted Spread, Percent Quoted Spread, Maximum Dollar
   Depth, Maximum Share Depth, Total Dollar Depth, and Total Share
   Depth by Time Inforce */
    wQuotedSpread_Dollar=(BestOfr-BestBid)*inforce;
    wQuotedSpread_Percent=((BestOfr-BestBid)/midpoint)*inforce;
    wTotalOfrDepth_Dollar=BestOfr*TotalOfrDepth*inforce;
    wTotalBidDepth_Dollar=BestBid*TotalBidDepth*inforce;
    wTotalOfrDepth_Share=TotalOfrDepth*inforce;
    wTotalBidDepth_Share=TotalBidDepth*inforce;
    wMaxOfrDepth_Dollar=BestOfr*MaxOfrDepth*inforce;
    wMaxBidDepth_Dollar=BestBid*MaxBidDepth*inforce;
    wMaxOfrDepth_Share=MaxOfrDepth*inforce;
    wMaxBidDepth_Share=MaxBidDepth*inforce;
run;

/* Find Average Across Firm-Day */
proc sql;
    create table QSpread3
    as select symbol,date,
    sum(inforce) as sumtime,
    sum(wQuotedSpread_Dollar) as swQuotedSpread_Dollar,
    sum(wQuotedSpread_Percent) as swQuotedSpread_Percent,
    sum(wTotalOfrDepth_Dollar) as swTotalOfrDepth_Dollar,
    sum(wTotalBidDepth_Dollar) as swTotalBidDepth_Dollar,
```

```sas
        sum(wTotalOfrDepth_Share) as swTotalOfrDepth_Share,
        sum(wTotalBidDepth_Share) as swTotalBidDepth_Share,
        sum(wMaxOfrDepth_Dollar) as swMaxOfrDepth_Dollar,
        sum(wMaxBidDepth_Dollar) as swMaxBidDepth_Dollar,
        sum(wMaxOfrDepth_Share) as swMaxOfrDepth_Share,
        sum(wMaxBidDepth_Share) as swMaxBidDepth_Share
        from QSpread2
        group by symbol,date
        order by symbol,date;
    quit;

/* Calculate Time-Weighted Dollar Quoted Spread, Percent Quoted Spread,
   Maximum Dollar Depth, Maximum Share Depth, Total Dollar Depth, and
   Total Share Depth */
data QSpread3;
        set QSpread3;
        QuotedSpread_Dollar=swQuotedSpread_Dollar/sumtime;
        QuotedSpread_Percent=swQuotedSpread_Percent/sumtime;
        TotalOfrDepth_Dollar=swTotalOfrDepth_Dollar/sumtime;
        TotalBidDepth_Dollar=swTotalBidDepth_Dollar/sumtime;
        TotalOfrDepth_Share=swTotalOfrDepth_Share/sumtime;
        TotalBidDepth_Share=swTotalBidDepth_Share/sumtime;
        MaxOfrDepth_Dollar=swMaxOfrDepth_Dollar/sumtime;
        MaxBidDepth_Dollar=swMaxBidDepth_Dollar/sumtime;
        MaxOfrDepth_Share=swMaxOfrDepth_Share/sumtime;
        MaxBidDepth_Share=swMaxBidDepth_Share/sumtime;
    run;


/* STEP 9: CALCULATE LIQUIDITY MEASURES - TRADE STATISTICS:
   Effective Spreads; Multiply Dollar Effective Spreads and
   Percent Effective Spreads by Dollar and Share Size of Trade */
data TSpread1;
        set TSpread1;
        wEffectiveSpread_Dollar=(abs(price-midpoint))*2;
        wEffectiveSpread_Percent=wEffectiveSpread_Dollar/midpoint;
        dollar=price*size;
        wEffectiveSpread_Dollar_DW=wEffectiveSpread_Dollar*dollar;
        wEffectiveSpread_Dollar_SW=wEffectiveSpread_Dollar*size;
        wEffectiveSpread_Percent_DW=wEffectiveSpread_Percent*dollar;
        wEffectiveSpread_Percent_SW=wEffectiveSpread_Percent*size;
    run;

/* Delete Trades Associated with Locked or Crossed Best Bids or Best
   Offers */
data TSpread2;
        set TSpread1;
        if lock=1 or cross=1 then delete;
    run;

/* Find average across firm-day */
proc sql;
        create table TSpread3
        as select symbol,date,
        sum(dollar) as sumdollar,
        sum(size) as sumsize,
        mean(wEffectiveSpread_Dollar) as EffectiveSpread_Dollar,
```

