



ELSEVIER

Journal of Financial Markets 6 (2003) 233–257

Journal of
FINANCIAL
MARKETS

www.elsevier.com/locate/econbase

Issues in assessing trade execution costs[☆]

Hendrik Bessembinder*

*David Eccles School of Business, 1645 E. Campus Center Drive, University of Utah,
Salt Lake City, UT 84112, USA*

Abstract

This study assesses the sensitivity of trading cost estimates derived from publicly-available trade and quote data to two methodological issues: the time adjustment made before comparing trades to quotes, and the procedure used to designate trades as buyer or seller-initiated. The results indicate that making *no* allowance for trade reporting lags is optimal when assessing whether trades are buyer or seller initiated, for both Nasdaq and NYSE stocks. However, trade prices are best compared to earlier quotations when assessing trade execution costs, in order to capture the effect of systematic quotation revisions in the seconds before trades are reported. A technique for inferring trade direction recommended by Ellis et al. (*J. Financial Quant. Anal.* 35 (2000) 529) leads to significantly smaller estimates of trading costs than the well-known Lee and Ready (*J. Finance* 46 (1991) 733) algorithm. Despite the sensitivity of trading cost measures to these methodological issues, inference as to whether the Nasdaq dealer market or the NYSE auction market provides lower trade execution costs is not sensitive.

© 2002 Elsevier Science B.V. All rights reserved.

JEL classification: G1

Keywords: Trade execution costs; Market structure; Cost measurement

0. Issues in assessing trade execution costs

What determines the price that investors pay to have their orders completed in financial markets? Obtaining accurate measures of trade execution costs and

[☆]The author thanks Jeff Bacidore, Paul Irvine, Bruce Lehmann, Sunil Wahal, an anonymous referee, and seminar participants at the University of Florida for many useful comments. The New York stock exchange provided data and financial support for this study.

*Tel.: +1-801-581-8268; fax: +1-801-581-7214.

E-mail address: finhb@business.utah.edu (H. Bessembinder).

assessing the reasons for their systematic variation is important to individual investors, portfolio managers, those evaluating brokerage firm or financial market performance, and corporate managers considering where to list their shares. Interest in measuring trading costs appears to have increased in recent years. In November of 2000 the U.S. SEC issued Rule 11Ac1-5, which requires each U.S. market center to provide monthly reports detailing execution quality for its trades in common stocks.¹

The ideal measure of trading costs may be the “implementation shortfall” described by [Perold \(1988\)](#). This measure involves comparing the value-weighted average trade price (after commissions and fees) for an investor’s order to a reference price (e.g. the quote midpoint) at the time of the trading decision. Although investors could conceivably construct such a measure for their own trades, researchers and regulators who are interested in measuring trading costs in aggregate do not have sufficient information to do so.

Most broad-based studies of trading costs and market quality are conducted using publicly-available databases that are limited to trade prices and quotations. One shortcoming of trade and quote data is that whether a trade was initiated by a buyer or a seller must be imperfectly inferred from the data. A second issue is that, although trade prices can be readily compared to quotes in effect at the trade report time, the appropriate comparison might be to quotes in effect at an earlier time—such as the time of the trading decision or at the time the order arrived at the market, and these times are generally not known. This paper evaluates the practical impact of two methodological choices made when estimating trading costs from public data: the method used to classify trades as resulting from customer buy versus sell orders, and the relationship between trade report times and the time of the quote chosen as a reference point.

Some researchers have been able to refine their estimates of trading costs by using customer order data in their studies. However, access to the order data is often restricted, and the order data that has been available typically applies to a limited set of institutional customers, or to a small group of stocks.² Also, as [Lee and Radhakrishna \(2000\)](#) and [Finucane \(2000\)](#) have emphasized, even having access to proprietary order data does not allow for all trades to be identified as buyer or seller initiated. Individual market centers will use their internal order data to measure trade execution costs as required by SEC rule 11Ac1-5. But, as [Bacidore et al. \(2001\)](#) emphasize, measures of trade execution quality are quite sensitive to detailed methodological choices, including the method used to average across stocks, the timing and precise definition of the reference quote, which orders and trades are

¹The new SEC rules can be reviewed at <http://www.sec.gov/rules/final/34-43590.htm>.

²The Trade, Order, Record, and Quote (TORQ) database is publicly available, and has been used by several researchers, including [Chung et al. \(1999\)](#), [Finucane \(2000\)](#), [Odders-White \(1999\)](#) and [Lee and Radhakrishna \(2000\)](#), among others. However, the TORQ database is limited to 144 NYSE stocks during a three-month period during 1990 and 1991. [Peterson and Sirri \(1999\)](#) and [SEC \(2000\)](#) use proprietary data obtained by the Securities and Exchange Commission. Studies that have used proprietary databases pertaining to specific institutional investors include [Keim and Madhavan \(1997\)](#) and [Chan and Lakonishok \(1997\)](#).

included in the analysis, and order size. It is unclear the extent to which execution quality reports prepared by individual market centers will be comparable to each other. Researchers who wish to compare market quality measures while using completely uniform definitions and methods are likely to have to construct their own measures, and to continue using the publicly available trade and quote data. The same will be true for those who wish to evaluate market quality measures for time intervals finer than one month or those who wish to study measures (e.g. net order imbalances) other than those required by rule 11Ac1-5.

The central methodological issue considered here concerns the relative timing of trades and quotes. Lee and Ready (1991) recommend comparing trade prices to quotes in effect five seconds before the trade report time to allow for delays in trade reporting. However, comparing trade prices to an earlier reference quote might be justified even if trades were reported without delay. Traders are generally concerned with the possibility of adverse price changes between the time of their trade decision and trade execution. The possibility of adverse pre-trade price movements underlies Perold's recommendation to compare trade prices to a benchmark at the time of the trading decision, as well as the common use in the practitioner literature of prior day quotes or trade prices as the benchmark. Comparing trade prices to earlier benchmark quotes potentially captures the effect of systematic price movements ahead of trades.

Researchers who have adopted the five second (or other fixed time) lag recommended by Lee and Ready generally use the adjustment *both* when inferring whether trades are buyer or seller-initiated and for selecting a quote to be used as a benchmark to measuring effective trade execution costs. However, the optimal amount by which to adjust trade times before comparing trade prices to quotes could differ depending on the application. Suppose, for example, that trade report times lag actual trade execution times by five seconds, *and* that quotes tend to be systematically revised in the fifteen seconds before trades are completed. Then, comparing trade prices to quotes five seconds before the trade report time would be optimal for inferring trade direction. However, comparing trades to quotes in effect twenty seconds before the trade report time would provide a more accurate measure of trade execution cost, including the effect of *pre-trade price impact*. This study assesses the effect on measured trading costs of comparing trade prices to quotes in effect from zero to thirty seconds prior to the trade report time. I also compare measures of trading costs obtained when using the well-known algorithm recommended by Lee and Ready (1991) to assign trades as buyer or seller-initiated to those obtained using an alternate procedure recommended by Ellis et al. (2000).

Roll (1984) introduces a technique for inferring trade execution costs from the serial covariance of price changes. His method does not require that trades be signed or matched to quotation data. Schultz (2000) reports that the Roll technique provides good estimates of trade execution costs for his sample of Nasdaq stocks. This study assesses whether difficulties arising from the need to sign trades and to match trades with prevailing quotes can be avoided by simply using the Roll method in broader samples that include data from the New York Stock Exchange (NYSE) as well as Nasdaq.

These measurement issues are considered in the context of a broad comparison of trade execution costs across the Nasdaq dealer market and the NYSE's specialist-based auction market. The study examines 300 stocks traded on the Nasdaq stock market and 300 matched NYSE stocks during the July to December 1998 period.

The results of the study can be summarized as follows. First, measures of rates at which trades are executed at prices better (trades receive "price improvement") or worse (trades receive "price disimprovement") than the quotations are quite sensitive to whether trade prices are compared to contemporaneous or previous quotes. Comparing trade prices to earlier quotes decreases the percentage of trades that appear to receive price improvement while sharply increasing the percentage of trades that appear to be disimproved. Comparison of these results to those obtained in recent studies (Bacidore et al., 2001; Peterson and Sirri, 2002; and Ellis et al., 2000) that use proprietary order data suggests that contemporaneous comparisons are optimal when assigning trades as buyer or seller initiated.

Second, measures of effective trading costs increase with longer time adjustments because quotations systematically rise (fall) in the seconds before customer buy (sell) trades are reported. If the quotation movement prior to the trade report time occurs after the trading decision but before trade execution, then the increase in measured trading costs with a longer time adjustment is real, not illusory. Third, the Ellis, Michaely, and O'Hara method for inferring trade direction leads to smaller estimates of trading costs compared to the Lee and Ready method, for both NYSE and Nasdaq-listed stocks. Fourth, the trading cost estimator due to Roll (1984) provides estimates of effective trading costs on the Nasdaq market that are very similar to those obtained when comparing trade prices to quotations, a result that is consistent with the findings of Schultz (2000). However, the Roll estimates do not correspond as closely to the quote-based estimates for NYSE stocks or for large trades on either market, implying that the difficulties arising from the need to match trades with quotes and to assign trade direction cannot be avoided by simply using the Roll procedure instead.

