KATHOLIEKE UNIVERSITEIT
LEUVEN

Faculty of Economics and
Applied Economics

Department of Economics

How to Match Trades and Quotes for Nyse Stocks?

by

Olivier VERGOTE

Econometrics

Center for Economic Studies
Discussions Paper Series (DPS) 05.10
http://www.econ.kuleuven.be/ces/discussionpapers/default.htm

March 2005

DISCUSSION
PAPER

How to Match Trades and Quotes for NYSE Stocks?

Olivier Vergote

K.U.Leuven

March 2005

Abstract

This paper proposes a new procedure to determine the time of the prevailing quote relative to the time of the trade for NYSE stock data obtained from the TAQ database. The procedure tests whether the quote revision frequency around a trade is contaminated by quote revisions triggered by a trade, and then determines the smallest timing adjustment needed to eliminate this contamination. An application to various stocks and sample periods shows that the time difference between trade and quote reporting lags varies across stocks and time. The procedure takes this variation into account and hence offers a stock- and time-specific update to the Lee and Ready (1991) 5-second rule.

JEL: G10, G15, G20

Keywords: TAQ, Market Microstructure, Trade and Quote Reporting Lags, Trade and Quote Matching Algorithm.

1 Introduction

The extraction programme of the Trade and Quote (TAQ) database produces separate files for trades and quotes, each with its own time stamp. For many research questions related to transaction data, one needs to construct the sequence of quotes and trades. In other words, one needs to merge both data sets and rank their records chronologically. This would be a straightforward operation, were it not that trades and quotes can be subject to different reporting lags. This problem was first reported by Lee and Ready (1991). The solution suggested by these authors is to add five seconds to the reported times of quotes. So far, most studies follow this suggestion or do not adjust at all.

In this paper it becomes clear, however, that the difference between trade and quote reporting lags has changed over time and can vary between stocks. This implies that the 5-second rule is too rigid. A new procedure is proposed to deal with the varying difference in lags between trade and quote reports.

Section 2 describes the trade and quote reporting procedures at the NYSE. These procedures are known to drive the lags. In Section 3, I discuss the 5-second rule and related reporting lags found in the literature. In Section 4, I discuss the data and have a preliminary look at the quote revision frequency around a trade. In Section 5, a new procedure to determine the prevailing quote at the time of a trade is presented. The procedure treats each stock individually and tests whether the quote revision frequency around a trade is contaminated by quote revisions triggered by a trade, and then determines the smallest timing adjustment needed to take this contamination into account. The procedure is applied to several stocks and sample periods between 1993 and 2003. Section 6 concludes.

2 Trade and quote reporting procedures

Anyone who noticed the rapid increase in market volume handled by the NYSE will find it natural that trade and quote reporting procedures have evolved over time. Subsequent chapters show that changes to reporting procedures affected the difference between trade and quote reporting lags. For later reference, this section contains a short overview of the procedures, which is mainly based on Hasbrouck et al. (1993) and information provided by nyse.com.

The NYSE records trades via the Consolidated Tape System (CTS) and revisions of the best quote via the Consolidated Quote System (CQS). The TAQ database is an extraction of these systems. The way that trades and quotes reach CTS and CQS has changed over the years.

After the 1987 market crash electronic workstations were introduced to deal with high volumes of trades and quotes. Until June 1989 the procedure was as follows: The specialist calls out the details of trades and new quotes as they happen. These trades and quotes are recorded by the specialist assistant or by floor reporters. It is the specialist who determines whether a floor reporter is involved in the recording of trades and quotes. The specialist assistant controls the Display Book, an electronic workstation that keeps track of all limit orders and incoming market orders and assists in the recording and dissemination of trades and quotation changes. The floor reporter, employed by the exchange, records trades and quotes by filling in boxes on a mark-sense card and feeding it into an optical reader. Trade reports travel through the Post Support System (PSS) to the exchange's Market Data System (MDS). MDS performs certain validation checks, before it sends the information to the CTS. Quote revisions travel through PSS to MDS and then to the CQS.

On 19 June 1989 the exchange began to abandon quote reporting by floor reporters. By September 1989, already 95% of quotes changed from the Display Book (Hasbrouck et al., 1993). By 2000, 99.9% of all quotes were updated by the Display Book, the exceptions

being for trading halts and other related events (NYSE, 2000 and 2001).

Through the years 1987-2001, the floor reporter was also less frequently used for recording trades, as trades were more and more Display Book reported. In 1994, the mark-sense card system was abolished and the floor reporter began to use a hand-held device to report trades. In the year 2000, 91% of all trades were Display Book reported (NYSE, 2000). Since 24 July 2001, the floor reporter position is eliminated and all trade reporting is done directly through the Display Book (NYSE, 2001). Prior to the elimination 99% of all trades were Display Book reported.

