Trade Size and Components of the Bid-Ask Spread

Ji-Chai Lin
Gary C. Sanger
G. Geoffrey Booth
Louisiana State University

The relation between theorized components of the bid-ask spread and trade size for a sample of NYSE firms is examined. We find that the adverse selection component increases uniformly with trade size. Conversely, order processing costs decrease with increases in trade sire for all but the largest trades. We find that or&persistence decreases with trade size. The adverse selection component is highest at the beginning of the day and lowest at the end of the day for all but the largest trades. Trades of NYSE firms executed on regional exchanges or NASDAQ contain a large order processing cost component but no significant adverse information effect.

A dealer or specialist making a market for a security provides immediacy services to traders who want to transact promptly. To compensate the dealer for these services, traders pay relatively high prices when purchasing or receive relatively low prices when selling shares of the security.[1] Early work on the bid-ask spread by Benston and Hagerman (1974), Demsetz (1968), Hamilton (1976), and Tinic (1972) examined the cross-sectional relation between quoted spreads

---

We wish to thank two anonymous referees and Andrew Lo (the editor) for very helpful suggestions in developing this manuscript. We also benefited from the comments of participants at the 1993 Financial Management Association meeting and the discussant, Paul Seguin, and participants at the 1994 American Finance Association meeting and the discussant, Puneet Handa.
Address correspondence and reprint requests to Ji-Chai Lin, Department of Finance, Louisiana State University, Baton Rouge, LA 70803.

[1] In addition to income generated by the bid-ask spread, specialists also earn revenues through trading profits.

The Review of Financial Studies Winter 1995 Vol. 8, No. 4, pp. 1153-1183

and variables such as trading volume, security price, and security risk. More recently an effort has been made to decompose the spread into adverse information and dealer gross profit components [e.g., George, Kaul, and Nimalendran (1991), Glosten and Harris (1988), and Stoll (1989)]. The dealer's gross profit represents compensation for inventory holding costs and order processing costs. Estimates by George, Kaul, and Nimalendran (1991), Hasbrouck (1988), Madhavan and Smidt (1991), and Stoll (1989) suggest that inventory holding costs (e.g., the cost of holding inventory that may deviate from the desired level) appear to be relatively small. Hence, order processing costs and adverse information costs seem to be the dealer's main costs of making a market.

The purpose of this article is to explore empirically the relation between the components of the spread (or dealer costs) and trade size. In particular we examine the extent to which the predictions of existing theories concerning the components of the spread and trade size are supported by the data.

Copeland and Stoll (1990) argue that "[o]rder costs represent the clerical costs of carrying out a transaction, the cost of the dealer's time, and the cost of the physical communications and office equipment necessary to carry out the transaction. To a considerable degree, order costs are fixed with respect to any particular transaction." Because of these fixed costs, the average order processing cost *per share* should decrease as trade size increases.

The effect of informed traders on spreads has been studied by Bagehot (1971), Copeland and Galai (1983), and Glosten and Milgrom (1985), among others. In these models the spread includes an adverse information component that compensates the specialist for losses from trading with better-informed agents. The specialist also adjusts quotes following trades to reflect the information conveyed by trading. Furthermore, Easley and O'Hara (1987) develop a model in which informed traders prefer to trade a larger amount (ceteris paribus) at any given price. If this is true, the adverse information component of the spread should increase with trade size.[2]

Our analysis of the relation between the components of the bid-ask spread and trade size proceeds as follows. Section 1 presents an empirical model of the components of dealer costs. Section 2 describes the transactions data and sample construction methods used in the study, Section 3 reports the empirical relation between the adverse in-

[2] Rational expectations models of informed trading such as Admati and Pfleiderer (1988), Kyle (1985), Subrahmanyam (1991), and others also predict that larger trades will results in a larger price impact (e.g., less market depth, where depth is the reciprocal of the sensitivity of equilibrium price to trade volume).

formation component of the spread and trade size. Section 4 presents a similar analysis for the order processing cost component (dealer's gross profits). Section 5 examines the order arrival process. Section 6 presents a breakdown of previous results by time of day. Section 7 examines trades occurring on the regional exchanges or NASDAQ. Section 8 concludes the paper.

1. Estimating the Components of the Bid-Ask Spread

1.1 The model

In this section we follow Huang and Stoll (1994), Lin (1992), and Stoll (1989) to develop empirical estimates of the components of the bid-ask spread. Consider first a market sell order that is executed at the specialist's bid price B_t at time t. The probability of order persistence (i.e., a buy order following a buy order, or a sell order following a sell order) is d, and the probability of order reversal (i.e., a buy order following a sell order, or a sell order following a buy order) is 1 - d. Thus the probability of the next trade occurring at the specialist's bid B_{t+1} is δ and is 1 - d at the specialist's ask A_{t+1}. Following Stoll (1989), the specialist's expected gross profit at time t + 1, conditional on the trade at time t, is

$$\delta(B_{t+1} - B_t) + (1 - \delta)(A_{t+1} - B_t) = E_t(P_{t+1}) - P_t, \qquad (1)$$

where $E_t(P_{t+1}) = \delta B_{t+1} + (1 - \delta)A_{t+1}$ is the expected future (e.g., equilibrium) market price conditioned on the trade at time t, and $P_t = B_t$ is the transaction price at time t.

Equation (1) measures the trade-by-trade expected gross profit earned by the specialist. As Stoll (1989) notes, the specialist participates in a relatively small portion of all trades on the NYSE.[3] Many incoming market orders are either matched with limit orders or other simultaneous offsetting market orders, or are handled by floor traders. To the extent that the cost structures of floor traders and those who have placed limit orders differ from that of the specialist, our measure of gross profit represents a trade-by-trade average of these profits across the various suppliers of immediacy. It also represents the trader's unconditional estimate of the cost of trading on the NYSE. We discuss the implications of differing cost structures across various suppliers of immediacy services in Section 5.

[3] According to the *New York Stock Exchange Fact Book* (1991), specialists acted as either the buyer or the seller for public or member firm customer orders in 18.4 percent of the share volume traded in 1988, the year which we study in this paper. More recently, specialist participation has varied between 19.0 percent and 19.9 percent of total share volume.

This measure of the specialist's gross profit can be mathematically related to the effective bid-ask spread. Let $Q_t = (A_t + B_t)/2$ be the quote midpoint at time *t*. Huang and Stoll (1994) posit that $z_t = P_t - Q_t$ is one-half the signed effective spread with $z_t < 0$ for a sell order and $z_t > 0$ for a buy order. The effective spread equals the quoted spread for trades executed at the quoted bid or ask price, but is smaller (larger) than the quoted spread for trades executed inside (outside) the quoted spread. For simplicity, we first assume that trades can only be executed at the quoted bid or ask price. This assumption is then relaxed when the regression model is derived.

To reflect possible adverse information revealed by the trade at time *t*, quote revisions are assumed to be $B_{t+1} = B_t + \lambda z_t$ and $A_{t+1} = A_t + \lambda z_t$, where $1 > \lambda > 0$ is the proportion of the spread due to adverse information. Hence, the specialist adjusts downward (upward) both the bid and ask prices following a sell (buy) order.

The dealer's gross profit for a sell order at time *t* is then related to the effective spread as

$$\begin{aligned} E_t(P_{t+1}) - P_t &= [\delta B_{t+1} + (1-\delta)A_{t+1}] - P_t \\ &= \lambda z_t + (1-2\delta)[(A_t + B_t)/2 - B_t] + [(A_t + B_t)/2] - P_t \\ &= -(1-\lambda-\theta)z_t, \end{aligned} \tag{2}$$

where $\theta = 2\delta - 1$. When $\delta = 0.5$ (i.e., following the sell order at time *t*), buy orders and sell orders arrive randomly, then $q = 0$. Choi, Salandro, and Shastri (1988), Hasbrouck (1991a), and Hasbrouck and Ho (1987) show that buy orders tend to follow buy orders, and sell orders tend to follow sell orders. This pattern of order arrivals implies $1 > \delta > 0.5$ and $1 > \theta > 0$.

The gross profit for a buy order at time *t* with the transaction price at the specialist's ask A_t can be similarly derived:

$$\begin{aligned} E_t(P_{t+1}) - P_t &= [\delta A_{t+1} + (1-\delta)B_{t+1}] - P_t \\ &= -(1-\lambda-\theta)z_t. \end{aligned} \tag{3}$$

Equation (3) indicates that $(1-\lambda-\theta)z_t$ is the specialist's expected revenue when the specialist sells at A_t.

Therefore, based on Equations (2) and (3), we measure the dealer's gross profit as a fraction of the effective spread as

$$\gamma = 1 - \lambda - \theta. \tag{4}$$

Since a buy order has $z_t > 0$ (a sell order has $z_t < 0$), Equations (2) and (3) imply that the equilibrium market price is less (greater) than the transaction price for a buy (sell) order. Hence, because of order processing costs, a temporary price effect can be expected for each buy or sell order. The temporary price effect is expressed as a fraction of the effective spread z_t. The magnitude of the temporary price effect depends on three variables: the extent of adverse information (A), the pattern of order arrivals (θ) and the size of the effective spread (z_t). Holding the effective spread constant, the average temporary price effect (e.g., trade price reversal) will be smaller the larger are the adverse information and order persistence parameters. The intuition is that larger quote revisions in response to trades and a greater tendency of buy (sell) orders to follow buy (sell) orders tends to reduce both the number and magnitude of price reversals.