```
        mean(wEffectiveSpread_Percent) as EffectiveSpread_Percent,
        sum(wEffectiveSpread_Dollar_DW) as waEffectiveSpread_Dollar_DW,
        sum(wEffectiveSpread_Dollar_SW) as waEffectiveSpread_Dollar_SW,
        sum(wEffectiveSpread_Percent_DW) as waEffectiveSpread_Percent_DW,
        sum(wEffectiveSpread_Percent_SW) as waEffectiveSpread_Percent_SW
        from TSpread2
        group by symbol,date
        order by symbol,date;
quit;

/* Calculate Dollar-Weighted (DW) and Share-Weighted (SW) Dollar Effective
   Spreads and Percent Effective Spreads */
data TSpread3;
    set TSpread3;
    EffectiveSpread_Dollar_DW=waEffectiveSpread_Dollar_DW/sumdollar;
    EffectiveSpread_Dollar_SW=waEffectiveSpread_Dollar_SW/sumsize;
    EffectiveSpread_Percent_DW=waEffectiveSpread_Percent_DW/sumdollar;
    EffectiveSpread_Percent_SW=waEffectiveSpread_Percent_SW/sumsize;
run;


/* STEP 10: CALCULATE LIQUIDITY MEASURES - TRADE STATISTICS:
   Realized Spreads and Price Impact; Find the NBBO Midpoint 5 Minutes
   Subsequent to the Trade */

/* Redefine the InterpolatedTime variable as 5 minutes earlier (e.g., quotes
   at 10:00:00 are redefined as occurring at 9:35:00 in order to match to
   trades occurring at 9:35:00. This way we match trades occurring at time T
   to NBBO quotes outstanding at T+5). */
data MidQ (keep=symbol date type midpointnew InterpolatedTime BestOfrnew
BestBidnew);
    set QSpread1;
    midpointnew=midpoint;
    InterpolatedTime=InterpolatedTime-300;
    BestOfrnew=BestOfr;
    BestBidnew=BestBid;
run;
data MidT (keep=symbol date tradetime InterpolatedTime type midpoint price
tradeLR
        tradeEOH tradeCLNV wEffectiveSpread_Dollar size dollar);
    set TSpread1;
run;
proc sort data=MidQ;
    by symbol date InterpolatedTime type;
run;
proc sort data=MidT;
    by symbol date InterpolatedTime type;
run;

/* Stack Trades at Time T with NBBO Quotes at Time T+5 */
data Mid1;
    set MidT MidQ;
    by symbol date InterpolatedTime type;
run;

/* For Each Trade at Time T, Identify the Outstanding NBBO at Time T+5 */
data Mid1;
```

```sas
    set Mid1;
    by symbol date;
    retain midpoint5 BestOfr5 BestBid5;
    if type='Q' then midpoint5=midpointnew;
    else midpoint5=midpoint5;
    if type='Q' then BestOfr5=BestOfrnew;else BestOfr5=BestOfr5;
    if type='Q' then BestBid5=BestBidnew;else BestBid5=BestBid5;
run;

/* Delete Trades at T Associated with Locked or Crossed Best Bids or Best
   Offers at T+5 */
data Mid2 (drop=midpointnew BestOfrnew BestBidnew);
    set Mid1;
    if BestOfr5=BestBid5 or BestOfr5<BestBid5 then delete;
run;

/* Create Indicator Variable "D" equal to "1" if Trade is a Buy and "-1" if
   Trade is a Sell for LR, EOH, and CLNV */
data Mid2;
    set Mid2;
    where type='T';
    if tradeLR='buy' then Dlr=1;
    if tradeLR='sel' then Dlr=-1;
    if tradeEOH='buy' then Deoh=1;
    if tradeEOH='sel' then Deoh=-1;
    if tradeCLNV='buy' then Dclnv=1;
    if tradeCLNV='sel' then Dclnv=-1;

    /* Compute Dollar and Percent Realized Spread for LR, EOH, and CLNV */
    wDollarRealizedSpread_LR=Dlr*(price-midpoint5)*2;
    wDollarRealizedSpread_EOH=Deoh*(price-midpoint5)*2;
    wDollarRealizedSpread_CLNV=Dclnv*(price-midpoint5)*2;
    wPercentRealizedSpread_LR=wDollarRealizedSpread_LR/midpoint5;
    wPercentRealizedSpread_EOH=wDollarRealizedSpread_EOH/midpoint5;
    wPercentRealizedSpread_CLNV=wDollarRealizedSpread_CLNV/midpoint5;

    /* Compute Dollar and Percent Price Impact for LR, EOH, and CLNV */
    wDollarPriceImpact_LR=wEffectiveSpread_Dollar-wDollarRealizedSpread_LR;
    wDollarPriceImpact_EOH=wEffectiveSpread_Dollar-wDollarRealizedSpread_EOH;
    wDollarPriceImpact_CLNV=wEffectiveSpread_Dollar-
    wDollarRealizedSpread_CLNV;
    wPercentPriceImpact_LR=wDollarPriceImpact_LR/midpoint5;
    wPercentPriceImpact_EOH=wDollarPriceImpact_EOH/midpoint5;
    wPercentPriceImpact_CLNV=wDollarPriceImpact_CLNV/midpoint5;