The best quote was not updated automatically when a trade affected the best quote until 27 May 2003, when the exchange introduced 'auto-quoting' for all stocks. This procedure implies that the NYSE automatically updates the NYSE's best bid or offer whenever a limit order is transmitted to the Display Book at a better price than the previous best bid or offer. When a trade occurs that involves the best bid or offer, the NYSE automatically updates the best bid or offer, and the associated depths, according to the specialist's book. Auto-quoting also includes adding size to the best quote as additional limit orders arrive and reducing size of the best quote as limit orders are executed or cancelled. Only in cases where the specialist trades for his own account are quotes not automatically updated. In conclusion, most quotes are automatically updated following each trade (NYSE, 2003a,b). In practice, only 5 percent of the quoting is performed manually on the Display Book (NYSE, 2003c).

3 Review of the literature

In an often-cited paper, Lee and Ready (1991) report a problem with the then existing reporting procedure to reconstruct the sequence of trades and quote revisions, see Section 2. As they point out, if the specialist assistant is faster in recording a quote revision than the floor reporter in recording a trade, the corresponding quote update can be recorded before the trade that triggered it. This is problematic if one seeks to determine the

prevailing quote at the time of a trade. Lee and Ready (1991) investigate the lag of trade reporting relative to quote reporting. Their results suggest using the prevailing quote at five seconds prior to the trade as the prevailing quote at the time of a trade. However, the dataset on which their analysis was based, dates back to 1988 and the result was obtained for a cross-section of 150 stocks. Furthermore, the reporting procedure changes pointed out in Section 2 undoubtedly affected reporting lags and it seems unlikely that the 5-second rule is universal.

Although Lee and Ready (1991, footnote 10) realize that the delay can vary with the sample period, their 5-second rule has been used in many studies based on TAQ data of the nineties: see e.g. Ball and Chordia (2001), Busse and Green (2002), Chan et al. (2002), Chordia et al. (2001, 2002), Easley et al. (2001), Edelen and Gervais (2003), Engle and Patton (2004), Huang and Stoll (2001), Kryzanowski and Zhang (2002), Nyholm (2003), Schultz (2000), Stoll (2000), Venkataraman (2001). This list is incomplete, but gives an idea of the popularity of the 5-second rule. I do not claim that the use of a better rule would alter the results of these studies, but few studies check the robustness of their results with respect to the 5-second rule.

Delays of NYSE trade reports have already been studied. Blume and Goldstein (1997) report a median delay of sixteen seconds for NYSE trades between execution and reporting, for the period July 1994 - June 1995. Peterson and Sirri (2003) use a two-week sample of 1997 of the NYSE System Order Database Daily File (SOD file), which contains details of order entry and execution. This allows them to compare the execution time and the reporting time of trades. They report a median delay for trades of only 2 seconds for NYSE stocks. Piwowar and Wei (2003) study the impact of different trade and quote matching algorithms on estimates of the effective spread for Nasdaq and NYSE stocks. In order to determine an optimal matching algorithm, they search for the trade time adjustment that minimizes the rate of small trades occurring outside the prevailing spread. Small trades are defined as 1000 shares or less. Their results clearly show the sensitivity of the effective spread estimates to the algorithm and that the sensitivity has increased

over time. However, their statistics are averages across several stocks.

Delays for Nasdaq stocks have been studied by Bessembinder (2003), Ellis et al. (2000) and Piwowar and Wei (2003). In general, there is no consensus on how to deal with reporting delays.

4 Evidence on trade and quote reporting lags

It is intuitively clear that trade reporting lags depend on the way trades are reported. Hasbrouck et al. (1993) already pointed out that Display Book reported trades have a much smaller reporting delay than trades reported by floor reporters. They report 15% of trades to be Display Book reported for a sample of five days in November 1990, and this percentage to be increasing fast. As shown in Section 5, the increasing popularity of the Display Book has decreased the overall lag of trade reporting over time. The success of the Display Book is also described in NYSE (2003c).

4.1 The data

I consider five 3-month samples, between 1993 and 2003, of NYSE trades and quotes from the TAQ database. The periods are April - June 1993, April - June 1997, April - June 2001, October - December 2001 and October - December 2003. In the discussion below, I refer to the different 3-month periods by the year only (1993, 1997, 2001a, 2001b and 2003). Most of the papers that make use of the Lee and Ready (1991) 5-second rule are based on samples that cover at least one of the first two sample periods. The third and fourth sample periods are just before and after the abolition of reporting by floor reporters. The last sample period is a period after auto-quoting was introduced.

I select five groups of five stocks based on the trading activity of the stocks during 2001a. Ranking the stocks in ascending order according to dollar volume traded that period, I first select three groups starting from the three stocks found at the 33%, 67% and 100% quantiles and each time moving down until I have five stocks that existed

during the period 1993 - 2003. Secondly, I take two more groups of actively traded stocks, because floor reporters were especially used for actively traded stocks, which allows us to have a better look at non Display Book reported trades. More specifically, using the same procedure, I take two more groups starting from the two stocks found at the 90% and 98% quantile. The five groups of stocks are {GE, IBM, EMC, PFE, TYC}, {F, DD, BBY, HAL, ADI}, {NSC, BCR, MYL, LTD, JCI}, {KWD, NAB, NC, ACG, ESL} and {PYM, GTY, SAF, NNJ, EY}, see Table 1. In Section 5 it turns out that the choice of the groups does not have any further implications, because the results are not significantly different between the groups.