Finally, despite the sensitivity of trading cost measures to the methodological issues considered here, inference as to whether the dealer market or the auction market provides lower trade execution costs is not sensitive. Trade execution costs are higher on the dealer market regardless of the method used to sign trades or of allowances for trade reporting lags. The cross-market differences in average execution costs across markets are modest for large-capitalization stocks, but are substantially greater for small-capitalization stocks.

This analysis, which relies wholly on the publicly available trade and quote data, complements those provided by Bacidore et al. (2001); Peterson and Sirri (2002) and Werner (2002), each of which also exploit proprietary NYSE system order data. The results here could be described as an extensive sensitivity analysis, assessing how measured costs vary when research methods are altered. The Bacidore, Ross, and Sofianos, Peterson and Sirri, and Werner analyses provide partial insights into why measured trading costs for NYSE stocks vary as they do. For example, Peterson and Sirri verify that quotes tend to move adversely in the seconds between order arrival and trade execution times. A complete understanding of the reasons for variation in

measured trading costs would also require knowledge of the behavior of quote changes in the time between trade decision and order arrival times, for both NYSE and Nasdaq stocks.

This paper is organized as follows. Section 1 describes the sample firms and the trade and quote data that are used. Section 2 describes the measures of trading costs and procedures for inferring trade direction. Section 3 reports the main empirical results, and Section 4 concludes.

1. The sample

This study focuses on trades in 300 Nasdaq and 300 NYSE-listed common stocks during the period July 1 to December 31, 1998. The Nasdaq and NYSE samples are matched based on beginning-of-sample market capitalization, and include subsets of large, medium, and small capitalization stocks.³ Capitalization matches are generally quite close—the median (mean) absolute difference in market capitalization between matched NYSE and Nasdaq stocks is 0.13% (0.28%). All but ten of the 300 firm pairs are matched within 1%. Sample selection procedures are described in Appendix A. [Table 1](#) provides some sample summary statistics.

I obtain all trades and quotes in the 600 sample stocks that occur during normal trading hours (9:30 a.m. to 4:00 p.m.) between July 1 and December 31, 1998 from the Trade and Quote (TAQ) database. A small number of these (less than 2%) were omitted from the sample because they did not pass the error filters described in [Bessembinder \(1999\)](#). I also exclude non-NYSE quotations for NYSE stocks and non-Nasdaq quotations for Nasdaq stocks.⁴ The final sample includes 43.5 million trades and 25.4 million quotes.

Sample Nasdaq stocks are traded more frequently, averaging 946 trades per stock/day in the full sample, compared to 215 trades for sample NYSE stocks. However, trading frequencies are not directly comparable across samples due to interdealer trading on the Nasdaq market and because customer orders interact directly more often on the NYSE specialist market than on the Nasdaq dealer market. Quotes are also updated more frequently on Nasdaq, 1094 times per stock/day, compared to 496

³ A natural question is whether key results of this study would vary if the Nasdaq and NYSE samples were matched by criteria other than market capitalization. The large sample sizes (69 million trades and quotes in the present sample) involved made sensitivity testing impractical. The available evidence indicates that the results of cross-market comparison studies are quite robust to reasonable alternative matching procedures. [Chung et al. \(2000\)](#) match Nasdaq and NYSE samples based on share price, return volatility, and trading activity to compare trading costs in the wake of the 1997 Nasdaq market reforms, and obtain results very similar to those of [Bessembinder \(1999\)](#), who matched based on market capitalization. [LaPlante and Muscarella \(1997\)](#) investigate several alternatives to size-matching in their smaller scale (ten firms from each market) study of large trade execution costs, and report that results are very insensitive to the use of alternative matching techniques.

⁴ A large number (over 21 million for the present sample) of quotations for NYSE stocks originate at the regional stock exchanges or with NASD dealers. However, as [Blume and Goldstein \(1997\)](#) document, these off-NYSE quotations only rarely improve on the NYSE quote. A very small number of quotations in sample Nasdaq stocks originate at the Midwest stock exchange.

Table 1

Sample descriptive statistics.

The sample consists of 300 Nasdaq and 300 NYSE-listed stocks during the interval July 1 to December 31, 1998. Market capitalization is measured as of the beginning of the sample interval and is reported in millions of dollars. The share price is the traded-weighted sample average. Trades include those completed on any market, while quote updates include only those originating at the listing market. The return standard deviation pertains to daily returns computed on the basis of 4 p.m. quotation midpoints. Reported are means computed across stocks.

Firm size group	Listing market	Market capitalization	Share price	Trades per day	Quote updates per day	Return standard deviation
All	NYSE	4139.5	26.54	214.9	223.4	3.18
All	Nasdaq	4079.8	29.02	945.6	452.8	4.42
Large	NYSE	11,365.8	39.60	526.7	495.9	2.95
Large	Nasdaq	11,185.4	42.74	2451.0	1093.9	4.14
Medium	NYSE	837.2	21.71	93.8	131.0	3.64
Medium	Nasdaq	838.8	29.08	289.5	202.6	4.33
Small	NYSE	215.6	18.32	24.2	43.1	2.94
Small	Nasdaq	215.3	15.24	96.3	61.9	4.80

on the NYSE. Returns on sample Nasdaq stocks are more volatile: the cross-sectional average standard deviation of daily returns computed from 4 p.m. quotation midpoints is 4.1% for sample Nasdaq stocks, compared to 3.2% for sample NYSE stocks.

2. Research methods employed

2.1. Measures of trading costs

This study considers four measures of trade execution costs.⁵ The first is the quoted bid–ask half-spread, defined as half the difference between the inside ask quote and the inside bid quote. The average quoted half-spread for each sample stock is computed on a time-weighted basis. Reported are cross-sectional averages of the firm means.

The quoted half-spread does not accurately measure trading costs when trades are executed at prices away from the quotes. A measure of trading costs that incorporates actual execution prices is the effective bid–ask spread, measured for

⁵ Brokerage commission charges are not studied, for two reasons. First, comprehensive data on commission payments is not available. Second, individual traders are aware of the commissions they pay, since these are reported directly to them. Trade execution costs are not reported to traders, but must be inferred from trade prices.

trade t in stock i as

$$\text{Effective Half-spread}_{it} = D_{it}(P_{it} - M_{it}),$$

where D_{it} is an indicator variable that equals one for customer buy orders and negative one for customer sell orders, P_{it} is the price at which the trade is executed, and M_{it} is the midpoint of the reference bid and ask quotes, viewed as a proxy for the underlying value of the stock.

If traders possess private information about security values on average, market prices will tend to rise after customer buys and fall after sells. These price movements reflect what [Glosten \(1987\)](#) refers to as adverse selection costs. Some observers (e.g. [Seppi, 1997](#)) have argued that trading costs should be measured based on trades' *temporary* or *non-informational* price impact. The realized bid–ask half-spread is such a measure. It is defined as

$$\text{Realized Half-spread}_{it} = D_{it}(P_{it} - M_{it+n}),$$

where M_{it+n} is midpoint of the quotations in effect n periods after the trade, used as a proxy for the post-trade value of the stock. The post-trade movement in the quote midpoint reflects (on average) the market's assessment of the private information that the trade conveys, referred to as the trade's price impact:

$$\begin{aligned} \text{Price Impact}_{it} = D_{it}(M_{it+n} - M_{it}) = & \text{Effective Half-spread}_{it} \\ & - \text{Realized Half-spread}_{it}. \end{aligned}$$

I use the midpoint of the quotes in effect 30 min after the time of the reference quote, or the 4 p.m. quotations during the last half hour of trading, as M_{it+n} . Effective and realized half-spreads are measured for each trade and averaged across trades for each stock. Reported are simple cross-sectional averages of the firm-by-firm means.

Obtaining estimates of effective or realized spreads requires that trades be assigned as being buyer or seller-initiated, and that trades be matched with prevailing quotes. A measure of average trade execution costs that does not impose these requirements is the [Roll \(1984\)](#) spread, which exploits the movement of trade prices between the market's effective bid and ask prices, and infers the effective width of the bid–ask spread from the magnitude of the negative serial correlation induced in transaction price changes.

I use a modification of the Roll estimator suggested by [Schultz \(2000\)](#) to alleviate small sample bias. Letting P_{it} denote trade t in stock i and COV denote covariance, the estimator is

$$S_i = \frac{\sqrt{-\text{COV}(P_{it} - P_{it-1}, P_{it+1} - P_{it})}}{1 - 7/8(n - 1)},$$

where n is the number of trades in the sample. Due to limitations on the number of trades that can be handled by available computing systems, I implement the Roll measure for each firm on a weekly basis, using all trades completed during the week. The final estimate for each stock is obtained as the trade-weighted average of the

weekly estimates. Reported are cross-sectional averages of the firm-by-firm estimates.