1.2 Estimation

Since λ reflects the quote revision (in response to a trade) as a fraction of the effective spread z_t, and since θ reflects the extent of order persistence, following Huang and Stoll (1994) and Lin (1992), these parameters can be estimated using the following regressions:

$$Q_{t+1} - Q_t = \lambda z_t + e_{t+1}, \tag{5}$$

$$z_{t+1} = \theta z_t + \eta_{t+1}, \tag{6}$$

where the disturbance terms e_{t+1} and η_{t+1} are assumed to be uncorrelated.[4]

Alternatively, with Equations (5) and (6) and $z_{t+1} = P_{t+1} - Q_{t+1}$, the order processing costs for the trade at time t can be related to the effective spread z_t as

$$\begin{aligned} P_{t+1} - P_t &= (Q_{t+1} - Q_t) + z_{t+1} - z_t \\ &= -\gamma z_t + u_{t+1}, \end{aligned} \tag{7}$$

where $\gamma = 1 - \lambda - \theta$ and $u_{t+1} = e_{t+1} + \eta_{t+1}$. Equation (7) suggests that, to the extent that transaction prices tend to bounce between bid and ask prices, changes in transaction prices are somewhat predictable. The extent of the temporary price effect reflects the dealer's costs of order processing. In fact, Equation (7) is a regression form of Equations (2) and (3). Therefore, the parameter y is empirically

[4] As shown by Hasbrouck and Ho (1987), under the assumption of a constant effective spread (i.e., $|z_t| = k$ for all t), Equation (6) implies that if the transaction at time t is a buy (sell) order, there is a probability of $(1+\theta)/2$ that the next transaction will be a buy (sell) order. The conditional probability of order reversal is then equal to $(1-\theta)/2$. The probability structure is more complex if the effective spread changes over time. However, in our model, the parameter θ is a sufficient statistic for the pattern of order arrivals. The disturbance term e_{t+1} may reflect the arrival of public information and market frictions such as price discreteness and the lag in price adjustment due to limit orders. These market frictions may cause serial dependence In the disturbance term. However, we find that, for most of our sample firms, the autocorrelations in the disturbance term are small and that adjusting for these autocorrelations does not have a material effect on the estimate of A.

estimable using a time series of intraday transaction data. Notice that, given Equations (5) and (6), Equation (7) holds whether trades occur at quoted prices or inside the spread.

The fact that order processing costs are largely fixed suggests that our estimate of these costs, y, should be inversely related to trade size. Conversely, based on Easely and O'Hara (1987), Seppi (1990), and others, our estimates of the adverse information component of the spread, λ, should be directly related to trade size. However, one caveat is in order. Seppi (1990, p. 88) points out that most transactions data sets, including the Institute for the Study of Security Markets (ISSM) data utilized here, pool trades that are executed via different market-making mechanisms. For example, most small trades for NYSE firms will be executed in the specialist market (perhaps almost automatically through Super Dot), while larger block trades may be negotiated in the upstairs market. In Seppi's equilibrium model this pooling of trade types could result in a nonmonotonic relation between trade size and price impact. For example, a block trade negotiated in the upstairs market with a credible guarantee that it is not information motivated could have a smaller price impact than a typical small market order executed in the specialist market. Since our data do not allow us to separate trades executed by different trading mechanisms, our estimates of the adverse information component of the spread for different trade sizes is a weighted average of the effect across different trading mechanisms.

To test our hypotheses, we classify the trades of a set of sample firms into seven trade size groups (described in the next section) and examine how the parameter estimates λ, y, and δ vary across the size groups. Although components of the spread may depend on trade size, they may not be linearly related to trade size, and thus the grouping approach is used.

In the empirical tests that follow, we use the logarithms of the transaction price and the quote midpoint in each of the equations. This transformation yields a (continuously compounded) rate of return for the dependent variable and a relative spread for the independent variable. The transformation produces estimates of the cost components of the spread as a percent of the effective spread for easy cross-sectional comparisons, and may also reduce the problem of price discreteness.

2. Data and Sample Description

To estimate empirically the components of the spread we select a sample of 150 common stocks from the 1988 ISSM files. These 150 common stocks are the first 150 NYSE common stocks in the files that have more than 2500 trades and no stock splits in 1988.[5]

All trades in 1988, except for opening transactions on each day, are included for analysis.[6] For each trade, the transaction time, the trade price, and the prevailing bid and ask prices are identified. Lee and Ready (1991) find that prevailing quotes may sometimes be recorded ahead of trades. Following their suggestion, we identify the prevailing quotes for each transaction as the quotes that are in effect five seconds earlier and are BBO eligible (i.e., eligible for inclusion in the National Market System and NASD best bid and offer calculation). The prevailing quote midpoint is then computed as the average of the prevailing bid and ask prices.[7]

Table 1 contains descriptive statistics for our sample of 150 firms. To check the robustness of our results the full sample is partitioned into three subsamples of 50 firms each based on average daily dollar trading volume (the third data item described in Table 1). This partitioning also allows us to explore the possibility of any systematic variations in spread components among thinly versus heavily traded stocks. Such variations may be the result of a difference in the rate of participation by specialists in infrequently versus heavily traded stocks. Average figures in the table are computed by averaging across the 253 trading days in 1988 for each firm and then averaging across firms. Minimum and maximum values are the lowest and highest of the individual (253 day) firm averages. In the description that follows we focus on the full sample figures.

Share price ranges from a minimum of $1.04 to a maximum of $335.79, with an average value of $29.98. This compares reasonably closely with the average share price of all NYSE firms for 1988 of $33.20 (NYSE 1991, p. 11). Average daily share volume is 129,700 shares, with a minimum of 5000 shares and a maximum of 773,200

[5] The requirement of more than 2500 transactions in the sample period is to ensure sufficient observations to carry out reliable estimations in each trade size category. We discard firms with stock splits because several studies show empirically that stock splits tend to change firms' stock returns variances and bid/ask spreads [see Conroy, Harris, and Benet (1990) and the references therein].

[6] Opening transactions are excluded from the sample because they are conducted in a call market, while transactions after the opening are generally conducted in a continuous auction market. In this section of the article, we also discard trades and quotes that are initiated on regional exchanges.

[7] Occasionally, trades with exactly the same time, price, and trading volume are observed on the ISSM files. If there are several trades reported at exactly the same time, price, and volume, we take the first trade and discard the remaining trades. Additionally, Hasbrouk and Ho (1987) point out that a trade maybe reported as multiple trades if different parts of a market order clear at different prices. Thus, they consider only those trades that are separated by at least two minutes. We apply the two-minute requirement to our sample and find that the requirement does not alter the results materially.

Table 1
Summary statistics for sample firms[1]

	Full sample, $N = 150$			Low volume, $N = 50$			Medium volume, $N = 50$			High volume, $N = 50$		
	Average	Min.	Max.	Average	Min.	Max.	Average	Min.	Max.	Average	Min.	Max.
Share price ($)	29.98	1.04	335.79	13.98	1.04	33.09	25.79	6.16	62.04	50.17	12.88	335.79
Daily share volume (00's)	1297	50	7,732	249	50	1128	925	160	2662	2,718	288	7,732
Daily dollar volumes ($000's)	4348	18	24,043	315	18	676	2,011	679	4276	10,717	4,444	24,043
Firm size ($000,000's)	2332	10	19,378	210	10	652	1,124	161	9026	5,664	685	19,378
Relative spread (%)	1.24	0.29	11.46	2.10	0.79	11.46	1.03	0.53	2.55	0.60	0.29	1.28
Dollar spread ($)	0.24	0.11	1.10	0.22	0.11	0.33	0.24	0.15	0.40	0.25	0.15	1.10
Trade size in shares (00's)	20.34	4.00	55.00	14.00	4.00	42.00	22.08	7.00	43.00	24.94	5.00	55.00
Trade size in dollars (000's)	57.2	2.6	184.7	17.7	2.6	43.5	51.5	18.5	94.4	102.4	32.0	184.7
Daily number of trades	82	12	479	24	12	74	60	13	207	160	43	479

[1] The full sample of 150 NYSE firms is partitioned into thirds based on average (per firm) daily dollar trading volume. All figures represent cross-sectional averages, minimums, and maximums of per firm averages over the 253 trading days in calendar year 1988. Firm size is defined as share price times the number of shares outstanding. Relative spread is defined as dollar spread divided by share price.

shares. Average daily dollar volume is $4.3 million, with a minimum (maximum) of $18,000 ($24.0 million). Firm size, measured by average market value of equity is approximately $2.3 billion. The smallest firm in our sample has a market value of $10 million while the largest has a value of $19.4 billion. Relative spreads range from 0.29 percent to 11.46 percent, with an average of 1.24 percent. Corresponding dollar (raw) spreads range from $0.11 to $1.10, with an average of $0.24. Trade size in shares averages 2034 shares, with a minimum of 400 shares and a maximum of 5500 shares. Average trade size for all firms on the NYSE in 1988 was 2303 shares (NYSE 1991, p. 13). Trade size in dollars ranges from $2600 to $184,700, with an average of $57,200. Finally, the average number of trades per day is 82, with a low of 12 per day and a high of 479 per day. In all respects the sample appears to be representative of the NYSE universe.[8]

When comparing the volume-ranked subsamples, the relations among most variables are intuitive. Moving from low- to high-volume firms, the following variables increase on average monotonically: share price, daily share and dollar volume, firm size, trade size in shares and dollars, and daily number of trades. Average quoted dollar spread increases only slightly from $0.22 for low-volume firms to $0.25 for high-volume firms. However, due to the increase in average price level, average relative spread decreases monotonically from 2.10 percent for the low-volume subsample to 0.60 percent for the high-volume subsample. The wide differences in trading activity and liquidity across the volume-ranked subsamples will provide an interesting opportunity to examine the robustness and properties of the spread model.