For each stock only trades and quotes are selected that meet all of the following conditions:

- trades and quotes need to occur within the trading day: 9:30 - 16:00;
- trades need to be regular trades, which were not corrected, changed, or signified as cancel or error; this is indicated by a zero value of the correction indicator (CORR);
- trades need to be regular way or NYSE Direct+ trades; this is indicated by a blank or 'E' value of the Condition indicator;
- quotes need to stem from normal trading conditions; this is indicated by the Mode indicator taking the value 1, 2, 3, 6, 10, 12 or 18.

Another feature that is used below is whether trades are Display Book reported. This is indicated by specific values of the G127 indicator; see the TAQ2 user's guide for more details.

4.2 Quote revision frequencies around trades

In order to accurately determine the prevailing quote at the time of a trade, it is necessary that each quote update, triggered by a trade, can be distinguished from other quote revisions. If the quote update is recorded before the trade, then one would take the prevailing quote at one second before the quote update as the prevailing quote at the time of the trade. Unfortunately, the data does not contain a field that links trades with

Table 1: Dollar volume of selected stocks for 2001a

Symbol	Dollar Volume	Symbol	Dollar Volume
GE	52,560,633,160	LTD	1,557,559,751
IBM	49,375,424,041	JCI	1,555,545,716
EMC	31,445,660,352	KWD	126,363,410
PFE	27,688,446,737	NAB	124,842,589
TYC	26,389,434,725	NC	122,857,484
F	8,028,909,391	ACG	120,519,787
DD	7,944,040,652	ESL	115,849,712
BBY	7,683,355,617	PYM	8,707,448
HAL	7,657,386,156	GTY	8,654,034
ADI	7,606,474,675	SAF	8,648,920
NSC	1,566,265,741	NNJ	8,637,559
BCR	1,565,307,184	EY	8,501,009
MYL	1,562,611,688		

their quote update. Moreover, it is not always possible to detect the link between a trade and its quote update from the trade and quote sequence, by comparing trade sizes with changes in quote depth and trade prices with quote prices. This is caused by active trading and cancellations of limit orders, which complicate the interpretation of the trade and quote sequence.

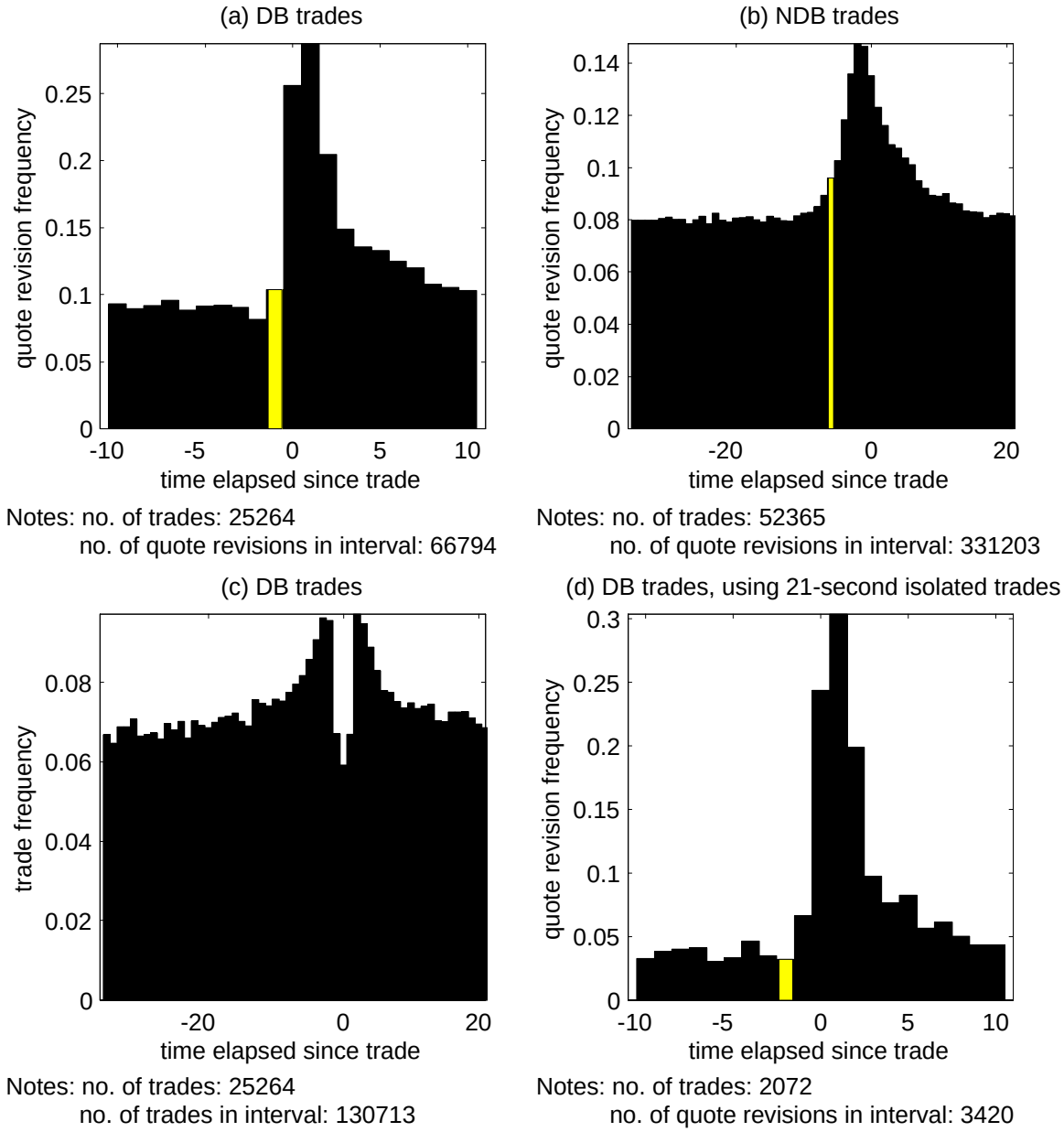
Although we cannot distinguish a quote update triggered by a trade from other quote revisions, we can compute the average timing of the quote updates relative to the trades. For this purpose, I compute the frequency of quote revisions at each second of a 20-second interval around a trade, for each stock and each 3-month period. If quote revisions are triggered by trades, then one would expect to find a peak in this frequency distribution at the time these quotes are reported. Other quote revisions are either the result of incoming orders or cancelled orders and their timing should be approximately independent of the timing of trades. The 20-second intervals $[-10, 10]$ centered around trades can overlap, which implies that some quote revisions are counted more than once. I come back to this below. As the quote updates, triggered by non Display Book reported trades, are expected to occur further away from the trade, I look at 55-second intervals $[-35, 20]$ around this type of trades.