2.2. Comparisons to common practitioner measures of trading costs

The trading cost measures used here have become fairly standard in the academic literature. The practitioner literature commonly uses a different set of benchmarks, comparing trade prices to quote midpoints or other benchmarks at the time of the trading decision (as in the Perold measure), to closing prices or quote midpoints on the day of the trade, or to the volume weighted average price (VWAP) on the day of the trade. The recommendation made here to select a reference quote from a time earlier than the trade is similar in intent to using the quote at the time of the trade decision as the reference point, and can be implemented without knowing the decision time. Measures of trading costs that use end of day prices or quotes as the reference point are similar to the realized spread measures described here, in that the reference price is observed after the trade and will have been affected by any information the trade conveyed to the market. Since the VWAP for the day is an average of prices for trades occurring both before and after the trade of interest, comparing the trade price to VWAP for the day is similar to computing a weighted average of numerous effective and realized spread measures, using individual trade prices as reference points. The use of both pre and post-trade prices clouds the interpretation of VWAP-based trading cost measures. However, it may be possible to implement trading cost measures based on VWAP in cases where intraday quote data is not available.

2.3. Algorithms for assigning trade direction

The most widely used technique for categorizing trades as buyer or seller-initiated is that recommended by [Lee and Ready \(1991\)](#), henceforth LR. Their algorithm assigns trades completed at prices above (below) the prevailing quote midpoint as customer buys (sells). Trades executed at the quote midpoint are assigned by the “tick test”, by which trades at a higher (lower) price as compared to the most recent trade at a different price are classified as buys (sells).

Several studies have recently emerged that use specialized datasets containing order information to assess the accuracy of the Lee and Ready algorithm. [Finucane \(2000\)](#), [Odders-White \(2000\)](#) and [Lee and Radhakrishna \(2000\)](#) all use the TORQ dataset, which provides trade, order, and quote data for a sample of 144 NYSE-listed stocks during a three-month period in 1990 and 1991. [Ellis et al. \(2000\)](#) use a proprietary sample that includes order data following initial public offerings in 313 Nasdaq stocks during a twelve-month period during 1996 and 1997. Peterson and Sirri use recent NYSE system order data. These papers indicate that, while the Lee and Ready algorithm works fairly well overall, classifying about 85% of trades correctly, alternative algorithms may perform better.

Ellis, Michaely, and O’Hara (henceforth EMO) propose assigning trades executed at the ask (bid) quote as customer buys (sells), while using the tick test for all other

trades. They show that their proposed method outperforms the LR method in their Nasdaq sample. Finucane (2000) does not directly evaluate the EMO method, but his results indicate that, while the tick test and the LR method have similar overall success rates, the LR algorithm works better than the tick test for trades executed at the quotes. The EMO method uses the tick test for all trades *except* those executed at the quotes, so we can infer from the Finucane evidence that the EMO method would have outperformed the tick test in the NYSE-based TORQ sample as well. Peterson and Sirri verify this result for more recent NYSE data.

2.4. Which quote midpoint should be used as the reference?

Researchers assessing trade execution costs generally compare trade prices to quotes in effect somewhat before the trade report time. The usual rationale is that trade reports are sometimes delayed, so that report times lag actual execution times.⁶ Lee and Ready (1991) recommended a 5-second allowance which most researchers have adopted so that trade times are reduced by five seconds before the trade and quote series are merged.⁷ Even if trades were reported without delay, comparing trade prices to earlier quotes might be appropriate to capture any systematic pre-trade price impacts.

2.5. Statistical significance

I assess statistical significance based on variability in the firm-by-firm estimates. For each trading cost measure, I compute the cross-sectional average trading cost by market, and conduct a simple *t*-test of the hypothesis that the average trading cost is the same on each market.⁸ I also examine each pair of matched stocks and assess the proportion of the sample firm-pairs where the observation for the NYSE stock exceeds that of its Nasdaq counterpart. If there is no systematic variation in trading costs across markets this percentage should not deviate significantly from fifty. I test this hypothesis using a simple binomial sign test, and report the resulting two-tailed *p*-value. I also assess whether results differ depending on the adjustment made to

⁶An additional complication is that the trade report times shown in the TAQ database have, during some periods, lagged actual trade report times for Nasdaq stocks. The problem apparently arose due to delays in data transmission from Nasdaq's Automated Confirmation Transaction System to the Securities Industry Automation Corporation (SIAC) system, which is the source of the TAQ data. Schultz (2000) provides a more complete discussion, and notes that the problem appeared to be most severe during 1995 and 1996. NASD officials have indicated in private correspondence that this problem has been solved as of early 1998, and that trade report times in the TAQ database have since been generally accurate.

⁷There is, however, some evidence that average delays in reporting trades may exceed five seconds. Hasbrouck et al. (1993) report a median delay of 14 seconds on the NYSE. EMO report a median delay between actual trade execution times and TAQ trade report times of 15 seconds for their sample of Nasdaq stocks. Blume and Goldstein match trades and quotes after adjusting downward reported time stamps by the median delay for the reporting market, ranging from 16 seconds on the NYSE to 34 seconds at the Boston Stock Exchange. Bessembinder (1999) adjusts reported time stamps by twenty seconds.

⁸Note that, although the sample includes over 43 million trades, the final test statistics are based on the relatively small sample size of 600, the number of firm-level means employed.

Table 2

Average quoted bid–ask half-spreads for a capitalization-matched sample of 300 NYSE and 300 Nasdaq stocks, during the July to December 1998 period.

Average spreads are computed for each firm on a time-weighted basis. Reported are cross-sectional averages of the firm means. The *t*-statistic is for the hypothesis that the mean difference is zero. The binomial *p*-value is for the hypothesis that the percentage of cases where the NYSE observation exceeds that of the matched Nasdaq observation equals fifty.

	NYSE mean	Nasdaq mean	Mean difference	<i>t</i> -Statistic	% where NYSE larger	Binomial <i>p</i> -value
<i>Panel A: In cents per share</i>						
All stocks	8.71	15.69	−6.98	−6.84	32.67	0.000
Large stocks	6.95	9.76	−2.81	−2.22	46.00	0.484
Medium stocks	8.41	17.73	−9.32	−4.38	21.00	0.000
Small stocks	10.77	19.59	−8.82	−5.07	31.00	0.000
<i>Panel B: As percent of share price</i>						
All stocks	0.486	0.739	−0.253	−8.62	32.33	0.000
Large stocks	0.212	0.238	−0.026	−1.72	49.00	0.920
Medium stocks	0.497	0.621	−0.124	−3.31	34.00	0.002
Small stocks	0.750	1.358	−0.608	−9.38	14.00	0.000

trade times by using simple “differences of means” tests that compare each mean obtained with a positive time adjustment to the corresponding mean obtained with a zero lag.

3. Empirical results

3.1. Quoted bid–ask half-spreads

Though the main focus of this paper is on methodological issues that arise when computing effective or realized spreads, it is useful to initially examine quoted half-spreads as a basis for comparison. Table 2 reports average quoted half-spreads; Panel A reports spreads in cents per share while Panel B reports spreads as a percentage of the bid–ask midpoint.

Consistent with results reported by Bessembinder (1999) and Huang and Stoll (1996) for earlier samples, quoted half-spreads are significantly wider on Nasdaq than the NYSE. The full-sample mean quoted half-spread on Nasdaq is 15.7 cents (0.74% of share price), compared to 8.7 cents (0.49% of share price) on the NYSE. There is, however, variation across firm size groups. For large-capitalization stocks differences in average quoted half-spreads across markets are minimal.⁹ The mean quoted half-spread for large stocks is 7.0 cents (0.21% of share price) on the NYSE versus 9.7 cents (0.24%) of share price on Nasdaq. The *t*-statistic indicates that these

⁹Finding that quoted spreads for large stocks are quite similar across Nasdaq and the NYSE is consistent with the results reported by Weston (2000), who examines only large-capitalization stocks.

differences in means are only marginally significant, and the binomial test does not reject the hypothesis that NYSE spreads exceed Nasdaq spreads for half of sample firm pairs.

In contrast, average quoted half-spreads for medium and small-capitalization stocks remain substantially larger on Nasdaq. The mean quoted half-spread for medium-size (small-size) stocks is 8.4 cents (10.8 cents) on the NYSE, compared to 17.7 cents (19.6 cents) on Nasdaq. The *t*- and binomial tests confirm the statistical significance of the difference in quoted spreads across markets for the smaller stocks.