Table 2 reports average quoted and effective bid-ask spreads in each of seven trade size categories. The trade size categories are created in the following way. For each firm in the sample, all trades in 1988 (except the opening trade each day) are ordered by their size in shares. Each firm's trades are then partitioned into the following seven percentile categories: less than 25 percent, 25-50 percent, 50-75 percent, 75-90 percent, 90-95 percent, 95-99 percent, and greater than 99 percent. The share size cutoffs corresponding to each percentile are determined based on each firm's trade size distribution, thus they vary in absolute share size across firms. The reason for this procedure is to define large and small trades for each firm relative to the order flow experience for the firm over the entire sample period.[9]

[8] Not surprisingly, several of the descriptive variables exhibit considerable skewness across the sample. Median values for daily share volume, daily dollar volume, and firm size are, respectively, 77,700 shares, $1.7 million, and $865 million. All other medians are within 25 percent of their respective means.

[9] A similar procedure is followed by Chan and Lakonishok (1993) in studying institutional trades.

For reference, the median share size corresponding to the percentile cutoffs for the full sample are as follows: 25 percent, 200 shares; 50 percent, 500 shares; 75 percent, 1500 shares; 90 percent, 4200 shares; 95 percent, 7000 shares; 99 percent, 20,500 shares.

The effective dollar spread is the average value of $2|P_t - Q_t|$ while the effective relative (percent) spread is the average of $2|\log(P_t/Q_t)|$.[10] For the full sample the quoted relative spread vanes only slightly, from 1.23 percent for all trade sixes except the largest trades, where it increases significantly, to 1.335 percent.[11] Similar patterns are evident in each of the volume-ranked subsamples. Relative quoted spreads decrease uniformly from the low-volume to the high-volume subsamples for all trade size categories. As we show below this is due to a price effect.

The average effective relative spread increases almost monotonically with trade size in the full sample and in each of the volume-ranked subsamples. Except for trades in the 90 to 95 percent category the effective relative spread increases significantly from one trade size category to the next. The average effective relative spread is less than the average quoted relative spread in each trade size category, indicating that a substantial portion of trades are executed inside the quoted spread. For the full sample, trades above the 99th percentile have an average effective relative spread that is approximately 84 percent of the average quoted relative spread. These figures indicate that demanders of immediacy services rarely received prices which were less favorable than prevailing quotes on the NYSE. In fact, for our 150 sample firms we find that less than 2 percent of all trades above the 95th percentile of the trade size distribution for each firm are executed outside the quoted bid-ask spread.

Quoted dollar spreads average approximately $0.23 for all trade sixes in the full sample (again, except for the largest trades). Note that there is very little variation in average quoted dollar spreads

---

It is motivated by the fact that a (say) 50,000 share trade in a large high-volume firm will likely have a very different impact on the price and quotes of a small low-volume firm. To check the robustness of our results all estimates presented in the paper were computed using absolute shale size categories of 100 shares, 101-500 shares, 501-2099 shares, 2100-9999 shares, 10,000-19,999 shares, 20,000-30,099 shares, and greater than 30,099 shares. Share cutoffs of 2099 shares and 30,099 shares were chosen to detect any possible impact of the NYSE's electronic order routing system, Super Dot. Super Dot. allowed the electronic routing of public market orders of up to 2099 shares and member firm market orders of up to 30,099 shares in 1988. Qualitatively all of the paper's results are insensitive to this change in trade size classification.

10 This measure of the effective percentage spread is approximately equal to $2|(P_t - Q_t)/Q_t|$, where P_t is the trade price and Q_t is the quote midpoint in dollar terms.

11 This increase in quoted spreads for very large trades is also documented by Choe, McInish, and Wood (1991). Significance is determined by performing two-sample *t*- test between adjacent trade size categories.

Table 2
Average quoted and effective bid-ask spread by trade size[1]

Trade size (percentile)	Quoted relative spread (percent)[2]	Effective relative spread (percent)	Quoted dollar spread	Effective dollar spread
Full sample (150 firms)				
< 25	1.238%	0.799%	$0.234	$0.139
25–50	1.247	0.850 + +	0.234	0.152 + +
50–75	1.240	0.898 + +	0.234	0.165 + +
75–90	1.232	0.938 + +	0.232 – –	0.173 + +
90–95	1.222 – –[3]	0.931	0.230 – –	0.172
95–99	1.242	0.963 + +	0.234	0.176 + +
> 99	1.335 + +	1.115 + +	0.249 + +	0.208 + +
Low volume (50 firms)				
< 25	2.087%	1.435%	$0.220	$0.136
25–50	2.106	1.501 + +	0.220	0.147 + +
50–75	2.091	1.563 + +	0.220	0.158 + +
75–90	2.081	1.644 + +	0.218–	0.167 + +
90–95	2.063–	1.634	0.217–	0.168
95–99	2.092	1.684+	0.220	0.170
> 99	2.245 + +	1.860 + +	0.232 + +	0.188 + +
Medium volume (50 firms)				
< 25	1.033%	0.616%	$0.236	$0.137
25–50	1.037+	0.674 + +	0.237	0.152 + +
50–75	1.028	0.721 + +	0.235	0.166 + +
75–90	1.017 – –	0.745 + +	0.233 – –	0.172 + +
90–95	1.012 – –	0.735	0.231 – –	0.169 –
95–99	1.035	0.768 + +	0.237	0.175 +
> 99	1.121 + +	0.924 + +	0.255 + +	0.212 + +
High volume (50 firms)				
< 25	0.594%	0.346%	$0.245	$0.143
25–50	0.598	0.374 + +	0.246	0.156 + +
50–75	0.603 + +	0.410 + +	0.248 + +	0.173 + +
75–90	0.598	0.424 + +	0.246	0.179 + +
90–95	0.590 – –	0.424	0.243 –	0.179
95–99	0.599	0.437 + +	0.247	0.182
> 99	0.639 + +	0.561 + +	0.261 + +	0.225 + +

[1] The spread in each trade size category is computed by first averaging over all transactions for each firm and then across firms within each sample. All but opening transactions in 1988 are used in the analysis.

[2] Denote A_t, B_t, and P_t as the ask, bid, and trade prices at time t. The quoted relative spread is $(A_t - B_t)/Q_t$, where $Q_t = (A_t + B_t)/2$; the effective relative spread is $2|\log(P_t/Q_t)|$; the quoted dollar spread is $A_t - B_t$; and the effective dollar spread is $2|P_t - Q_t|$.

[3] The signs ++ and + (–– and –) indicate that the average spread in the trade size category is significantly larger (smaller) than the average spread in the preceding size category at the 1 percent and 5 percent level, respectively.

across the volume-ranked subsamples. Recall from Table 1 that average share price increases from low- to high-volume stocks. It is this price effect that leads to lower quoted relative spreads for high-volume stocks. Finally, like effective relative spreads, average effective dollar spreads increase significantly with trade size for the full sample and the volume-ranked subsamples.

3. The Adverse Information Component of the Effective Spread

Why does the effective spread increase with trade size? We conjecture that the increase is due to adverse information. The conjecture is based primarily on Easley and O'Hara (1987). They show that "trade size introduces an adverse selection problem into security trading because, given that they wish to trade, informed traders prefer to trade large amounts at any given price." This suggests that large trades tend to convey more information to the market and move quoted prices more than small trades. Thus, subject to the qualifications mentioned in Section 1, we expect the adverse information component of the effective spread to increase with trade size.

To test this hypothesis, we estimate Equation (5) for each of the 150 sample firms in each of the seven trade size categories. The estimate of λ in Equation (5) reflects the adverse information component of the effective spread. We report the average estimate of λ along with average t-statistics and R^2 values in each trade size category in Table 3. The results are virtually the same when Equations (5) and (7) are simultaneously estimated. The fifth column of Table 3 reports the average number of trades per firm within each trade size category. To obtain reliable inferences we include a firm's estimate in a particular trade size category only if at least 30 trades are available.

For the full sample the average estimate of λ is 0.198 for trades below the 25th percentile of the trade size distribution (median trade size below 200 shares). As trade size increases, the average estimate of A increases significantly for all trade size categories through the 95 to 99 percent level (median trade size between 7000 and 20,500 shares). For trades exceeding the 99th percentile of the trade size distribution (median trade size above 20,500 shares) the average estimate of A is 0.626.

These results indicate that the average quote revision is slightly greater than 60 percent of half the effective spread z_t following a large trade, while the average quote revision is only about 20 percent of half the effective spread following small trades. The evidence clearly suggests that the adverse information component of the effective spread increases monotonically with trade size.[12] The estimates are all highly significant, with average t-statistics ranging between 12.94 and 33.50.