As an example, Figures 1a and 1b display the frequency distribution for BBY Display Book (DB) reported and non Display Book (NDB) reported trades of 2001a, respectively. The horizontal axis shows the number of seconds before and after the recording of a trade. The vertical bars represent the quote revision frequency around a trade, computed as the number of quote revisions divided by the number of trades. A grey bar indicates the time of the prevailing quote that results from the procedure described in Section 5. We notice a clear difference between both figures, which indicates that DB reported trades and NDB reported trades are subject to different lags. Both figures show a hump, which indicates the time relative to a trade at which quote revisions triggered by trades are recorded. The humps are situated at different times relative to the time the trade is recorded. In Figure 1a, the hump is steep and shows that quote revisions triggered by trades are recorded at the time the trade is recorded or up to a few seconds later. Hence, to avoid taking a quote update triggered by a trade as the prevailing quote at the time of that trade, it would seem optimal to take the prevailing quote *one second before the trade* (-1). For the NDB trades in Figure 1b, however, quote revisions triggered by a trade are also recorded before the trade. Again, to avoid taking a quote update triggered by a trade as the prevailing quote at the time of that trade, it appears to be a good choice to take the prevailing quote about *ten seconds before the trade* as the prevailing quote at the time of the trade (-10). The marked difference between the two figures suggests that it is better to treat the two types of trades separately when determining the prevailing quote.

Unfortunately, there is a problem with the type of figures above if overlapping intervals around trades are used, because it implies that quote revisions due to one trade can be counted as a quote revision around another trade. If the arrival of trades around a trade is uniformly distributed, then this has no side effects, because the quote revisions that the trades trigger would also be uniformly distributed. Figure 1c gives an example of the frequency distribution of trades around a DB trade. Across all stocks, the distribution is typically bimodal and approximately symmetric around zero, with a steep slope between zero and the mode at ± 3 to 4 seconds away from the trade, and after the mode the

frequency decays (often to a higher level than that close to the trade). As a consequence, the quote revisions that these trades trigger are not expected to be uniformly distributed. The latter causes a bias in the frequency of quote revisions around a trade. Corresponding with this expectation, I find in quite a few cases that the quote revision frequency is relatively high in the 10-second interval before the trade and drops just before the trade. Therefore, I use non-overlapping intervals in the analysis below. In order to avoid any overlap of the 20-second intervals for Display Book reported trades, I select trades that are at least 21 seconds away from any other trade. For the non Display Book reported trades, I take 56-second isolated trades, because a high quote revision frequency is found up to 30 seconds before the trade and 20 seconds after the trade and in Section 5 the first five seconds are used for computations. It implies far fewer observations and less smooth figures, but avoids the possible bias in the number of quote revisions around a trade. Compare e.g. Figure 1a with Figure 1d, where the isolated trades in the latter figure have a lower frequency before (and after) the trade compared to the frequency at the time of the trade. The figures also show how the bias can lead to a different result when determining the time of the prevailing quote (see Section 5). In Figure 1a, the grey bar indicates that the prevailing quote is found at -1, while in Figure 1d it is found at -2.