3.2. Price improvement and allowances for trade reporting lags

Recognizing that trade reports can be delayed, researchers have generally compared trade prices to quotes reported slightly earlier than the trade report time. The adjustment is typically accomplished by reducing trade report times by a fixed number of seconds before comparing trades to contemporaneous quotes. This section assesses the effect of alternative allowances for trade lags. I first report the percentage of trades that are completed at prices inside (Table 3) and outside (Table 4) the matching quotes, when the matching quote is defined as that in effect from 0 to thirty seconds before the trade report time.

Comparing to earlier quotations monotonically decreases the percentage of trades that appear to be executed within the quotes. For the NYSE sample the percentage of trades apparently receiving price improvement decreases from 36.4% for contemporaneous quotes to 34.0% when five seconds are deducted from trade report times to 32.5% when trade times are adjusted by thirty seconds. For the Nasdaq sample the effect is more dramatic, as the rate of price improvement declines from 28.1% with no lag, to 17.2% at a five-second lag, and to 14.5% at a thirty-second lag.

In contrast, comparing trade prices to earlier quotes monotonically *increases* the percentage of trades apparently receiving price disimprovement.¹⁰ For the NYSE sample, the percentage of trades at prices greater than the ask or lower than the bid is 0.6% when comparing trade prices to quotes at the trade report time. The apparent rate of price disimprovement rises to 1.9% if trade times are reduced by five seconds, and to 5.7% at a thirty-second lag. A similar effect is observed for Nasdaq firms, as the percentage of trades apparently completed outside the quotes rises from 5.0% with no trade reporting lag to 6.0% at a five-second lag and to 14.7% with a thirty-second lag. These results are generally quite robust across the large, medium, and small capitalization sub-samples. The changes in reported rates of price improvement and disimprovement as the trade lag allowance is increased are statistically significant in all cases.

¹⁰NYSE trades may be executed outside the quotations if they are larger than the quote size, or if multiple orders arrive simultaneously (see Bacidore et al., 2001). NASD regulations do not specifically prohibit trading at prices inferior to the best quotes (see Smith et al., 1998), but do require (NASD rule 2320) market makers to “buy or sell ... so that the resultant price to the customer is as favorable as possible under prevailing market conditions”.

Table 3

Percentage of trades receiving price improvement.

The percentage of trades at each firm that are executed at prices less than the matched ask quote and greater than the matched bid quote is computed. Matching quotes are assessed by comparing trade prices to quotes in effect from 0 to 30 seconds earlier than the trade report time. Trade lag is the number of seconds deducted from the trade report time before comparing to quotes. Reported are simple cross-sectional averages across stocks. The t -statistic is for the hypothesis that the mean difference is zero. The binomial p -value is for the hypothesis that the percentage of cases where the NYSE observation exceeds that of the matched Nasdaq observation equals fifty. An asterisk indicates that the mean differs significantly (t -statistic > 1.96) from the corresponding mean obtained with a zero lag.

Firm size	Trade lag	NYSE mean	Nasdaq mean	Mean difference	t -Statistic	% NYSE larger	Binomial p -value
All	0	36.44	28.14	8.30	12.03	73.00	0.000
All	5	34.04*	17.24*	16.80	33.76	98.00	0.000
All	10	33.31*	14.72*	18.59	40.15	99.33	0.000
All	15	32.99*	14.26*	18.73	41.08	99.67	0.000
All	20	32.78*	14.26*	18.52	40.58	99.67	0.000
All	25	32.66*	14.38*	18.28	39.87	99.67	0.000
All	30	32.56*	14.48*	18.08	39.29	99.67	0.000
Large	0	35.31	20.92	14.39	12.29	88.00	0.000
Large	5	32.83*	13.11*	19.73	23.17	99.00	0.000
Large	10	31.86*	11.90*	19.96	25.22	99.00	0.000
Large	15	31.31*	12.06*	19.24	24.67	99.00	0.000
Large	20	30.97*	12.27*	18.71	24.09	99.00	0.000
Large	25	30.73*	12.43*	18.30	23.56	99.00	0.000
Large	30	30.52*	12.54*	17.98	23.13	99.00	0.000
Medium	0	36.19	31.29	4.89	4.24	62.00	0.012
Medium	5	34.01*	19.09*	14.92	17.38	96.00	0.000
Medium	10	33.45*	16.42*	17.03	20.62	100.00	0.000
Medium	15	33.17*	16.05*	17.11	20.73	100.00	0.000
Medium	20	33.01*	16.17*	16.83	20.33	100.00	0.000
Medium	25	32.93*	16.40*	16.53	19.76	100.00	0.000
Medium	30	32.89*	16.57*	16.31	19.40	100.00	0.000
Small	0	37.83	32.22	5.61	5.49	69.00	0.000
Small	5	35.28*	19.51*	15.77	19.57	99.00	0.000
Small	10	34.64*	15.85*	18.79	24.45	99.00	0.000
Small	15	34.49*	14.66*	19.84	26.69	100.00	0.000
Small	20	34.37*	14.34*	20.03	27.08	100.00	0.000
Small	25	34.33*	14.32*	20.01	27.23	100.00	0.000
Small	30	34.27*	14.33*	19.94	27.14	100.00	0.000

Some indication of the optimal amount by which to adjust trade times for purposes of assessing trade direction can be obtained by assessing the adjustment that leads to the most accurate measures of the proportion of trades executed inside and outside the quotations. This cannot be ascertained using the TAQ data, since the true trade times are not known. However, the percentages reported in Tables 3 and 4 can be compared to data reported for specialized samples where actual trade times are available.

Table 4

Percentage of trades executed outside the quotations.

The percentage of trades at each firm that are executed at prices greater than the contemporaneous ask quote or less than the contemporaneous bid quote is computed. Trade lag is the number of seconds deducted from the trade report time before comparing to quotes. Reported are simple cross-sectional averages across stocks. The *t*-statistic is for the hypothesis that the mean difference is zero. The binomial *p*-value is for the hypothesis that the percentage of cases where the NYSE observation exceeds that of the matched Nasdaq observation equals fifty. An asterisk indicates that the mean differs significantly (*t*-statistic > 1.96) from the corresponding mean obtained with a zero lag.

Firm size	Trade lag	NYSE mean	Nasdaq mean	Mean difference	<i>t</i> -Statistic	% NYSE larger	Binomial <i>p</i> -value
All	0	0.59	4.96	-4.38	-23.49	1.00	0.000
All	5	1.85*	5.97	-4.12	-17.15	11.00	0.000
All	10	3.01*	7.48*	-4.47	-13.84	19.00	0.000
All	15	3.85*	9.28*	-5.43	-13.63	19.00	0.000
All	20	4.52*	11.25*	-6.73	-14.54	16.67	0.000
All	25	5.10*	13.15*	-8.05	-15.56	13.33	0.000
All	30	5.65*	14.74*	-9.09	-16.30	12.33	0.000
Large	0	0.81	6.96	-6.15	-20.78	0.00	0.000
Large	5	1.68*	8.47	-6.78	-16.69	1.00	0.000
Large	10	2.77*	10.86*	-8.08	-14.19	4.00	0.000
Large	15	3.74*	13.59*	-9.85	-13.73	5.00	0.000
Large	20	4.65*	16.50*	-11.85	-14.11	3.00	0.000
Large	25	5.53*	19.23*	-13.71	-14.50	3.00	0.000
Large	30	6.38*	21.48*	-15.10	-14.73	3.00	0.000
Medium	0	0.52	4.60	-4.08	-14.39	0.00	0.000
Medium	5	1.76*	5.50	-3.73	-10.60	7.00	0.000
Medium	10	2.87*	6.78*	-3.91	-8.40	14.00	0.000
Medium	15	3.60*	8.38*	-4.78	-8.31	13.00	0.000
Medium	20	4.18*	10.14*	-5.97	-8.92	10.00	0.000
Medium	25	4.65*	11.84*	-7.19	-9.63	10.00	0.000
Medium	30	5.07*	13.26*	-8.19	-10.16	9.00	0.000
Small	0	0.43	3.32	-2.89	-9.57	3.00	0.000
Small	5	2.10*	3.94	-1.84	-5.60	25.00	0.000
Small	10	3.40*	4.80*	-1.40	-3.38	39.00	0.035
Small	15	4.21*	5.88*	-1.67	-3.34	39.00	0.035
Small	20	4.73*	7.11*	-2.37	-4.16	37.00	0.012
Small	25	5.13*	8.38*	-3.25	-5.08	27.00	0.000
Small	30	5.50*	9.49*	-3.99	-5.74	25.00	0.000

EMO report that 25.2% of the trades in their proprietary Nasdaq sample are completed at prices within the quotes, while 4.3% of trades are executed outside the quotes. Their results, based on actual trade times, correspond most closely to the Nasdaq results obtained here when trades are compared to quotes in effect at trade report times, without any adjustment. They also compare actual trade times for their sample to trade times reported in the TAQ database, observing a mean (median) delay of eleven (fifteen) seconds. They indicate that incorporating a fifteen second delay does not improve the accuracy of the LR algorithm, and recommend that

Nasdaq trades be assigned as buyer or seller initiated without making any allowance for reporting lags. The results reported here support the same conclusion.