[12] In addition to adverse information revealed by trades, in theory, the inventory effect could also partly explain the observed quote revisions. However, Madhavan and Smidt (1991) use specialist Inventory data and show that inventory has a very weak effect on intraday dealer pricing. Similarly, Hasbrouck (1988) uses a sophisticated VAR model and finds the Inventory effect to be insignificant. Finally, Hasbrouck and Sofianos (19993) use NYSE specialist inventory data and find "little support for the classic inventory control mechanism." Hence, we focus on adverse information in the interpretation of our quote revision results.

Table 3
Estimates of the adverse selection component of the effective spread[1]

Trade size (percentile)	Average estimate λ	Average t-statistic on λ	Average R^2	Average number of trades per firm
Full sample (150 firms)				
< 25	0.198	22.51	17.36%	4382
25–50	0.287 + +[2]	26.66	24.34	3385
50–75	0.392 + +	33.50	30.91	3256
75–90	0.483 + +	31.41	34.98	2005
90–95	0.569 + +	21.15	39.52	649
95–99	0.609 + +	20.29	39.63	488
> 99	0.626	12.94	32.45	133
Low volume (50 firms)				
< 25	0.213	14.76	17.86%	1378
25–50	0.311 + +	17.29	24.90	1075
50–75	0.419 + +	20.82	31.72	942
75–90	0.501 + +	19.09	35.35	543
90–95	0.590 + +	12.88	40.87	190
95–99	0.642 + +	12.74	41.66	147
> 99	0.670	8.71	37.68	47
Medium volume (50 firms)				
< 25	0.215	22.29	19.73%	3183
25–50	0.316 + +	26.26	27.55	2519
50–75	0.424 + +	33.30	33.64	2409
75–90	0.523 + +	30.59	37.43	1423
90–95	0.601 +	20.66	41.73	499
95–99	0.644	20.19	40.54	370
> 99	0.658	12.91	32.08	104
High volume (50 firms)				
< 25	0.164	30.48	14.50%	8583
25–50	0.233 + +	36.44	20.59	6560
50–75	0.331 + +	46.40	27.36	6418
75–90	0.424 + +	44.56	32.21	4051
90–95	0.516 + +	29.91	35.97	1258
95–99	0.541 +	27.92	36.70	949
> 99	0.550	17.20	27.58	249

[1] The quote revision model is estimated for each of the 150 sample firms in each of the seven trade size categories, and for low-, medium-, and high-volume samples:

$$\Delta Q_{t+1} = \lambda z_t + e_{t+1}, \tag{5}$$

where $\Delta Q_{t+1} = Q_{t+1} - Q_t$, Q_t is the log quoted midpoint at time $z_t = P_t - Q_t$, P_t is the log trade price at time t; λ is the adverse information component of the effective spread. All but opening transactions in 1988 are used in the analysis.

[2] The signs + + and + (-- and -) indicate that the average estimate of λ in the bade size category is significantly larger (smaller) than the average estimate in the preceding size category at the 1 percent and 5 percent level, respectively.

Results for the low-, medium-, and high-volume subsamples are qualitatively similar to those for the full sample. Estimates of the adverse information component of the spread are somewhat lower in each trade size category for the high-volume firms. This could be due to greater informational efficiency for large widely held firms. Such firms tend to have a high level of monitoring and public information production by securities analysts and others [e.g., Bhide (1993) and Ho and Michaely (1988)]. This effect could also be due partly to the fact that the minimum quote adjustment on the NYSE is 1/8 and the high-volume sample firms also have higher prices making the minimum proportional quote adjustment smaller. Overall our results are consistent with Easley and O'Hara's (1987) model and supportive of our conjecture that the increase in effective spreads with trade size is the result of an information effect.

Our measure of the adverse information (permanent) component of the spread could be biased for several reasons, especially for large trades. Continuity requirements and the presence of limit orders may prevent the specialist from immediately adjusting quotes to a new equilibrium level. As a test of robustness, we reestimated Equation (5) using the change in quote midpoint from trade t through trade $t + 3$ as the dependent variable. The slope coefficient measures the impact of the current trade on the quote change over the time interval spanned by the following three trades.[13] Using this procedure, estimates of the adverse information component are slightly larger for each trade size category, though the pattern of larger coefficients for larger trade sizes remains unchanged.

Another potential problem with our estimate of the adverse information component of the spread is that it does not use information contained in previous trades or quotes. Hasbrouck (1991b) develops a method which does utilize information from past trades and quotes.[14] Specifically, he uses a vector autoregression on trades and quotes to decompose the variance of changes in the (implicit) efficient price into components that are correlated and uncorrelated with trades. The tradecorrelated component in Hasbrouck's model is interpreted as a measure of trade informativeness. Hasbrouck (1991b) reports impulse response coefficients for market value quartiles of NYSE firms using ISSM data from the first quarter of 1989. The impulse response coefficient measures the persistent impact of a 1000 share trade on the log quote midpoint. In our model, this effect can be estimated as $|z_t|\lambda$ with $|z_t|$ and λ estimated using trades of 1000 shares. Trades of 1000 shares fall in the 50 to 75 percent trade size category for each of the low-, medium-, and high-volume subsamples.[15] Thus, using

---

[13] The choice of a three-trade lag is motivated by Holthausen, Leftwich, and Mayers (1990) who found that price adjustment following large block trades is essentially completed by the third trade following the block. We experimented with other lag lengths and found little variation in results.

[14] Also see Blume, Easley, and O'Hara (1994) who develop a model in which sequences of prices and trading volumes are informative.

[15] The average of the 50th and 75th percentiles of the trade size distribution for the low-, medium-,

estimates of $|z_t|$ from Table 2 and estimates of λ from Table 3, we obtain the following measures of the persistent impact of a 1000 share (approximate) trade: 1.563/2 × 0.419 = 0.328 for the low-volume sub sample; 0.153 for the medium-volume subsample; and 0.068 for the high-volume subsample.[16] Our three subsamples compare reasonably closely with Hasbrouck's subsamples 2, 3, and 4 in terms of market capitalization and quoted spread. His impulse response coefficients are, respectively, 0.324, 0.162, and 0.078 [Hasbrouck (1991b), p. 586, Table 1]. Given the approximate nature of this comparison and the fact that different calendar periods are used in the two papers, we find the estimates of the permanent effects of trades to be surprisingly similar. We next examine the relation between trade size and the order processing cost component of the spread.

4. Empirical Estimates of the Order Processing Cost Component of the Effective Spread

To obtain estimates of the order processing cost component of the effective spread, we estimate *y* using Equation (7) for each of the 150 sample firms in each of the seven trade size categories. The average estimates of *y* and corresponding *t*- statistics are reported in columns 2 and 3 of Table 4. The average number of observations per firm in each trade size category is the same as in Table 3 and, hence, is not repeated here. For the full sample, the average estimate of *y* for the smallest trades is 0.460. It declines significantly to 0.353 for trades in the 25 to 50 percent category, and continues to decline to 0.201 for trades in the 95 to 99 percent category. The average estimate of *y* then increases to 0.224 for the largest trades. Average t-statistics range from 3.28 to 27.73, indicating that estimates of dealers' order processing costs are significant for all trade sizes. Results for the volume-ranked subsamples are qualitatively very similar to those for the full sample. They decrease as trade size increases for all but the largest trades.

Recall that the dealer's gross profit should largely reflect order processing costs. We continue to assume that inventory holding costs are minor. To better compare dealer gross profits across trade sizes, we report order processing costs as a percent of the trade price, computed as *y* times mean $|z_t|$ in column 4. On average, for the full sample, order processing costs are 0.196 percent of trade price for small trades, and they gradually decrease as trade size increases reaching 0.112 per-

and high-volume subsamples are respectively 500:1000 shares, 500:1500 shares, and 600:1850 shares. Note that our measure only aproximates the persistent impact of a 1000 share trade because all trades in the 50–75 percent trade size category are used in estimating $|z_t|$ and λ.

[16] Note that Table 2 reports values for $2 \cdot |z_t|$.

Table 4
Estimates of the order processing Cost component of the effective spread[1]

Trade size (percentile)	Average estimate γ	Average t-statistic on γ	Average[2] OPC (%)	Average[3] OPC (cents/share)
Full sample (150 firms)				
< 25	0.460	27.73	0.196%	3.00¢
25–50	0.353 – –[4]	19.94	0.158	2.45
50–75	0.282 – –	15.69	0.147	2.10
75–90	0.238 – –	10.57	0.125	1.94
90–95	0.215 – –	5.55	0.112	1.76
95–99	0.201	4.88	0.114	1.69
> 99	0.224	3.28	0.148	2.46
Low volume (50 firms)				
< 25	0.471	16.87	0.359%	3.20¢
25–50	0.379 – –	12.25	0.287	2.69
50–75	0.324 – –	10.01	0.278	2.42
75–90	0.279 – –	6.79	0.238	2.26
90–95	0.262	3.76	0.221	2.14
95–99	0.242	3.20	0.231	1.98
> 99	0.257	2.09	0.280	2.41
Medium volume (50 firms)				
< 25	0.443	24.29	0.145%	2.94¢
25–50	0.328 – –	17.36	0.117	2.33
50–75	0.263 – –	13.72	0.103	1.99
75–90	0.223 – –	9.11	0.089	1.84
90–95	0.188 – –	4.64	0.071	1.55
95–99	0.175	4.18	0.071	1.56
> 99	0.180	2.48	0.095	2.29
High volume (50 firms)				
< 25	0.466	42.03	0.085%	2.86¢
25–50	0.352 – –	30.20	0.071	2.34
50–75	0.260 – –	23.34	0.059	1.87
75–90	0.213 – –	15.80	0.049	1.71
90–95	0.194	8.25	0.043	1.60
95–99	0.186	7.27	0.042	1.53
> 99	0.236 +	5.27	0.069	2.69

[1] The order processing cost (OPC) model is estimated for each of the 150 sample firms in each of the seven trade size categories, and for low-, medium-, and high-volume subsamples:

$$\Delta P_{t+1} = -\gamma z_t + u_{t+1}, \tag{7}$$

where $\Delta P_{t+1} = P_{t+1} - P_t$; $z_t = P_t - Q_t$, P_t is the log trade price at time t and Q_t is the log quote midpoint at time t; γ is the cost of order processing as a fraction of the effective spread z_t. All but opening transactions in 1988 are used in the analysis.