Figure 1: Quote revision and trade frequencies around BBY trades, 2001a.



5 Towards a new adjustment rule

The distribution of quote revisions around a trade typically shows an increase in the frequency of quote revisions at some time before the trade. Thus, we seek a rule for determining the start of this increase, and then we can take the prevailing quote at one

second before the increase as the prevailing quote at the time of the trade.

I suggest the following rule. Let s be the elapsed time, in seconds, since an isolated trade, e.g. 21-second isolated trades with $s = -10, ..., 10$. Let t be the second in this interval at which quote revisions triggered by a trade start to arrive. Then $t-1$ is the second at which we can find the prevailing quote. Assume that quote revisions that are not triggered by a trade arrive according to a Poisson process with rate λ_s at s. A test can then be developed based on the idea that λ_s is expected to remain approximately constant over time, while estimates of λ_s will be biased upward from t onward due to the arrival of quote revisions triggered by trades.

Let q_s be the number of quote revisions at s and n the total number of isolated trades. Then, λ_s can be estimated by

$$\hat{\lambda}_s = \frac{q_s}{n}.$$

Let λ_s^p be the partial mean of λ_s, for example

$$\lambda_{-5}^p = \frac{\lambda_{-10} + ... + \lambda_{-5}}{6},$$

and let $\hat{\lambda}_s^p$ be the partial mean of $\hat{\lambda}_s$. I will interpret a statistically significant increase in $\hat{\lambda}_s$, above $1.2\hat{\lambda}_{s-1}^p$ as evidence that, at s, at least part of the quote revisions are triggered by a trade.

The null hypothesis that λ_s is not more than 20% higher than λ_{s-1}^p, or that $s < t$, can be tested by:

$$\frac{\hat{\lambda}_s - 1.2\hat{\lambda}_{s-1}^p}{s.e.} \overset{a}{\sim} N(0,1).$$

Under the null hypothesis ($s < t$), the standard error can be estimated as

$$s.e. = \left(\frac{\hat{\lambda}_s}{n} + \frac{1.44}{n(s+10)^2} \sum_{j=-10}^{s-1} \hat{\lambda}_j \right)^{1/2}.$$

The null hypothesis $H_0 : \lambda_s \leq 1.2\lambda_{s-1}^p$ is sequentially tested for $s = -9, ..., 10$ and a

significance level of 2.5%. The smallest s for which the null hypothesis can be rejected is inferred as t. The prevailing quote can be found at $t-1$.

Judging from the quote revision distributions, λ_s appears to vary within the interval before a trade. This motivates the 20% bound above λ^p_{s-1} used when determining the prevailing quote. It implies that an increase of λ_s of less than 20% is interpreted as no or not enough contamination by quote revisions triggered by a trade.

I apply this method to determine the adjustment rule for all 25 stocks in all 5 periods. In case of NDB trades, I let $s = -35, ..., 20$. In Figures 1 and 2, a grey bar indicates the time of the prevailing quote that results from the procedure. As not all figures for all stocks can be presented here, the derived adjustment rules are presented in table 2. The table contains the time of the prevailing quote relative to the time of the trade (in seconds), computed for Display Book reported and non Display Book reported trades. A blank indicates that there are no trades of that type, i.e. not even among the non-isolated trades. If there are less than 100 trades of a certain type, then that case is ignored and is indicated by $\cdot$. For the last two periods, non Display Book reported trades are the exception and therefore they are not included in table 2. With less than 100 isolated trades or less than 200 quote revisions in the interval centered around isolated trades, one cannot obtain a clear picture. The best solution would be to increase the sample size, but if this is not possible, one can also use the figures based on all trades. I stick to the sample size to construct table 2 and indicate the values for which all trades were used by $^\circ$. Although the distribution of quote revisions based on all trades may be slightly biased, it is expected to provide a better idea about what an appropriate adjustment should be than simply applying the 5-second rule or not adjusting at all. Results based on less than 200 (500) DB (NDB) trades or less than 300 (750) quote revisions around these trades are indicated by *. Finally, there are a few cases for which, even when using all trades, there are less than 200 quote revisions around the trades. This leads to an unclear figure and I indicate these cases by $\times$. Now I will discuss the results for the different periods.