    /* Multiply Realized Spreads and Price Impact by Dollar and Share Size of
          Trade for LR, EOH, and CLNV */
    wDollarRealizedSpread_LR_SW=wDollarRealizedSpread_LR*size;
    wDollarRealizedSpread_LR_DW=wDollarRealizedSpread_LR*dollar;
    wPercentRealizedSpread_LR_SW=wPercentRealizedSpread_LR*size;
    wPercentRealizedSpread_LR_DW=wPercentRealizedSpread_LR*dollar;
    wDollarPriceImpact_LR_SW=wDollarPriceImpact_LR*size;
    wDollarPriceImpact_LR_DW=wDollarPriceImpact_LR*dollar;
    wPercentPriceImpact_LR_SW=wPercentPriceImpact_LR*size;
    wPercentPriceImpact_LR_DW=wPercentPriceImpact_LR*dollar;
    wDollarRealizedSpread_EOH_SW=wDollarRealizedSpread_EOH*size;
    wDollarRealizedSpread_EOH_DW=wDollarRealizedSpread_EOH*dollar;
```

```
            wPercentRealizedSpread_EOH_SW=wPercentRealizedSpread_EOH*size;
            wPercentRealizedSpread_EOH_DW=wPercentRealizedSpread_EOH*dollar;
            wDollarPriceImpact_EOH_SW=wDollarPriceImpact_EOH*size;
            wDollarPriceImpact_EOH_DW=wDollarPriceImpact_EOH*dollar;
            wPercentPriceImpact_EOH_SW=wPercentPriceImpact_EOH*size;
            wPercentPriceImpact_EOH_DW=wPercentPriceImpact_EOH*dollar;
            wDollarRealizedSpread_CLNV_SW=wDollarRealizedSpread_CLNV*size;
            wDollarRealizedSpread_CLNV_DW=wDollarRealizedSpread_CLNV*dollar;
            wPercentRealizedSpread_CLNV_SW=wPercentRealizedSpread_CLNV*size;
            wPercentRealizedSpread_CLNV_DW=wPercentRealizedSpread_CLNV*dollar;
            wDollarPriceImpact_CLNV_SW=wDollarPriceImpact_CLNV*size;
            wDollarPriceImpact_CLNV_DW=wDollarPriceImpact_CLNV*dollar;
            wPercentPriceImpact_CLNV_SW=wPercentPriceImpact_CLNV*size;
            wPercentPriceImpact_CLNV_DW=wPercentPriceImpact_CLNV*dollar;
    run;

    /* Find average across firm-day */
    proc sql;
        create table Mid3
        as select symbol,date,
        sum(dollar) as sumdollar,
        sum(size) as sumsize,
        mean(wDollarRealizedSpread_LR) as DollarRealizedSpread_LR,
        mean(wDollarRealizedSpread_EOH) as DollarRealizedSpread_EOH,
        mean(wDollarRealizedSpread_CLNV) as DollarRealizedSpread_CLNV,
        mean(wPercentRealizedSpread_LR) as PercentRealizedSpread_LR,
        mean(wPercentRealizedSpread_EOH) as PercentRealizedSpread_EOH,
        mean(wPercentRealizedSpread_CLNV) as PercentRealizedSpread_CLNV,
        mean(wDollarPriceImpact_LR) as DollarPriceImpact_LR,
        mean(wDollarPriceImpact_EOH) as DollarPriceImpact_EOH,
        mean(wDollarPriceImpact_CLNV) as DollarPriceImpact_CLNV,
        mean(wPercentPriceImpact_LR) as PercentPriceImpact_LR,
        mean(wPercentPriceImpact_EOH) as PercentPriceImpact_EOH,
        mean(wPercentPriceImpact_CLNV) as PercentPriceImpact_CLNV,
        sum(wDollarRealizedSpread_LR_SW) as waDollarRealizedSpread_LR_SW,
        sum(wDollarRealizedSpread_LR_DW) as waDollarRealizedSpread_LR_DW,
        sum(wPercentRealizedSpread_LR_SW) as waPercentRealizedSpread_LR_SW,
        sum(wPercentRealizedSpread_LR_DW) as waPercentRealizedSpread_LR_DW,
        sum(wDollarPriceImpact_LR_SW) as waDollarPriceImpact_LR_SW,
        sum(wDollarPriceImpact_LR_DW) as waDollarPriceImpact_LR_DW,
        sum(wPercentPriceImpact_LR_SW) as waPercentPriceImpact_LR_SW,
        sum(wPercentPriceImpact_LR_DW) as waPercentPriceImpact_LR_DW,
        sum(wDollarRealizedSpread_EOH_SW) as waDollarRealizedSpread_EOH_SW,
        sum(wDollarRealizedSpread_EOH_DW) as waDollarRealizedSpread_EOH_DW,
        sum(wPercentRealizedSpread_EOH_SW) as waPercentRealizedSpread_EOH_SW,
        sum(wPercentRealizedSpread_EOH_DW) as waPercentRealizedSpread_EOH_DW,
        sum(wDollarPriceImpact_EOH_SW) as waDollarPriceImpact_EOH_SW,
        sum(wDollarPriceImpact_EOH_DW) as waDollarPriceImpact_EOH_DW,
        sum(wPercentPriceImpact_EOH_SW) as waPercentPriceImpact_EOH_SW,
        sum(wPercentPriceImpact_EOH_DW) as waPercentPriceImpact_EOH_DW,
        sum(wDollarRealizedSpread_CLNV_SW) as waDollarRealizedSpread_CLNV_SW,
        sum(wDollarRealizedSpread_CLNV_DW) as waDollarRealizedSpread_CLNV_DW,
        sum(wPercentRealizedSpread_CLNV_SW) as waPercentRealizedSpread_CLNV_SW,
        sum(wPercentRealizedSpread_CLNV_DW) as waPercentRealizedSpread_CLNV_DW,
        sum(wDollarPriceImpact_CLNV_SW) as waDollarPriceImpact_CLNV_SW,
        sum(wDollarPriceImpact_CLNV_DW) as waDollarPriceImpact_CLNV_DW,
        sum(wPercentPriceImpact_CLNV_SW) as waPercentPriceImpact_CLNV_SW,
```