[2] The percentage order processing cost for each sample firm in a given trade size category is computed as γ times mean $|z_t|$.

[3] The order processing cost in cents per share for each sample firm in any trade size category is computed as γ times mean $|z_t|$ times mean Q_t, where $Q_t = (A_t + B_t)/2$, the quote midpoint in dollar terms.

[4] The signs ++ and + (– – and –) indicate that the average estimate of γ in the trade size category is significantly larger (smaller) than the average estimate in the preceding size category at the 1 percent and 5 percent levels, respectively.

cent for trades in the 90 to 95 percent category. Order processing cost as a percent of trade price then levels off and increases somewhat, though not significantly, for very large trades.

The corresponding cost of order processing in cents per share for each trade size category is reported in column 5 of Table 4. The average order processing cost in cents per share is computed as y times mean $|z_t|$ times mean Q_t, where Q_t is the prevailing quote midpoint. For the full sample, the average order processing cost is 3.0 cents per share for the smallest trades, which is the highest among the seven trade size groups. The lowest cost of order processing is 1.69 cents per share for trades in the 95 to 99 percent size category. For trades above 99 percent in the size distribution, the cost of transacting is 2.46 cents per share. The same patterns are apparent in the volume-ranked subsamples.

The mean order processing cost averaged over all seven trade size categories for the 150 sample firms is 0.16 percent, or 2.23 cents per share. Order processing costs should reflect the gross profits earned by the dealer. From NYSE and SEC data, Smidt (1988) estimates the gross profits earned by NYSE specialists based on *specialist transactions* for the 11 year period from 1970 through 1980. He reports that “on each transaction the average gain of the specialist is about 2.25 cents.” Therefore, our mean estimates of the order processing cost component of the spread accurately reflect the historical average gross profits of NYSE specialists. Our mean estimates are also very close to those estimated by Hasbrouck and Schwartz (1988) who compute order processing costs based on excessive short-term price volatility. They estimate that average order processing costs are about 2.5 cents per share on the AMEX, and about 3.0 cents per share on the NYSE.

The fact that order processing costs decline as trade size increases is likely the result of a spreading of fixed costs over more shares for large orders. This is especially likely for orders coming in over the NYSE’s Super Dot system that are matched with either an existing limit order, the specialist’s inventory, or a market or discretionary order arriving nearly simultaneously from the trading crowd. Order processing costs may increase for the largest trades because their size requires extra effort and/or costs to execute. In fact, Hasbrouck, Sofianos, and Sosebee (1993, Table 5) show that the frequency of upstairs facilitated blocks increases with block size. Thus our results could reflect in part the difference in cost structures between the specialist and upstairs traders.

The fact that order processing costs or dealer’s gross profits are highest for small trades may explain one feature of the practice of paying for order flow. Non-NYSE dealers have recently begun paying brokers on average a penny per share for directing trades to them [see Lee (1993)]. However, there are usually size limitations placed on the solicited orders. Based on our results these orders have the lowest

Table 5
Estimates of the persistence of order arrival[1]

Trade size (percentile)	Average estimate θ	Average t-statistic on θ	Implied[2] δ	Average R^2
Full sample (150 firms)				
< 25	0.343	26.67	0.67	12.80%
25–50	0.360 ++[3]	23.76	0.68	14.35
50–75	0.326 −−	24.52	0.66	12.09
75–90	0.279 −−	16.36	0.64	9.52
90–95	0.216 −−	6.73	0.61	6.43
95–99	0.190	5.65	0.60	5.56
> 99	0.150 −−	2.78	0.58	4.77
Low volume (50 firms)				
< 25	0.315	11.35	0.66	10.60%
25–50	0.310	10.23	0.66	10.68
50–75	0.257 −−	8.09	0.63	7.43
75–90	0.220 −	5.57	0.61	6.19
90–95	0.148 −−	2.29	0.57	3.98
95–99	0.117	1.55	0.56	2.64
> 99	0.073	0.82	0.54	3.43
Medium volume (50 firms)				
< 25	0.342	18.64	0.67	12.29%
25–50	0.355	18.31	0.68	13.36
50–75	0.313 −−	16.42	0.66	10.22
75–90	0.254 −−	10.67	0.63	6.95
90–95	0.211 −−	5.36	0.61	5.07
95–99	0.181 −	4.34	0.59	4.53
> 99	0.163	2.35	0.58	4.20
High volume (50 firms)				
< 25	0.370	50.04	0.69	15.52%
25–50	0.415 ++	42.72	0.71	19.02
50–75	0.409	49.05	0.70	18.62
75–90	0.363 −−	32.86	0.68	15.41
90–95	0.289 −−	12.54	0.64	10.23
95–99	0.273	11.07	0.64	9.52
> 99	0.215 −−	5.17	0.61	6.69

[1] The order persistence model is estimated for each of the 150 sample firms in each of the seven trade size categories, and for low-, medium-, and high-volume subsamples:

$$z_{t+1} = \theta z_t + \eta_{t+1}, \tag{6}$$

where $z_t = P_t - Q_t$, P_t is the log trade price at time t; Q_t is the log quote midpoint at time t; θ is the measure of order persistence.

[2] δ is the probability of a buy (sell) order following a buy (sell) order implied by the spread model; $\delta = (\theta + 1)/2$.

[3] The signs ++ and + (−− and −) indicate that the average estimate of θ in the trade size category is significantly larger (smaller) than the average estimate in the preceding size category at the 1 percent and 5 percent level, respectively.

adverse information content and the highest dealer gross profits. If non-NYSE dealers have relatively lower costs for small trades, it could be profitable to "buy" this part of the order flow.

5. The Persistence of Order Flow

Several previous researchers, including Choi, Salandro, and Shastri (1988), Hasbrouck (1991a), and Hasbrouck and Ho (1987) find that buy (sell) orders tend to follow buy (sell) orders. We refer to this phenomenon as order persistence. It could be caused by several factors including the splitting of large orders into two or more small-orders, "working" of orders by floor brokers, slow adjustmerit of limit orders when new information arrives in the market, or requirements that the specialist maintain an "orderly" market. As can be seen from the spread model developed in Section 1, the extent of order persistence influences the decomposition of the spread into adverse information and order processing cost components.

To examine order persistence we estimate Equation (6) of Section 1 in a manner analogous to that used for the adverse information and order processing cost equations. Table 5 contains mean values, *t*-statistics, and R^2 values for the order persistence parameter θ. We also present $\delta = (\theta + 1)/2$, which is the probability of a buy (sell) order following another buy (sell) order.

For the full sample the average order persistence parameter is 0.343 for the smallest trades, implying the probability of a buy or sell order following a similar order of 67 percent. For both the full sample and the high-volume subsample the order persistence parameter increases significantly for trades in the 25 to 50 percent size category. However, this translates into only a slight increase in the probability of a buy or sell following another like order. As trade size increases further, the order persistence parameter decreases significantly to reach 0.150 for the largest trades in the full sample. This is equivalent to a 58 percent probability of a buy or sell order following another buy or sell. A similar pattern is evident in the volume-ranked subsamples. Intuitively we expect order persistence to decline as trade size increases because larger trades produce larger quote revisions (Table 3). Larger quote revisions should, in turn, discourage same-direction trades. Interestingly, even for very large trades, which generate the largest quote revisions, a buy (sell) order is still more likely to follow a large buy (sell) order than an order of the opposite type.

One final point concerns the degree of similarity of our estimates of the spread components across the volume-ranked subsamples for any given trade size category. To the extent that the mix of suppliers of immediacy services varies across the volume groups (e.g., greater specialist participation for low-volume stocks), our estimates imply that large differences in participants cost structures do not exist. This could be due in part to the fact that traders competing with the specialist who might have a temporary cost advantage need only match specialist quotes to take priority.