Table 2 shows that by 1993 the Display Book is well in use, because for quite a few

Table 2: Time of the prevailing quote relative to the time of the trade (in seconds)

	1993		1997		2001a		2001b	2003
	DB	NDB	DB	NDB	DB	NDB	DB	DB
GE	·	−21	·	−3	−2*	−3°	−1	−2
IBM	·	−2	0°	−8	−1	−3°	−1	−2
EMC	0	−15*	−2		−1		−1	−2
PFE	×	−17	·	−6	−1	·	−1	−2
TYC	0		−1		−2		−1	−2
F	1*	−15	−1*	−6	−1		−1	−2
DD	−2	·	−2	4°*	−1		−2	−2
BBY	−2	−10°*	−1°*	−13	−2	−9	−1	−2
HAL	×	−10	−1	−6	−1	−3°	−2	−2
ADI	·	−32	−2	−3°	−1		−1	−2
NSC	−1		−2	−7°*	−1		−2	−2
BCR	−1	−18	−6	−8*	−2	−1°*	−2	−2
MYL	−1		−1		−1		−1	−2
LTD	·	−14	·	−10	−1		−2	−2
JCI	−4	−21°*	−5		−2		−2	−2
KWD	−1		−2		−1		−2	−2
NAB	·	×	−1		−2		−1	−2
NC	−1		−6		−2		−1	−2
ACG	−3		−3		−2		−2	−2
ESL	−2*		−2		−1		−1	−2
PYM	·	−13*	−2		−2		−3	−2
GTY	−1*		−1*		−7		−1	−2
SAF	−1		−2		−2		−2	−2
NNJ	×		−2		−1		−1	−2
EY	0		−2		−2		−2	−2

Notes: Blank = no trades; · = exceptions among all trades (<100 trades), which are ignored; ° = less than 100 isolated trades or less than 200 quote revisions in the interval around isolated trades, in which case all trades are used; * = based on less than 200 (500) DB (NDB) trades or less than 300 (750) quote revisions in the interval around DB (NDB) trades; × = even using all trades there are less than 200 quote revisions in the interval around trades.

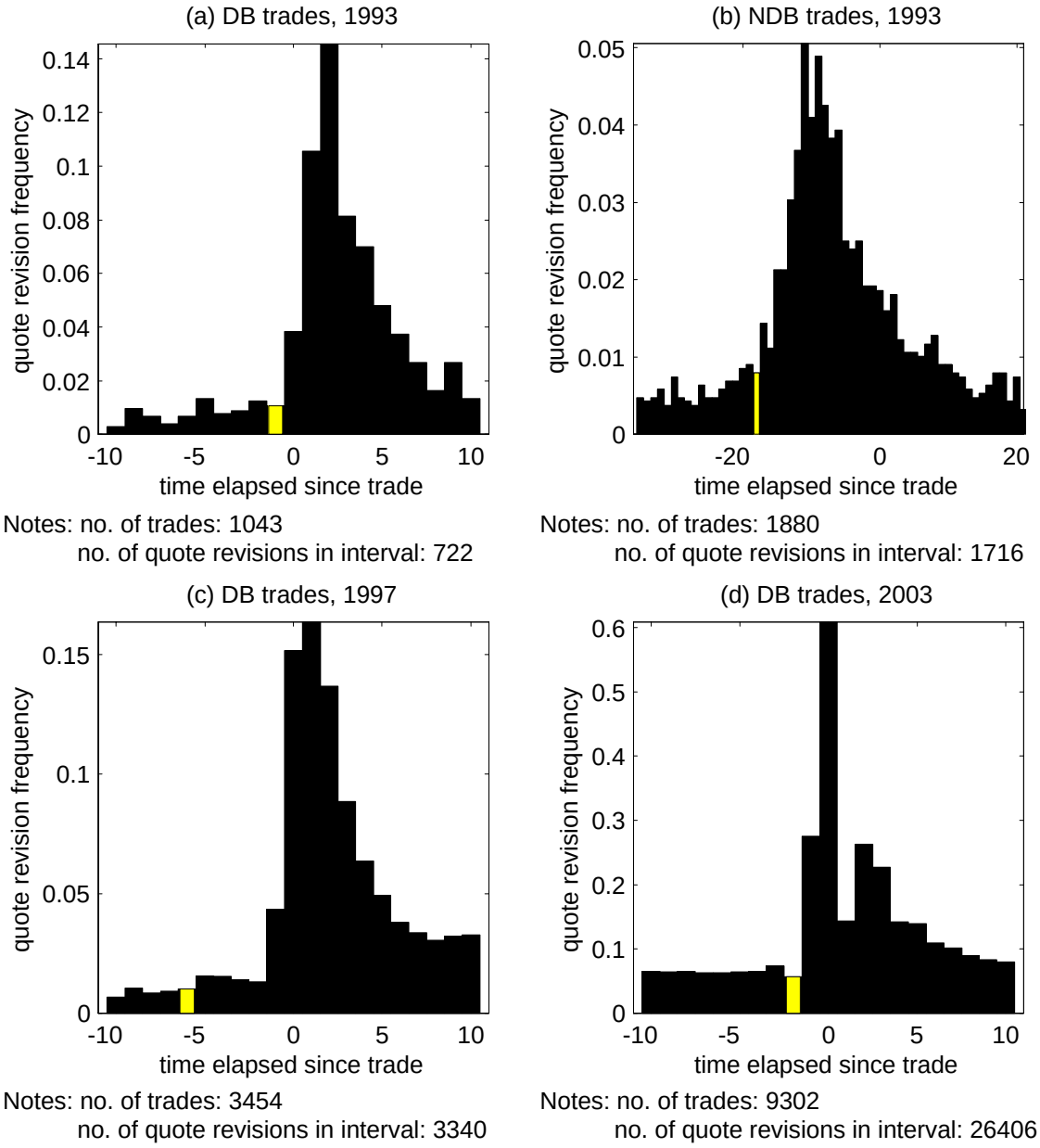
stocks all trades are Display Book reported. For the DB trades of 1993, depending on the stock, the adjustment rules vary moderately between 0 and -4 seconds; see e.g. Figure 2a for BCR. However, there is one case for 1993 (and one for 1997) in table 2 for which the adjustment rule is positive. The reason is that the increase in the frequency of quote revisions, caused by quote revisions triggered by trades, is not really clear from the figures due to a low number of observations. As a result the algorithm passes the time of the trade and proposes a positive adjustment rule.