Bacidore et al. (2001) use proprietary NYSE data to report that 0.7% of NYSE trades resulting from system market orders are executed at prices outside the quotes in effect at the trade time, while 41.5% of trades are executed at prices within the quotes. Again, the closest correspondence between results obtained here using publicly available data and results obtained when using the proprietary database arises when *no* allowance for trade reporting lags is made.

These results obtained here suggest that trades are best compared to contemporaneous quotations when measuring price improvement rates, and by implication, when assigning trades as buyer or seller initiated. Peterson and Sirri (2002) provide direct evidence on this issue, using proprietary NYSE order data drawn from 1997. Their results confirm that the success rates of the LR and EMO algorithms are greatest when trade prices are compared to contemporaneous rather than earlier quotes.

3.3. Measures of trading costs with varying trade time adjustments

Most studies of trading costs have compared trade prices to earlier quotes *both* when signing trades and when computing effective and realized bid–ask spreads. It is of interest to know whether measured trading costs are sensitive to the time adjustment used. To assess this issue I next report average effective and realized bid–ask spreads for the present sample, when the time of the reference quote precedes the trade report time between 0 and 30 seconds. The results reported rely on the EMO algorithm to sign trades. In light of the results reported above and by Peterson and Sirri, the EMO method is implemented without any adjustment to the trade time stamps. Similar results are obtained when using the LR method.

Table 5 displays average effective bid–ask half-spreads obtained when comparing trade prices to the quote midpoint in effect from 0 to thirty seconds before the trade report time. Measured effective spreads increase monotonically with the adjustment to trade times. For the sample of NYSE stocks, the measured average effective half-spread increases from 4.9 cents per share with no lag to 5.4 cents per share when trade report times are adjusted by thirty seconds. For the sample of Nasdaq stocks the measured effective half-spread increases from 9.4 cents with no lag to 10.8 cents with a thirty-second lag. *t*-Tests indicate that the increase in measured trading costs is statistically significant for time adjustments of 20 seconds or greater for NYSE stocks and for adjustments of 25 seconds or greater for Nasdaq stocks. Similar results are observed for each market capitalization subsample.

Although measured effective spreads on both markets vary with the trade lag, inference regarding cross-market comparisons do not. Regardless of the trade lag employed, the average effective half-spread for Nasdaq stocks exceeds that for NYSE stocks. The cross-market differences in average effective half-spreads are highly significant, with *t*-statistics for the hypothesis that average spreads are equal exceeding 6.0 for each subsample and exceeding 11.0 for the full sample regardless of the trade lag.

Table 5

Average effective bid–ask half-spreads.

The effective bid–ask spread is computed for every trade as $D_{it}(P_{it} - M_{it})$, where P_{it} is the trade price at time t for firm i , M_{it} is the matching quote midpoint, and D_{it} is a trade direction indicator that equals one (negative one) for buyer (seller)-initiated trades, assigned by the EMO algorithm while using a zero lag. Average effective spreads for each firm are obtained by weighting each trade equally. Trade lag is the number of seconds deducted from the trade report time before comparing to quotes. Reported are simple cross-sectional averages across stocks. The t -statistic is for the hypothesis that the mean difference is zero. The binomial p -value is for the hypothesis that the percentage of cases where the NYSE observation exceeds that of the matched Nasdaq observation equals fifty. An asterisk indicates that the mean differs significantly (t -statistic > 1.96) from the corresponding mean obtained with a zero lag.

Firm size	Trade lag	NYSE mean	Nasdaq mean	Mean difference	t -Statistic	% NYSE larger	Binomial p -value
All	0	4.93	9.36	-4.43	-11.21	16.00	0.000
All	5	5.08	9.86	-4.78	-11.56	17.00	0.000
All	10	5.17	10.09	-4.92	-11.69	16.67	0.000
All	15	5.24	10.29	-5.05	-11.84	15.67	0.000
All	20	5.29*	10.47	-5.18	-12.08	14.67	0.000
All	25	5.33*	10.65*	-5.32	-12.35	14.00	0.000
All	30	5.37*	10.81*	-5.44	-12.50	14.00	0.000
Large	0	4.05	6.71	-2.66	-6.00	17.00	0.000
Large	5	4.18	7.07	-2.89	-6.18	19.00	0.000
Large	10	4.26	7.28	-3.02	-6.21	19.00	0.000
Large	15	4.32*	7.49	-3.16	-6.38	16.00	0.000
Large	20	4.39*	7.71	-3.32	-6.64	14.00	0.000
Large	25	4.45*	7.92	-3.47	-6.88	13.00	0.000
Large	30	4.51*	8.11*	-3.60	-7.07	13.00	0.000
Medium	0	4.68	9.96	-5.28	-7.46	10.00	0.000
Medium	5	4.81	10.50	-5.69	-7.78	10.00	0.000
Medium	10	4.89	10.76	-5.88	-7.83	11.00	0.000
Medium	15	4.94	10.98	-6.04	-7.84	11.00	0.000
Medium	20	4.99	11.16	-6.18	-7.98	11.00	0.000
Medium	25	5.02*	11.33	-6.31	-8.14	11.00	0.000
Medium	30	5.06*	11.48	-6.42	-8.30	11.00	0.000
Small	0	6.07	11.43	-5.36	-6.54	20.00	0.000
Small	5	6.26	12.02	-5.76	-6.68	22.00	0.000
Small	10	6.37	12.24	-5.87	-6.77	20.00	0.000
Small	15	6.44	12.40	-5.96	-6.84	20.00	0.000
Small	20	6.48	12.53	-6.05	-6.91	19.00	0.000
Small	25	6.51	12.69	-6.18	-7.02	18.00	0.000
Small	30	6.55	12.84	-6.29	-7.02	18.00	0.000

Table 6 reports on average realized half-spreads measured with various trade lag allowances. In contrast to the effect on effective spreads, measures of average realized spreads are not sensitive to the adjustment for trade-reporting lags. For the NYSE sample, the mean realized half spread is generally quite small, but varies only from 0.23 to 0.24 cents as the trade lag adjustment changes. For Nasdaq firms the average realized half-spread declines very slightly, from 4.78 cents without any adjustment to trade report times to 4.72 cents with a 30-second adjustment. Similar

Table 6

Average realized bid–ask half-spreads, in cents per share.

The realized half-spread is computed for every trade as $D_{it}(P_{it} - M_{it+30})$, where P_{it} is the trade price at time t for firm i , M_{it+30} is the midpoint of the first quotes posted at least 30 minutes after the trade, and D_{it} is a trade direction indicator that equals one (negative one) for buyer (seller)-initiated trades, assigned by the EMO algorithm while using a zero lag. Average effective spreads for each firm are obtained by weighting each trade equally. Reported are simple cross-sectional averages across stocks. The t -statistic is for the hypothesis that the mean difference is zero. The binomial p -value is for the hypothesis that the percentage of cases where the NYSE observation exceeds that of the matched Nasdaq observation equals fifty. An asterisk indicates that the mean differs significantly (t -statistic > 1.96) from the corresponding mean obtained with a zero lag.

Firm size	Trade lag	NYSE mean	Nasdaq mean	Mean difference	t -Statistic	% NYSE larger	Binomial p -value
All	0	0.23	4.78	-4.55	-13.76	10.00	0.000
All	5	0.24	4.78	-4.54	-13.63	10.00	0.000
All	10	0.24	4.76	-4.52	-13.46	10.00	0.000
All	15	0.23	4.75	-4.51	-13.42	10.00	0.000
All	20	0.23	4.73	-4.50	-13.43	10.00	0.000
All	25	0.23	4.73	-4.50	-13.43	10.00	0.000
All	30	0.23	4.72	-4.49	-13.34	10.00	0.000
Large	0	0.63	2.86	-2.23	-6.42	19.00	0.000
Large	5	0.63	2.85	-2.23	-6.47	20.00	0.000
Large	10	0.62	2.84	-2.22	-6.51	20.00	0.000
Large	15	0.62	2.84	-2.22	-6.51	20.00	0.000
Large	20	0.61	2.83	-2.22	-6.50	20.00	0.000
Large	25	0.61	2.82	-2.21	-6.48	20.00	0.000
Large	30	0.61	2.81	-2.20	-6.49	20.00	0.000
Medium	0	-0.02	4.87	-4.89	-9.27	7.00	0.000
Medium	5	-0.02	4.88	-4.91	-9.21	5.00	0.000
Medium	10	-0.03	4.87	-4.90	-9.25	5.00	0.000
Medium	15	-0.03	4.86	-4.88	-9.30	5.00	0.000
Medium	20	-0.03	4.83	-4.85	-9.44	5.00	0.000
Medium	25	-0.03	4.83	-4.86	-9.39	5.00	0.000
Medium	30	-0.03	4.83	-4.86	-9.35	5.00	0.000
Small	0	0.09	6.62	-6.53	-9.22	5.00	0.000
Small	5	0.11	6.59	-6.48	-9.08	5.00	0.000
Small	10	0.11	6.56	-6.45	-8.83	5.00	0.000
Small	15	0.11	6.54	-6.44	-8.76	5.00	0.000
Small	20	0.11	6.54	-6.43	-8.71	5.00	0.000
Small	25	0.11	6.53	-6.42	-8.74	5.00	0.000
Small	30	0.11	6.52	-6.41	-8.65	5.00	0.000

results are observed for each market capitalization subsample. Inference regarding cross-market comparisons is again consistent with earlier studies, as average realized half-spreads for Nasdaq stocks significantly exceed those for the NYSE stocks, regardless of the lag employed.