6. Order Flow, Spread Components, and Time of Day

Several previous studies have reported intradaily patterns in price variability, trading volume, bid-ask spreads, and the components of bid-ask spreads. Both trading volume and price volatility follow a U-shaped pattern, being higher at the beginning and end of the day than during the middle of the day [Foster and Viswanathan (1993), Harris (1986), Jain and Joh (1988), Lockwood and Linn (1990), McInish and Wood (1990a,b), and Wood, McInish, and Ord (1985)]. A similar pattern in quoted bid-ask spreads is documented by McInish and Wood (1992). Wei (1992) finds a U-shaped pattern for the adverse information component of the spread and an inverted U-shaped pattern for the order processing component, while both Foster and Viswanathan (1993) and Hasbrouck (1991b) report a higher adverse information component at the beginning of the day relative to later in the day. Finally, Lee, Mucklow, and Ready (1993) find quoted depth to be relatively low early in the day and relatively high during the middle of the day.

Theories predicting intradaily variation in trading volume, price variability, and trading costs have been developed by Admati and Pfleiderer (1988) and Brock and Kleidon (1992). Admati and Pfleiderer (1988) predict periods of concentrated trading in which both discretionary liquidity traders and informed traders participate. Such periods are characterized as having high trading volume and low adverse information costs (due primarily to competition among informed traders). Brock and Kleidon (1992) predict high trading volume, due to portfolio rebalancing, and high trading costs, due to monopoly power of market makers, both before and after periods when the market is closed.[17] In this section of the article we present evidence of intradaily variation in trading volume, effective bid-ask spreads, components of the bid-ask spread, and order persistence broken down by trade size. For brevity we present only a general description of the more salient of our results for the full sample of 150 firms. The numerical results and results for the volume-ranked subsamples (which are similar to those discussed here) are available on request. To test for the significance of differences by time of day for each variable examined, we perform Wilcoxen sign rank tests comparing morning (9:30-10:30 A.M.)

[17] As Brock and Kleidon (1992) argue, the assumption of monopoly power by market makers is not necessarily to achieve their results.

with midday (10:30 A.M.- 3:00 P.M.), and afternoon (3:00-4:00 P.M.) with midday.[18]

Consistent with earlier studies we find that average share volume per hour (in millions of shares) follows a U-shaped pattern for all trade size categories, with midday volume being lower than either morning or afternoon volume. This tendency is somewhat more pronounced for larger trades, and larger trades (above the 50th percentile of the per-firm trade size distribution) account for the bulk of total trading volume. These differences are significant at the 1 percent level for all trade size categories.

The number of trades per hour also follows a U-shaped pattern for all trade size categories. The number of trades per hour is significantly higher at the beginning and end of the day for all trade size categories.

The average effective relative bid-ask spread, shown in Figure 1, also follows a U-shaped (or reverse J-shaped) pattern during the day for all trade size categories. Both morning and afternoon spreads differ significantly from midday at the 1 percent level for all trade size categories except the largest, where the differences are significant at 5 percent. This result complements McInish and Wood (1992) who find a similar pattern in quoted spreads unconditional on trade size. Note that high effective relative spreads and high trading volume at the beginning and end of the day support Brock and Kleidon's (1992) periodic market closure model. This relation holds across all trade size categories. This result is also consistent with Lee, Mucklow, and Ready (1993) who find that market makers both widen the spread and decrease quoted depth at times of high informational asymmetry (e.g., morning versus later in the day). At such times, a wider quoted spread, along with fewer crosses inside the spread, would both tend to result in a higher effective spread. Next we consider whether high effective relative spreads at the beginning and end of the day are due primarily to adverse information or to order processing costs.

Figure 2 presents the estimated average adverse information component of the spread by trade size and time of day. Except for the largest 1 percent of trades, the adverse information component is highest during the first hour of the trading day. It decreases during the middle of the day and is the lowest in the last hour of trading. The adverse information component at the beginning of the day is significantly higher than midday for the first six trade. size categories (up to trades at the 99th percentile of the trade size distribution) at the 1 percent level. It is significantly lower at the end of the day for trades

[18] The results are qualitatively similar using either the Fisher sign test or the two-sample *t*- test.

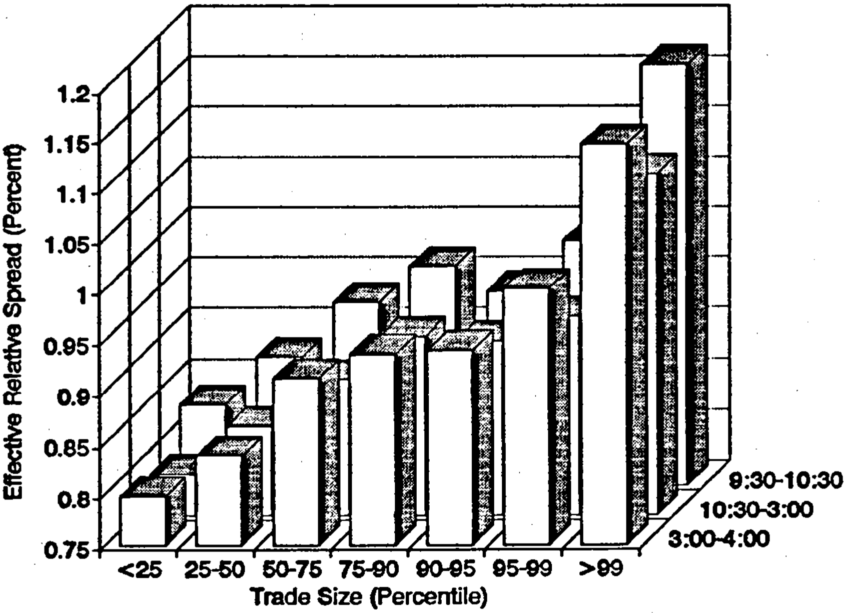

Figure 1
Average effective relative bid-ask spread by trade size and time of the day
The intraday pattern of average relative bid-ask spread for all trade size categories. Note the intraday pattern follows a U-shaped curve and morning and afternoon spreads differ significantly from midday for all trade size categories.

in the 50 to 75 percent, 75 to 90 percent, 90 to 95 percent, and 95 to 99 percent categories. This evidence is consistent with the notion that information acquired overnight would give informed traders the largest advantage early in the day before the information becomes more widely disseminated, either through trading or news releases. Even in the absence of private information, market participants face higher uncertainty regarding the true value of a security following the overnight nontrading period. This uncertainty could cause quotes to be more sensitive to trades early in the day. However, for the largest trades, time of day does not appear to be relevant. Again, this may be due to the fact that many of the largest trades are negotiated and executed via different mechanisms than smaller trades. The high adverse information component from 9:30 to 10:30 A.M. coincides with high trading volume, and hence, is inconsistent with the Admati and Pfleiderer (1988) model of concentrated trading. However, the lowest adverse information component of the day (3:00 4:00 P.M.) also coin-

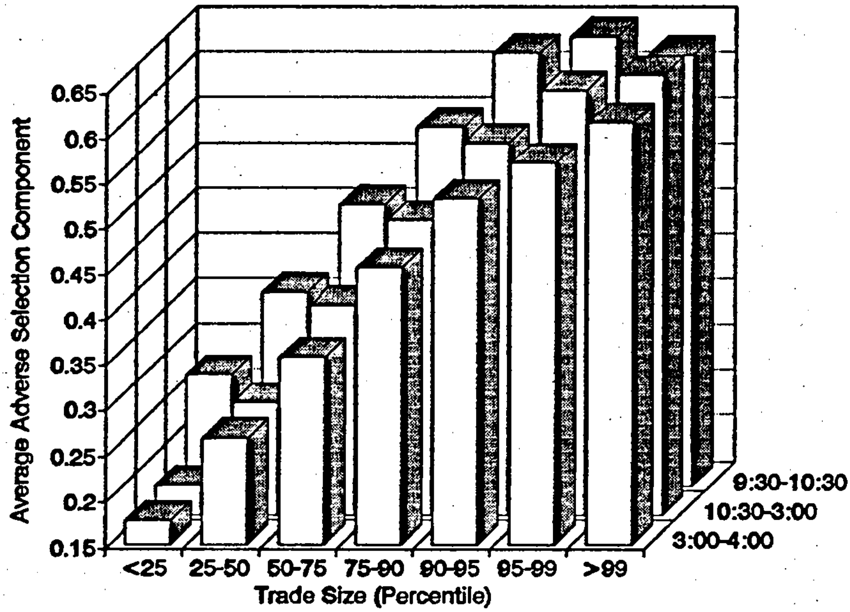

Figure 2
Average adverse information component of the spread by trade size and time of the day
The estimated average adverse information component of the spread by trade size and time of the day. Except for the largest 1 percent of trades, the adverse information component is highest during the first hour of the trading day.

cides with a period of high trading volume, which is consistent with Admati and Pfleiderer (1988).[19]

Consistent with the findings of Wei (1992), we find that the order processing cost component of the spread follows an inverse U-shaped pattern. Order processing costs are highest during the middle of the day for all trade size categories except for trades in the 95 to 99 percent category. Order processing costs are significantly lower in the 9:30 to 10:30 A.M. period (relative to midday) for all trade size categories at the 1 percent level. Order processing costs are significantly lower in the afternoon relative to midday for the first five trade size categories at the 5 percent level. Since there is no reason to expect the fixed costs of order processing to vary by time of day, we speculate that this result may reflect variation in specialists' *net* profit during the

[19] Note that the predictions of rational expectations models for volume and trading costs are sensitive to several underlying assumptions. For example, Subrahmanyam (1991) extends Admati and Pfleiderer (1988) to allow for risk averse informed traders, and shows that an increase in the number of informed traders may cause a decrease in market liquidity.

day. In fact, this pattern is consistent with the model of Glosten (1989) in which the specialist averages profits over time. The specialist may incur losses early in the day when information asymmetry is relatively high, and then make up these losses later in the day after learning from trading. The specialist's ability to average profits over time may be facilitated by two additional factors. First, as adverse information decreases during the day, the minimum tick size sets a lower bound on both quoted and effective spreads. Second, if higher price volatility early in the day removes limit orders from the specialist's book, competition from such orders may decrease later in the day.