```
      sum(wPercentPriceImpact_CLNV_DW) as waPercentPriceImpact_CLNV_DW
        from Mid2
      group by symbol,date
      order by symbol,date;
quit;

/* Calculate Dollar-Weighted (DW) and Share-Weighted (SW) Realized Spreads
   and Price Impact */
data Mid3;
      set Mid3;
      DollarRealizedSpread_LR_SW=waDollarRealizedSpread_LR_SW/sumsize;
      DollarRealizedSpread_LR_DW=waDollarRealizedSpread_LR_DW/sumdollar;
      PercentRealizedSpread_LR_SW=waPercentRealizedSpread_LR_SW/sumsize;
      PercentRealizedSpread_LR_DW=waPercentRealizedSpread_LR_DW/sumdollar;
      DollarPriceImpact_LR_SW=waDollarPriceImpact_LR_SW/sumsize;
      DollarPriceImpact_LR_DW=waDollarPriceImpact_LR_DW/sumdollar;
      PercentPriceImpact_LR_SW=waPercentPriceImpact_LR_SW/sumsize;
      PercentPriceImpact_LR_DW=waPercentPriceImpact_LR_DW/sumdollar;
      DollarRealizedSpread_EOH_SW=waDollarRealizedSpread_EOH_SW/sumsize;
      DollarRealizedSpread_EOH_DW=waDollarRealizedSpread_EOH_DW/sumdollar;
      PercentRealizedSpread_EOH_SW=waPercentRealizedSpread_EOH_SW/sumsize;
      PercentRealizedSpread_EOH_DW=waPercentRealizedSpread_EOH_DW/sumdollar;
      DollarPriceImpact_EOH_SW=waDollarPriceImpact_EOH_SW/sumsize;
      DollarPriceImpact_EOH_DW=waDollarPriceImpact_EOH_DW/sumdollar;
      PercentPriceImpact_EOH_SW=waPercentPriceImpact_EOH_SW/sumsize;
      PercentPriceImpact_EOH_DW=waPercentPriceImpact_EOH_DW/sumdollar;
      DollarRealizedSpread_CLNV_SW=waDollarRealizedSpread_CLNV_SW/sumsize;
      DollarRealizedSpread_CLNV_DW=waDollarRealizedSpread_CLNV_DW/sumdollar;
      PercentRealizedSpread_CLNV_SW=waPercentRealizedSpread_CLNV_SW/sumsize;
      PercentRealizedSpread_CLNV_DW=waPercentRealizedSpread_CLNV_DW/sumdollar;
      DollarPriceImpact_CLNV_SW=waDollarPriceImpact_CLNV_SW/sumsize;
      DollarPriceImpact_CLNV_DW=waDollarPriceImpact_CLNV_DW/sumdollar;
      PercentPriceImpact_CLNV_SW=waPercentPriceImpact_CLNV_SW/sumsize;
      PercentPriceImpact_CLNV_DW=waPercentPriceImpact_CLNV_DW/sumdollar;
run;
```