Contrary to DB trades, the adjustment rules for NDB trades of 1993 are much larger and go up to 32 seconds; see e.g. Figure 2b. IBM is a major exception to this with an adjustment rule of -2. The reason is that the distribution of quote revisions around NDB trades looks like the distribution around DB trades. Presumably it does concern Display Book reported trades, but have they not been indicated as such in the TAQ database.

In general, it is clear from these figures that the problem of reporting lags is severe for NDB trades, and in fact accurately determining the prevailing quotes for these trades is not really feasible. Furthermore, if a trade report can be delayed up to 32 seconds compared to the associated quote report, then this demands a considerable adjustment. At the same time, however, such a large adjustment implies that we can be far too conservative for trade reports that have no delay relative to the quote update report.

For Display Book reported trades of 1997, depending on the stock, I mainly find that the prevailing quote can be found 1 or 2 seconds before the trade. There are a few cases, however, where the algorithm stops at -6 or -5, because λ_s has already moved up by significantly more than 20%, which is interpreted as contamination by quote revisions triggered by a trade; see e.g. Figure 2c.

Figure 2: Quote revision frequency around BCR trades (using isolated trades).



When we compare the results for the same stock between the 1993 and 1997 periods, it appears that the relative trade report delays have decreased. Apparently, the switch from the mark-sense card system to a hand-held device for the floor reporter decreased the trade reporting lags; see section 2. This also suggests an extension to the procedure described above that might be useful for small samples. If the adjustment is relatively

small, then quote updates triggered by a trade are found relatively close to the trade. As a result, the interval over which trades are isolated can be decreased. This would increase the number of observations and the procedure can be run a second time.

For 2001a, the trades are predominantly Display Book reported, except for GE, for which 90% of trades are still non Display Book reported. For most stocks, the prevailing quote for Display Book reported trades is found at 1 or 2 seconds before the trade, and the same holds for 2001b. All trades are Display Book reported from mid-2001 on, although a few exceptional trades can be found among the TAQ data.

In the year 2003 autoquoting is introduced and, as expected, we notice more quote revisions at the time of the trade; see e.g. Figure 2d. It appears, however, that quote updates can still be reported before a trade is reported, because a substantial portion of the quote revisions is recorded one second before the trade is recorded. The reason is not clear, but a small discrepancy between the clocks of CTS and CQS can already have this effect. The optimal adjustment is therefore –2 seconds. Contrary to the other periods, the shape of the figure and the adjustment rule are very stable across stocks. The stabilising effect of autoquoting on the reporting of quotes makes determining the prevailing quote easier. This result suggests a new adjustment rule for NYSE stock data from mid-2003 onward. That is, to take the prevailing quote two seconds before the trade as the prevailing quote at the time of the trade.

The improvement of changing the adjustment rule from five to two seconds depends on whether there are actually changes in the quote between five and two seconds prior to a trade. In order to study this I compute the prevailing quote at the time of a trade based on the two rules separately, for each stock in the 2003 sample. The two rules yield the same ask quote between 69% and 98% of the time (depending on the stock), the same bid quote between 70% and 96% of the time, and the same spread between 62% and 94% of the time. Thus, the different rules lead to different quotes in a substantial number of cases.

6 Conclusion

This paper proposed a new procedure to match trades and quotes of NYSE stocks. The procedure tests whether the quote revision frequency around a trade is contaminated by quote revisions triggered by a trade, and then determines the smallest timing adjustment needed to take this contamination into account. The procedure was applied to a sample of 25 stocks in 5 sample periods. The results show that the difference between trade reporting lags and quote reporting lags varies across stocks and time. The variation can be mainly explained by changes in the reporting procedures of the NYSE and the co-existence of two reporting systems for trades, i.e. trades are Display Book reported or not. The non Display Book reported trades usually require a larger adjustment to match trade and quote times.