The persistence of the order arrival process also varies during the day. Except for the largest 1 percent of all trades, order persistence is lowest in the first hour of trading and highest in the last hour of trading. Order persistence is significantly lower in the morning relative to midday for trades in the 25 percent, 25 to 50 percent, 50 to 75 percent, and 99 percent categories at the 5 percent level. Order persistence in the afternoon is significantly greater than at midday for the first five trade size categories at the 1 percent level. This pattern for order persistence is not surprising since we have observed that trades move quotes by a larger amount earlier in the day.

Finally, we note that the patterns in the adverse information cost, order processing cost, and the persistence of order arrivals across trade size categories are invariant to time of day. Adverse information costs increase with trade size, order processing costs decrease with trade size except for the largest trades, and the persistence of order arrivals first increases (to trades in the 25 to 50 percent category) and then decreases with trade size.

7. Regional and NASDAQ Trades

All of the foregoing evidence excludes trades of NYSE firms on either the regional exchanges or on NASDAQ. The purpose of this was to focus on the execution of orders within the institutional constraints of the NYSE. However, as highlighted by Lee (1993) the degree of integration (or fragmentation) across electronically linked markets and its impact on the price discovery process is of growing interest. Hence, in this section of the article we examine the adverse information cost, the order processing cost, and the persistence of order arrivals for trades executed away from the NYSE.

All trades, both NYSE and non-NYSE, during 1988 are ranked by share size for each of the 150 sample firms. As before, trades are placed into seven trade size categories. The spread model defined by Equations (5) through (7) is then estimated for each trade size category for NYSE trades and non-NYSE trades. For all trades, $z_t = P_t - Q_t$ is computed using either the NYSE or non-NYSE trade price at time t and the prevailing quote midpoint. Subsequent trade price revisions, ΔP_{t+1}, and trade direction indicators, z_{t+1}, are computed using pooled NYSE and non-NYSE trades. Subsequent quote changes ΔQ_{t+1} use only NYSE quotes.

Table 6 contains the results of the analysis. The top panel reports results for the regional exchanges, while NASDAQ results appear in the bottom panel. Because sample sizes are rather small and results do not vary much among the regional exchanges, we report their results combined. The regional exchanges included in the analysis are Boston, Cincinnati, Midwest, Pacific, and Philadelphia. We also include trades executed on Instinet, which is an electronic order routing and execution system used primarily by institutions.

First notice that a firm is included in a particular trade size category only if it had a sufficient number of regional or NASDAQ trades in that size category to permit estimation of the model (15 trades).[20] This restriction ensures that comparisons of NYSE and non-NYSE trades are made on a same-firm basis. The number of firms in each trade size category is reported in column 1 of Table 6 and the average number of trades per firm on the NYSE/non-NYSE are reported in the last column. The number of firms in the analysis declines as trade size increases, indicating that as the size of a trade increases, the less likely it is to be executed away from the NYSE. For example, only 53 firms (5 firms) out of our total sample of 150 firm had 15 or more trades above the 99th percentile of the trade size distribution on any regional exchange (NASDAQ) during 1988.

The most striking feature of Table 6 is the absence of an adverse information effect for trades executed at either the regional exchanges or the NASDAQ market. The adverse information component of the spread (λ) for regional and NASDAQ trades is significantly lower than that for NYSE trades in all trade size categories. In most cases the average adverse information component for regional and NASDAQ trades does not differ significantly from zero. Furthermore, in contrast to NYSE trades, the adverse information component of the spread does not increase systematically with trade size for regional or NASDAQ trades.

Several possibilities may help to explain the lack of an information effect for non-NYSE trades. First, as Lee (1993) points out, inducements offered by the regional exchanges and NASDAQ market makers to brokers for order flow usually include restrictions that are likely

[20] Restricting the analysis to firms with at least 30 trades in a particular trade size category had only a minor effect on the results.

Table 6
Comparison of the adverse information component (λ), the order processing component (γ), and the persistence of order arrival (θ) for trades on the NYSE, regional exchanges, and NASDAQ

Trade size (percentile)	Number of firms	Adverse information λ	Order processing γ	Order persistence θ	Average number of trades per firm
NYSE/regional[1]					
< 25	145	0.183/0.093 (−9.26)	0.480/0.665 (13.67)	0.338/0.243 (−8.93)	4028/2255
25–50	144	0.256/0.085 (−13.35)	0.401/0.644 (15.80)	0.343/0.271 (−5.85)	3084/1324
50–75	144	0.351/0.101 (−18.32)	0.342/0.608 (16.77)	0.306/0.291 (−1.29)	3541/1125
75–90	139	0.424/0.079 (−24.80)	0.310/0.649 (16.49)	0.266/0.272 (0.54)	2433/353
90–95	119	0.494/0.049 (−26.64)	0.295/0.731 (14.88)	0.211/0.221 (1.09)	1049/97
95–99	108	0.542/0.034 (−31.63)	0.245/0.762 (22.79)	0.213/0.203 (−1.10)	844/75
> 99	66	0.490/0.033 (−18.04)	0.295/0.779 (15.89)	0.215/0.188 (−1.15)	340/38
NYSE/NASDAQ[1]					
< 25	90	0.142/0.018 (−13.10)	0.521/0.824 (14.31)	0.337/0.158 (−9.60)	5516/426
25–50	86	0.191/0.031 (−14.51)	0.448/0.787 (13.41)	0.361/0.182 (−8.01)	4300/246
50–75	79	0.283/0.043 (−15.27)	0.385/0.812 (16.91)	0.331/0.146 (−8.61)	5124/209
75–90	54	0.338/0.012 (−15.56)	0.359/0.842 (11.72)	0.304/0.146 (−4.08)	4334/77
90–95	19	0.379/0.069 (−5.15)	0.397/0.704 (4.16)	0.224/0.226 (0.04)	2478/41
95–99	16	0.486/0.090 (−6.13)	0.288/0.761 (5.78)	0.226/0.149 (−0.92)	1581/25
> 99	5	0.428/ −0.001 (−5.29)	0.253/0.888 (7.85)	0.319/0.113 (−3.57)	648/18

[1] The number to the left of the slash is the aveage estimate for the NYSE; the number to the right of the slash is the average estimate for the regional exchanges (top panel) or the NASDAQ (bottom panel). *t*-statistics (paired) for the difference between NYSE and non-NYSE estimates appear in parentheses below the estimates. The regional exchanges include Boston, Cincinnati, Midwest, Pacific, Philadelphia, and Instinet.

to exclude information-based trades (e.g., size and source restrictions [Lee (1993), p. 1014]). Second, large trades executed on the regional exchanges or NASDAQ may be negotiated in the "upstairs" market and may contain guarantees that they are not information motivated as described in Seppi (1990). Also, in 1988 the regionals had a "clean cross" rule in effect which allowed brokers to execute matched buy and sell orders without including the limit order book. Due to SEC restrictions the NYSE did not adopt a similar rule until 1992. Third, this result is consistent with Admati and Pfleiderer's (1988) prediction that informed traders prefer to go where trading volume is most concentrated, which here would be the NYSE. Finally, NYSE quotes might not respond to non-NYSE trades due to informational lags. To check this possibility we reestimate the model using the NYSE quote in effect 15 seconds after a regional or NASDAQ trade. The results were nearly identical to those presented.

Order processing costs (dealers' gross profits), in contrast to adverse information costs, are significantly larger for both regional and NASDAQ trades versus NYSE trades for all trade sixes. In most trade size categories this cost is at least twice as large for non-NYSE trades. It is also interesting to note that order processing costs do not decline as trade size increases for non-NYSE trades. The argument that (largely) fixed order costs should be spread over a larger number of shares, thus causing order processing costs to decline as trade site increases, does not hold for non-NYSE trades.

Finally, the persistence of the order flow is somewhat lower, following regional and NASDAQ trades than NYSE trades.[21] Order persistence is significantly lower for regional versus NYSE trades for the smallest two trade size categories. For NASDAQ trades, order persistence is significantly lower than for NYSE trades for all trade sizes except the 90 to 95 percent and 95 to 99 percent categories. It is interesting that order persistence is lower for non-NYSE trades despite the fact that quotes are not revised much (e.g., low adverse information component) following non-NYSE trades. This could, however, be explained by less activity by floor brokers on the regionals, as floor brokers frequently split up orders in an attempt to get better execution.[22]

[21] Here order persistence measures the tendency of a buy (sell) order to follow another buy (sell) order regardless of location (e.g., all trades are pooled).

[22] We repeated the analysis of this section using regional as well as NYSE quotes and obtained similar results. The adverse information component was slightly larger, indicating that regional trades do have some effect on regional quotes. However, in all trade size categories the adverse information effect is significantly lower for non-NYSE trades.