Some, e.g. [Heidle and Huang \(2002\)](#), assert that the decentralized structure of the Nasdaq dealer market leaves Nasdaq dealers more vulnerable to losses incurred in trades with better informed agents. These larger adverse selection costs would have

to be recovered through wider spreads. The results reported here indicate that inference regarding the magnitude of adverse selection costs *is* sensitive to the researcher's method of allowing for trade reporting lags.

Table 7 displays measures of trades' average price impact when the time of the benchmark quote is varied. The average price impact for the full sample of NYSE trades is 4.7 cents when quotes 30 minutes after the trade report time are compared to quotes at the trade report time. Measured price impact for NYSE trades increases to 5.1 cents when quotes 30 minutes later are compared to quotes in effect 30 seconds before the trade report time. For the Nasdaq sample, measured price impact increases more dramatically as earlier quotes are used as the reference point. The signed movement in the quote midpoint from the trade report time until 30 minutes later is 4.6 cents, while the change in the quote midpoint from 30 seconds before the trade report time to 30 minutes later (29.5 minutes after the trade) is 6.1 cents. The increase in measured price impact with a longer time adjustment is statistically significant for Nasdaq stocks, but is not significant (the *t*-statistic at a 30 second lag is 1.51) for NYSE-listed stocks.

These results indicate that quote midpoints move systematically away from trades (rising on buy orders and dropping on sell orders) in the seconds before trades are reported. On the NYSE the average movement in quote midpoints is 0.44 cents during the 30 seconds prior to the trade report. On Nasdaq, quote midpoints move away from the trade by an average 1.51 cents in the 30 seconds before the trade report. Greater adjustments to trade report times result in this *pre-trade price impact* being included in measures of effective trading costs.

Pre-trade price impacts could potentially reflect trade reporting lags, i.e., that the quote is updated during the lag between the actual trade time and the trade report time. However, EMO report that quotes are rarely updated during this interval in their Nasdaq sample, and Peterson and Sirri report that TAQ trade report times lag actual trade execution times by only two seconds, on average. The other, and more likely, explanation is that quotes are systematically and adversely updated in the seconds before trades are executed. Peterson and Sirri (2002) and Werner (2002) both analyze proprietary NYSE system order data and provide evidence consistent with this interpretation. The former study examines market orders, while the later also considers marketable limit orders and floor broker orders. Each reports that prices move significantly in the direction of the trades before execution. These adverse average movements in quotes ahead of trades could occur because information about pending orders leaks to the market ahead of trades. It could also reflect larger orders being broken up by brokerage firms into smaller orders sent to market in succession, or it may reflect several traders reacting similarly to common information events, with quotes being revised after the earliest orders are executed.

Regardless of the explanation for quote movements ahead of trade report times, these results show that inference as to whether trades' price impacts differ systematically across Nasdaq and the NYSE *is* sensitive to the researcher's adjustment to trade report times. When no adjustment is made, trades' price impacts differ by only 0.1 cents per share on average, which is not statistically significant (*t*-statistic = 0.44). When an adjustment of twenty seconds or more is

Table 7

Trades' average price impact, in cents per share.

Price impact is computed for every trade as $D_{it}(M_{it+30} - M_{it})$, where M_t is the matching quote midpoint at the time of firm i trade t , M_{it+30} is the midpoint of the first quotes posted at least 30 minutes after the trade, and D_{it} is a trade direction indicator that equals one (negative one) for buyer (seller)-initiated trades, assigned by the EMO algorithm. Average price impact for each firm is obtained by weighting each trade equally. Trade lag is the number of seconds deducted from the trade report time to identify the matching quote. Reported are simple cross-sectional averages across stocks. The t -statistic is for the hypothesis that the mean difference is zero. The binomial p -value is for the hypothesis that the percentage of cases where the NYSE observation exceeds that of the matched Nasdaq observation equals fifty. An asterisk indicates that the mean differs significantly (t -statistic > 1.96) from the corresponding mean obtained with a zero lag.

Firm size	Trade lag	NYSE mean	Nasdaq mean	Mean difference	t -Statistic	% NYSE larger	Binomial p -value
All	0	4.70	4.58	0.12	0.44	54.00	0.149
All	5	4.85	5.09	-0.24	-0.81	49.00	0.773
All	10	4.94	5.34*	-0.40	-1.27	47.33	0.387
All	15	5.00	5.54*	-0.54	-1.67	46.67	0.273
All	20	5.05	5.74*	-0.68	-2.08	45.33	0.119
All	25	5.10	5.92*	-0.82	-2.48	44.00	0.043
All	30	5.14	6.09*	-0.95	-2.80	43.33	0.024
Large	0	3.42	3.84	-0.43	-1.60	37.00	0.012
Large	5	3.55	4.22	-0.67	-2.29	36.00	0.007
Large	10	3.64	4.43	-0.80	-2.56	34.00	0.002
Large	15	3.71	4.65*	-0.94	-2.90	32.00	0.000
Large	20	3.77	4.88*	-1.10	-3.30	32.00	0.000
Large	25	3.84	5.10*	-1.26	-3.66	30.00	0.000
Large	30	3.90	5.30*	-1.40	-3.94	29.00	0.000
Medium	0	4.70	5.07	-0.37	-0.94	56.00	0.193
Medium	5	4.84	5.62	-0.78	-1.90	48.00	0.764
Medium	10	4.91	5.89	-0.98	-2.25	47.00	0.617
Medium	15	4.97	6.12	-1.15	-2.53	47.00	0.617
Medium	20	5.01	6.33*	-1.32	-2.78	45.00	0.368
Medium	25	5.05	6.50*	-1.45	-3.02	43.00	0.193
Medium	30	5.09	6.65*	-1.56	-3.23	43.00	0.193
Small	0	5.98	4.82	1.16	1.68	68.00	0.000
Small	5	6.15	5.43	0.72	0.97	63.00	0.007
Small	10	6.26	5.68	0.58	0.76	61.00	0.021
Small	15	6.33	5.85	0.48	0.62	61.00	0.021
Small	20	6.37	5.99	0.38	0.48	59.00	0.057
Small	25	6.41	6.16	0.25	0.31	59.00	0.057
Small	30	6.45	6.33	0.12	0.15	58.00	0.089

allowed, the average measured price impact on Nasdaq significantly exceeds that on the NYSE.

3.4. Measures of trading costs obtained using the LR and EMO algorithms

Tables 8 and 9 report average effective and realized half-spreads, respectively, for each market, when trade direction is assigned by the LR and EMO procedures.

Table 8

Comparison of mean effective bid–ask half-spreads when trade direction is assigned by the Ellis, Michaely, and O’Hara (EMO) and Lee and Ready (LR) methods. The effective bid–ask half-spread is computed for each trade as $D_{it}(P_{it} - M_{it})$, where P_{it} is the trade price at time t for firm i , M_{it} is the quote midpoint at the trade report time, and D_{it} is a trade direction indicator that equals one (negative one) for buyer (seller)-initiated trades, assigned by either the EMO or the LR algorithm. Average effective spreads for each firm are obtained by weighting each trade equally. Reported are simple means computed across firms. The t -statistic is for the hypothesis that the mean difference is zero.

Firm size	Mean using EMO method	Mean using LR method	Mean difference	t -Statistic
<i>Panel A: NYSE stocks</i>				
All	4.93	5.72	0.79	21.68
Large	4.05	4.64	0.59	16.92
Medium	4.68	5.50	0.81	17.78
Small	6.07	7.02	0.95	10.68
<i>Panel B: Nasdaq stocks</i>				
All	9.36	11.27	1.91	9.79
Large	6.71	7.74	1.04	5.39
Medium	9.95	12.12	2.18	6.08
Small	11.43	13.95	2.52	6.15

These results are based on comparisons of trade prices to quotes in effect at the trade report time, i.e. with no trade time adjustment. Comparisons with trade lag allowances from five to thirty seconds are similar.