In summary, given a sample of NYSE stocks from the TAQ database, it is recommended to take the following issues into account when applying the new procedure. For samples prior to mid-2001 it is better to treat Display Book reported trades separately. In case of large sample periods, it is better to split the sample in order to take further differences across time into account. To sum up, it is recommended to determine a separate adjustment rule for each trade type, each stock, and each sub-sample.

Finally, an interesting observation is that the system of auto-quoting, in which quotes are automatically updated since mid-2003, has a stabilising effect on the reporting of quotes. In particular, taking the prevailing quote at two seconds before the trade as the prevailing quote at the time of the trade appears an appropriate adjustment rule for all stocks.

References

[1] Ball, C.A., and T. Chordia (2001): "True Spreads and Equilibrium Prices," *Journal of Finance*, 56(5), 1801-1835.

[2] Bessembinder, H. (2003): "Issues in Assessing Trade Execution Costs," *Journal of Financial Markets*, 6, 233-257.

[3] Blume, M., and M. Goldstein (1997): "Quotes, Order Flow, and Price Discovery," *Journal of Finance*, 52, 221-244.

[4] Busse, J.A., and T.C. Green (2002): "Market Efficiency in Real Time," *Journal of Financial Economics*, 65(3), 415-437.

[5] Chan K., Y.P. Chung, and W.M. Fong (2002): "The Informational Role of Stock and Option Volume," *Review of Financial Studies*, 15(4), 1049-1075.

[6] Chordia, T., R. Roll, and A. Subrahmanyam (2001): "Market Liquidity and Trading Activity," *Journal of Finance*, 56(2), 501-530.

[7] Chordia, T., R. Roll, and A. Subrahmanyam (2002): "Order Imbalance, liquidity, and Market Returns," *Journal of Financial Economics*, 65(1), 111-130.

[8] Easley, D., M. O'Hara, and G. Saar (2001): "How Stock Splits Affect Trading: A Microstructure Approach," *Journal of Financial and Quantitative Analysis*, 36(1), 25-51.

[9] Edelen, R., and S. Gervais (2003): "The Role of Trading Halts in Monitoring a Specialist Market," *Review of Financial Studies*, 16(1), 263-300.

[10] Ellis, K., R. Michaely, and M. O'Hara (2000): "The Accuracy of Trade Classification Rules: Evidence from Nasdaq," *Journal of Financial and Quantitative Analysis*, 35(4), 529-552.

[11] Engle, R.F., and A.J. Patton (2004): "Impacts of Trades in an Error-Correction Model of Quote Prices," *Journal of Financial Markets*, 7(1), 1-25.

[12] Harris, L. (2003): *Trading and Exchanges: Market Microstructure for Practitioners*, Oxford University Press.

[13] Hasbrouck, J., G. Sofianos, and D. Sosebee (1993): "New York Stock Exchange Systems and Trading Procedures," NYSE Working Paper 93-01.

[14] Huang, R.D., and H.R. Stoll (2001): "Exchange Rates and Firms' Liquidity: Evidence from ADRs," *Journal of International Money and Finance*, 20(3), 297-325.

[15] Kryzanowski, L., and H. Zhang (2002): "Intraday Market Price Integration for Shares Cross-Listed Internationally," *Journal of Financial Quantitative Analysis*, 37(2), 243-369.

[16] Lee, C.M., and M.J. Ready (1991): "Inferring Trade Direction from Intraday Data," *Journal of Finance*, 46, 733-746.

[17] Nyholm, K. (2003): "Inferring the Private Information Content of Trades: A Regime-Switching Approach," *Journal of Applied Econometrics*, 18(4), 457-470.

[18] NYSE (2000): "Stock Market Activity," *2000 Factbook*.

[19] NYSE (2001): "Stock Market Activity," *2001 Factbook*.

[20] NYSE (2003): *2003 Annual Report*.

[21] NYSE (2003a): *Information Memo*, 03-21, May 15.

[22] NYSE(2003b): *Information Memo*, 03-27, June 11.

[23] NYSE (2003c): "How to Compete with the New York Stock Exchange," *The Exchange*, 10(8).

[24] Peterson, M., and E. Sirri (2003): "Evaluation of the Biases in Execution Cost Estimation Using Trade and Quote Data," *Journal of Financial Markets*, 6, 259-280.

[25] Piwowar,. M.S., and L. Wei (2003): "The Sensitivity of Effective Spread Estimates to Trade-Quote Matching Algorithms," NYSE Working Paper,1-45.

[26] Schultz, P. (2000): "Stock Splits, Tick Size, and Sponsorship," *Journal of Finance*, 55(1), 429-450.

[27] Stoll, H.R. (2000): "Friction," *Journal of Finance*, 55(4), 1479-1514.

[28] Teweles, R.J., and E.S. Bradley (1998): *The Stock Market,* John Wiley & Sons, 7th Edition.

[29] Venkataraman, K. (2001): "Automated versus Floor Trading: An Analysis of Execution Costs on the Paris and New York Exchanges," *Journal of Finance*, 56(4), 1445-1485.