Estimates of average effective spreads obtained when using the EMO procedure are, by construction, less than or equal to those obtained from the LR method.¹¹ The effective half-spread estimated when using the EMO procedure will be the same as that obtained using the LR procedure, except when the EMO procedure identifies a trade at a price below (above) the midpoint as a buy (sell). In this case the effective spread measured by the EMO method is negative and the opposite of the effective spread measured by the LR method. The only question is how much smaller the EMO estimate of effective spreads will be.

For NYSE stocks, the EMO estimate of effective half-spreads is 0.8 cents less than the LR estimate, on average. For Nasdaq stocks the differential is greater, with LR estimates exceeding EMO-based estimates by 1.9 cents per share. A simple F -test rejects the hypothesis that differences between LR and EMO-based estimates are uniform across firm size groups on both markets. However, point estimates indicate a greater size-based sensitivity on Nasdaq. The differential in cost estimates for Nasdaq (NYSE) stocks increases from 1.0 (0.6) cents for large stocks to 2.5 (0.9) cents for small stocks. Note, however, that inferences regarding cross-market

¹¹ As a consequence, statistical significance should be assessed using a one-tailed test. The estimated t -statistics for differences in means are, however, so large that statistical significance is not an issue.

Table 9

Comparison of mean realized bid–ask half-spreads when trade direction is assigned by the Ellis, Michaely, and O’Hara (EMO) and Lee and Ready (LR) methods. The realized bid–ask half-spread is computed for every trade for every trade as $D_{it}(P_{it} - M_{it+30})$, where P_{it} is the trade price at time t for firm i , M_{it+30} is the midpoint of the first quotes posted at least 30 minutes after the trade, and D_{it} is a trade direction indicator that equals one (negative one) for buyer (seller)-initiated trades, assigned by either the EMO or the LR algorithm. Average effective spreads for each firm are obtained by weighting each trade equally. Reported are simple means computed across firms. The t -statistic is for the hypothesis that the mean difference is zero.

Firm size	Mean using EMO method	Mean using LR method	Mean difference	t -Statistic
<i>Panel A: NYSE stocks</i>				
All	0.23	1.27	1.04	20.42
Large	0.63	1.44	0.82	11.94
Medium	−0.02	1.14	1.16	16.64
Small	0.09	1.23	1.15	9.96
<i>Panel B: Nasdaq stocks</i>				
All	4.78	6.41	1.63	9.86
Large	2.86	3.81	0.94	5.15
Medium	4.87	6.78	1.91	6.26
Small	6.61	8.65	2.04	6.05

comparisons are again largely unaffected: effective half-spread estimates are substantially smaller on the NYSE under either method, for each subsample.

Unlike measured effective spreads, measured realized spreads need not always be smaller when using the EMO procedure than when using the LR procedure. In light of the evidence provided by EMO, Finucane, and Peterson and Sirri that the EMO method is more accurate in signing trades we anticipate more accurate measures of realized spreads when using the EMO procedure. The practical effect is that estimates of average realized half-spreads are 1.0 cents narrower for NYSE stocks and 1.6 cents narrower for Nasdaq stocks when the EMO method is used. While this difference is non-trivial, it does not reverse the results of cross-market comparisons: realized bid–ask spreads on Nasdaq remain substantially greater than on the NYSE, when using either algorithm to assign trades.

3.5. Comparisons of effective and realized spreads to roll-implied spreads

The Roll (1984) technique provides estimates of trade execution costs that do not require that trades be signed or that trades be matched to quotes, so errors from these sources are avoided.¹² Here, I assess the extent to which trading cost estimates obtained using the Roll procedure conform with those obtained from the effective and realized bid–ask spread methods.

¹² However, the Roll method relies on a set of restrictive assumptions. Among these are the assumptions that the spread width is constant over time, and that trades do not convey private information about value, i.e. that trades’ average price impact is zero.

Table 10

Average bid–ask half-spreads estimated by the Roll (1984) technique, in cents per share.

The Roll estimator is applied to trade data for each stock during the July 1 to December 31, 1998 period. Reported are cross-sectional averages of the firm estimates. The *t*-statistic is for the hypothesis that the mean difference is zero. The binomial *p*-value is for the hypothesis that the percentage of cases where the NYSE observation exceeds that of the matched Nasdaq observation equals fifty.

Trade size	Firm size	NYSE mean	Nasdaq mean	Mean difference	<i>t</i> -Statistic	% where NYSE larger	Binomial <i>p</i> -value
All	All	3.38	9.88	−6.49	−14.30	5.00	0.000
Large	All	9.15	18.06	−8.91	−7.24	13.67	0.000
Medium	All	4.05	11.83	−7.81	−10.24	9.33	0.000
Small	All	3.41	9.70	−6.29	−13.48	6.00	0.000
All	Large	3.37	8.17	−4.80	−7.74	8.00	0.000
All	Medium	3.05	11.14	−8.10	−8.10	3.00	0.000
All	Small	3.74	10.32	−6.58	−10.04	4.00	0.000

Because the results are of some interest, I report the average Roll half-spread for trades of varying sizes, using the procedure developed by Schultz (2000). Small trades are defined here as those of 1000 shares or less, medium trades are those from 1001 to 9999 shares, and large trades are those of 10,000 or more shares. Table 10 reports on average bid–ask half-spreads estimated by the Roll method. Also, to facilitate comparison across measures, Fig. 1 displays average quoted, effective, realized, and Roll-implied half-spreads.

The results indicate that average Roll-implied half-spreads are uniformly and significantly greater for the Nasdaq stocks than for NYSE stocks. The full sample mean for NYSE stocks is 3.38 cents, compared to 9.88 cents for Nasdaq stocks. Similar differentials are seen for each market-capitalization subsample.

Schultz (2000) reports that Roll-implied spread estimates correspond closely to estimates of effective half-spreads for his sample of Nasdaq stocks. This close correspondence is somewhat surprising. The Roll technique essentially measures price reversals around trades. Glosten (1987) points out that the negative serial covariance of returns arises *only* from the portion of the spread due to sources other than trades' permanent information content. The realized bid–ask spread also measures trading costs net of the information component. We might expect the Roll estimate to correspond more closely to estimates of realized bid–ask spreads.¹³

Comparing spread measures across Tables 8–10, or in Fig. 1, we see that average Roll-implied half-spreads in the present sample are quite similar to average effective half-spreads for Nasdaq stocks, consistent with the findings of Schultz (2000). For the full Nasdaq sample the average effective half-spread is 9.4 cents, while the

¹³ Schultz investigates this issue in more detail for his sample of Nasdaq stocks, and reports that an upward bias in the Roll estimate attributable to time variation in spreads (which were assumed constant in the Roll model) offsets the effect of trades' information content. As a consequence the Roll estimates are similar to measured effective spreads.

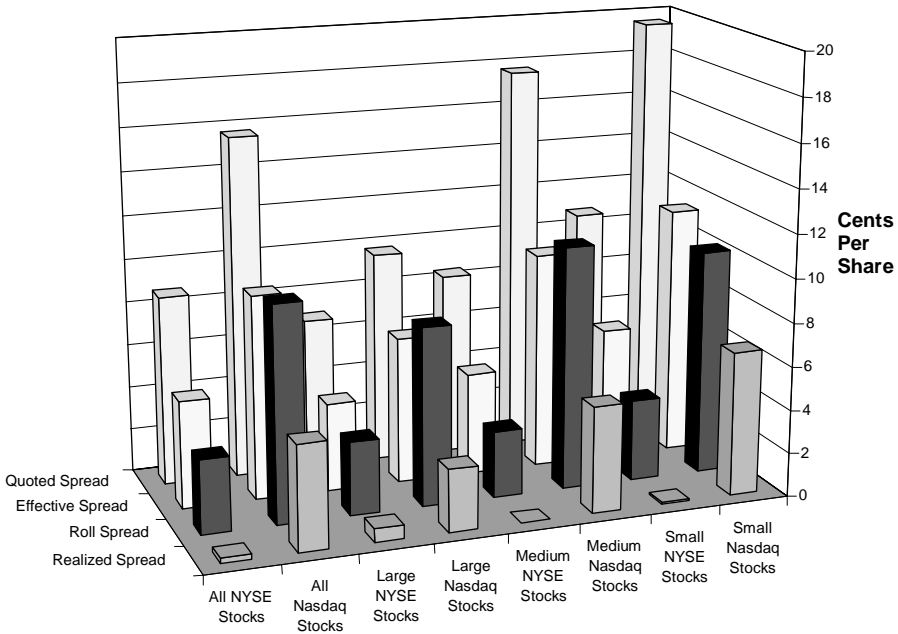


Fig. 1. Comparisons of bid-ask half-spread measures: quoted spreads, effective spreads, realized spreads, and roll-implied spreads.

average Roll-implied half-spread is 9.9 cents. Results for Nasdaq market capitalization subsamples are similar.

The close correspondence between Roll spreads and effective spreads on the Nasdaq market might be viewed as suggesting that trades need not be signed or matched with quotes at all, as the Roll estimate could simply be used as a general procedure. However, estimated Roll spreads for NYSE stocks do not match effective spread estimates as closely, particularly for small stocks. For small-capitalization NYSE issues the average estimated Roll half-spread is 3.7 cents, compared to an average effective half-spread of 6.1 cents. For the full sample of NYSE stocks the estimated Roll half-spread is 3.4 cents, which is closer to but still below the effective half-spread estimate of 4.9 cents.

Estimated Roll half-spreads are reasonably uniform across firm size groups, but not across trade sizes. On the NYSE, Roll half-spreads are 9.2 cents for large trades, 4.1 cents for medium trades, and 3.4 cents for small trades, indicating greater price reversals after large trades. A similar pattern is evident for Nasdaq stocks, where average Roll half-spreads are 18.1 cents for large trades, 11.8 cents for medium trades, and 9.7 cents for small trades.

The [Schultz \(2000\)](#) result that Roll implied spreads provide good alternative estimates of effective spreads for Nasdaq stocks does not generalize well to trades of varying sizes or to NYSE stocks. One possible explanation for the NYSE result is the presence of price continuity rules that increase the serial dependence of price

changes, thereby reducing Roll spread estimates. Regardless of the explanation, the implication is that difficulties in assessing trade execution costs resulting from the need to sign trades and to match trades with prevailing quotes cannot be readily sidestepped in studies involving NYSE stocks by simply using the Roll procedure instead.

4. Conclusions

This study conducts sensitivity analyses to assess the practical importance of some methodological issues that arise when attempting to measure trade execution costs using the publicly-available quotation and trade price databases. The publicly available databases contain trade report times, but not order submission or trade execution times, and do not indicate whether trades are buyer or seller-initiated. The results of the study indicate that a method for inferring trade direction suggested by Easley et al. (2000) leads to smaller estimates of trading costs as compared to the Lee and Ready (1991) method. Results reported by Easley et al. (2000), Finucane (2000), and Peterson and Sirri (2002) indicate that the smaller trading cost estimates obtained using the EMO method are more accurate measures of actual trading costs.

Adjusting trade report times as an allowance for possible reporting lags decreases the percentage of trades that appear to be executed within the quotes while increasing the percentage of trades that appear to be executed outside the quotes. Comparing results here with those obtained using proprietary databases that include accurate trade times and order data, it appears that trade direction and rates of price improvement are best assessed when making *no* adjustment for trade report lags.

Estimated trading costs increase if trade prices are compared to earlier rather than contemporaneous quotes, reflecting adverse quote movements prior to trade report times. If these adverse quote movements occur after order submission but before trade execution then they comprise a cost to traders that will not be captured if trade prices are compared to quotes in effect at trade report times.

On balance, the results obtained here support recommendations to (1) use the EMO technique in preference to the LR method to sign trades, (2) implement the EMO technique on the basis of contemporaneous rather than earlier quotations, and (3) use quotation midpoints in effect somewhat prior to the trade report time as the benchmark quote when measuring effective bid–ask spreads. This last recommendation is similar in spirit to the use of the quote midpoint at the time of the trading decision as the reference point (as in Perold, 1988), in order to include costs stemming from *pre trade price impact*.

Appendix A. Sample selection

The sample was selected as follows. A set of 1880 NYSE and 4241 Nasdaq common stocks was identified based on two requirements: that the stocks were continuously listed on either the NYSE or the Nasdaq stock market from June to

December 1998, and that data be available in both the Center for Research in Security Prices (CRSP) and the Trade and Quote (TAQ) databases throughout the six-month sample period. Of this set, 78 NYSE stocks and 1160 Nasdaq stocks with beginning-of-sample share prices less than \$5 or greater \$150 were eliminated, leaving an initial sample of 1802 NYSE stocks and 3081 Nasdaq stocks.

The final large-firm sample was selected as follows. The 100 Nasdaq common stocks with the largest beginning-of-sample market capitalization were included in the study. A total of 530 NYSE stocks in the initial sample had market capitalization as large as this set of Nasdaq firms. Of these, the 100 whose market capitalization most closely matched those of the 100 large Nasdaq firms were included in the study.

A small-capitalization subsample was selected by identifying the 450 firms comprising the smallest quartile of the initial NYSE sample. Of these, the smallest 50 firms were eliminated, to avoid including an unusual number of financially distressed firms. One hundred NYSE firms were chosen at random from the remaining 400 small-capitalization NYSE issues. A total of 2380 Nasdaq firms in the initial sample had market capitalization as small or smaller than the final sample of small-capitalization NYSE issues. The 100 whose market capitalization most closely matched those of the small NYSE firms were included in the final small-capitalization sample.

The medium capitalization sample was selected from the 824 NYSE stocks and 601 Nasdaq stocks with beginning-of-sample market capitalization greater than that of the largest firm in the small stock sample and less than that of the smallest firms in the large stock sample. Of these, 100 NYSE stocks were selected at random. The 100 Nasdaq stocks whose market capitalization matched the NYSE stocks most closely were included to complete the sample.

References

- Bacidore, J., Ross, K., Sofianos, G., 2001. Quantifying market order execution quality at the New York stock exchange. *Journal of Financial Markets*, this issue.
- Bessembinder, H., 1999. Trade execution costs on Nasdaq and the NYSE: a post reform comparison. *Journal of Financial and Quantitative Analysis* 34, 387–407.
- Blume, M., Goldstein, M., 1997. Quotes, order flow, and price discovery. *Journal of Finance* 52, 221–244.
- Chan, L., Lakonishok, J., 1997. Institutional equity trading costs: NYSE versus Nasdaq. *Journal of Financial Economics* 52, 713–735.
- Chung, K., Van Ness, B., Van Ness, R., 1999. Limit orders and the bid–ask spread. *Journal of Financial Economics* 255–287.
- Chung, K., Van Ness, B., Van Ness, R., 2000. Spreads, depths, and quote clustering on the NYSE and Nasdaq: evidence from the 1997 SEC Rule changes. Working paper, University of Memphis.
- Ellis, K., Michaely, R., O'Hara, M., 2000. The accuracy of trade classification rules: evidence from Nasdaq. *Journal of Financial and Quantitative Analysis* 35, 529–552.
- Finucane, T., 2000. A direct test of methods for inferring trade direction from intra-day data. *Journal of Financial and Quantitative Analysis* 35, 553–576.
- Glosten, L., 1987. Components of the bid–ask spread and the statistical properties of transactions prices. *Journal of Finance* 42, 1293–1307.
- Hasbrouck, J., Sofianos, G., Sosobee, D., 1993. New York stock exchange systems and trading procedures. NYSE working paper 93-01.

- Heidle, H., Huang, R., 2002. Information-based trading in dealer and auction markets: an analysis of exchange listings. *Journal of Financial and Quantitative Analysis* 37, 391–424.
- Huang, R., Stoll, H., 1996. Dealer versus auction markets: a paired comparison of execution costs on NASDAQ and the NYSE. *Journal of Financial Economics* 41, 313–358.
- Keim, D., Madhavan, A., 1997. Transactions costs and investment style: an inter-exchange analysis of institutional equity trades. *Journal of Financial Economics* 46, 265–292.
- LaPlante, M., Muscarella, C., 1997. Do institutions receive comparable executions in the NYSE and Nasdaq markets? A transactions study of block trades. *Journal of Financial Economics* 45, 97–134.
- Lee, C., Radhakrishna, R., 2000. Inferring investor behavior: evidence from the TORQ data. *Journal of Financial Markets* 3, 83–112.
- Lee, C., Ready, M., 1991. Inferring trade direction from intraday data. *Journal of Finance* 46, 733–746.
- Odders-White, E., 2000. On the occurrence and consequences of inaccurate trade classification. *Journal of Financial Markets* 3, 259–286.
- Perold, A., 1988. The implementation shortfall. *Journal of Portfolio Management* 14, 4–9.
- Peterson, M., Sirri, E., 2002. Evaluation of the biases in execution cost estimates using trade and quote data. *Journal of Financial Markets*, this issue.
- Roll, R., 1984. A simple measure of the effective bid–ask spread in an efficient market. *Journal of Finance* 39, 1127–1139.
- Schultz, P., 2000. Regulatory and legal pressures and the costs of Nasdaq trading. *Review of Financial Studies* 13, 917–958.
- Seppi, D., 1997. Liquidity provision with limit orders and a strategic specialist. *Review of Financial Studies* 10, 103–150.
- Smith, J., Selway, J., McCormick, T., 1998. The Nasdaq Stock market: historical background and current operation. NASD working paper 98-01, National Association of Securities Dealers, Washington, DC.
- Weston, J., 2000. Competition on the Nasdaq and the impact of recent market reforms. *Journal of Finance* 55, 2565–2598.
- Werner, I., 2002. NYSE execution costs. *Journal of Financial Markets*, this issue.