8. Conclusions

This article examines the relations between trade size and components of the bid-ask spread. The contributions of the article are summarized as follows. Following Huang and Stoll (1994), Lin (1992), and Stoll (1989), we decompose the spread into adverse information and order processing cost components. Empirically we estimate these cost components conditional on the relative size of the trade. As predicted by existing theory, we find that the adverse information component of the spread increases significantly and monotonically as trade size increases. This result is robust to several variations in the methodology.

The average order processing cost component is found to decrease monotonically for trades through the 99th percentile of the relative trade size distribution, and then to increase somewhat for the largest 1 percent of all trades. This may reflect extra costs necessary to negotiate large trades even when they do not convey information. Confirming previous research we find a high degree of order persistence, but find a lower probability of a buy (sell) following a large versus small buy (sell) order. The average effective spread is found to increase with trade size. The increase in the effective spread associated with large trades is largely a result of an increase in adverse information revealed from large trades. However, the average effective. spread is less than the average quoted spread, even for very large trades. The results suggest that price concessions (by those demanding immediacy) beyond the quoted bid or ask price for large trades is not a very common practice on the NYSE.

We find strong intradaily patterns in trading volume, relative spreads, and components of the spread for all trade size categories. Trading volume and effective relative spreads are highest at the beginning and end of the day for all trade size categories. However, the adverse information component of the spread declines monotonically during the day for all but the largest 1 percent of all trades. Order processing costs tend to be highest during the middle of the day, while the persistence of order arrival is highest at the end of the day for all trade size categories.

Finally, trades of NYSE securities occurring on a regional exchange or on NASDAQ have markedly different characteristics than those of NYSE trades. Non-NYSE trades convey little or no information regardless of trade size. Average estimates of the adverse information component of the spread do not differ significantly from zero for all trade size categories. This suggests that many non-NYSE trades are negotiated and contain a credible guarantee of being noninformation motivated.

Despite several robustness checks employed in this article, the results should of course be interpreted with caution. The model developed is a one-period horizon model, and, hence will not capture longer-term inventory rebalancing effects. Also, as mentioned, the data employed contain trades executed by many different methods and agents. Our estimates measure the weighted average of the spread components across different mechanisms.

The methods used here could be applied to other financial asset markets, such as options, futures, and over-the-counter markets, where intraday price and volume transaction data are available. Different trading arrangements and information structures might result in interesting comparisons of spread components and, hence, the cost of trading. We leave this issue for future research.

References

Admati, A. R., and P. Pfleiderer, 1988, "A Theory of Intraday Patterns: Volume and Price Variability," *Review of Financial Studies,* 1, 3-40.

Bagehot, Walter (pseud.), 1971, "The Only Game In Town," *Financial Analysts Journal,* 27, 12-14, 22.

Benston, G. J., and R. L. Hagerman, 1974, "Determinants of Bid-Asked Spreads in the Over-the-Counter Market," *Journal of Financial Economics,* 1, 353-364.

Bhide, A., 1993, "The Hidden Costs of Stock Market Liquidity," *Journal of Financial Economics,* 34, 31-52.

Blume, L, D. Easley, and M. O'Hara, 1994, "Market Statistics and Technical Analysis: The Role of Volume," *Journal of Finance,* 49, 153-182.

Brock, W. A., and A. W. Kleidon. 1992, "Periodic Market Closure and Trading Volume: A Model of Intraday Bids and Asks," *Journal of Economic Dynamics and Control*, 16, 451-489.

Chan, L. K. C., and J. Lakonishok, 1993, "Institutional Trades and Intraday Stock Price Behavior," *Journal of Financial Economics,* 33, 173-200.

Choe, H., T. H. McInish, and R. A. Wood, 1991, "Market Microstructure Effect on the Measurement of the Impact of Block Trades," working paper, Pennsylvania State University and Memphis State University.

Choi, J. Y., D. Salandro, and K. Shastri, 1988, "On the Estimation of Bid-Ask Spreads: Theory and Evidence," *Journal of Financial and Quantitative Analysis,* 23, 219-230.

Conroy, R. M., R. S. Harris, and B. A. Benet, 1990, "The Effects of Stock Splits on Bid-Ask Spreads,' *Journal of Finance,* 45, 1285-1295.

Copeland, T., and D. Galai, 1983, "Information Effects and the Bid-Ask Spread," *Journal of Finance,* 38, 1457-1469.

Copeland, T. E., and H. R. Stoll, 1990, "Trading Markets," In Dennis E. Logue (cd.), *Handbook of Modern Finance,* 2nd ed., Warren, Gorham & Lamont, Boston.

Demsetz, H., 1968, "The Cost of Transacting," *Quarterly Journal of Economics,* 82, 33-53.

Easley, D., and M. O'Hara, 1987, "Price, Trade Size, and Information in Securities Markets." *Journal of Financial Economics,* 19, 69-90.

Foster, F. D., and S. Viswanathan, 1993, "Variations in Trading Volume, Return Volatility and Trading Costs: Evidence on Recent Price Formation Models," *Journal of Finance.* 48, 187-212.

George, T. J., G. Kaul, and M. Nimalendran, 1991, "Estimation of the Bid-Ask Spread and Its Components A New Approach," *Review of Financial Studies,* 4, 623-656.

Glosten, L. R., 1989, "Insider Trading, Liquidity, and the Role of the Monopolist Specialist," *Journal of Business,* 62, 211-236.

Glosten, L. R., and L. E. Harris, 1988. "Estimating the Components of the Bid/Ask Spread," *Journal of Financial Economics,* 21, 123-142.

Glosten, L. R., and P. Milgrom, 1985, "Bid, Ask and Transaction Prices in a Specialist Market with Heterogeneously Informed Traders," *Journal of Financial Economics,* 14, 71-100.

Hamilton, J. L., 1976, "Competition, Scale Economies, and Transaction Cost in the Stock Market," *Journal of Financial and Quantitative Analysis,* 11, 779-802.

Harris, L, 1986, "A Transaction Data study of weekly and Intradaily Patterns in Stock Returns," *Journal of Financial Economics,* 16, 99-118.

Hasbrouck, J., 1988. "Trades, Quotes, Inventories and Information," *Journal of Financial Economics,* 22, 229-252.

Hasbrouck, J., 1991a, "Measuring the Information Content of Stock Trades," *Journal of Finance,* 46, 179-207

Hasbrouck, J., 1991b, "The Summary Informativeness of Stock Trades: An Econometric Analysis," *Review of Financial Studies,* 4, 571-595.

Hasbrouck, J., 1992. "Assessing the Quality of a Security Market A New Approach to Transaction Cost Measurement," *Review of Financial Studies,* 6, 191-212.

Hasbrouck, J., and T. S. Y. Ho, 1981. "Order Arrival, Quote Behavior, and the Return-Generating Process," *Journal of Finance,* 42, 1035-1048.

Hasbrouck, J., and R. A. Schwartz, 1988, "Liquidity and Execution costs in Equity Markets," and *of Portfolio Management,* 14, 10-16.

Hasbrouck, J., and G. Sofianos, 1993, "The Trades of Market Makers: An Empirical Analysis of NYSE Specialists," *Journal of Finance,* 48, 155-1593.

Hasbrouck, J., G. Sofianos. and D. Sosebee. 1993, "New York Stock Exchange Systems and Trading Procedures," Working Paper 93-01, New York Stock Exchange.

Ho, T. S. Y., and R Michaely, 1988, "Information Quality and Market Efficiency," *Journal of Financial and Quantitative Analysis,* 23,53-70.

Holthausen, R. W., R. W. Leftwich, and D. Mayers, 1990, "Large Block Transactions, the Speed of Response, and Temporary and Permanent Stock Price Effects," *Journal of Financial Economics,* 26,75-91

Huang, R. D., and H. R. Stoll, 1994, "Market Microstructure and Stock Return Predictions," *Review of Financial Studies,* 7, 179-213.

Jain, P., and G.-H. Joh, 1988, "The Dependence Between Hourly Prices and Trading Volume *Journal of Financial and Quantitative Analysis,* 23,269-284.

Kyle, A. S., 1985, "Continuous Auctions and Insider Trading," *Econometrica,* 53, 1315-1335.

Lee, c. M. C., 1993, "Market Integration and Price Execution for NYSE Securities," *Journal of Finance,* 48, 1009-1038.

Lee, C. M. C., B. Mucklow, and M. J. Rudy, 1993, "Spreads, Depths and the Impact of Earnings Information: An Intraday Analysis," *Review of Financial Studies,* 6, 345-374.

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

Lin, J.C., 1992, "Order Persistence, Adverse Selection, and Gross Profits Earned by NYSE Specialists," working Paper, Louisiana State University.

Lockwood, L., and S. Linn, 1990. "A Examination of Stock Market Return Volatility During overnight and Intraday Periods, 1964-1989," *Journal of Finance,* 45, 591-601.

Madhavan, A., and S. Smidt, 1991, "A Bayesian Model of Intraday Specialist Pricing," *Journal of Financial Economics,* 30, 99-134

McInish, T. H., and R. A. Wood, 1990a. "A Transactions Data Analysis of the Variability of Common Stock Returns During 1980-1984," *Journal of Banking and Finance,* 14, 99-112.

McInish, T. H., and R. A. Wood, 1990b, "An Analysis of Transactions Data for the Toronto Stock Exchange: Returns Patterns and End-of-Day Effect," *Journal of Banking and Finance,* 14, 441-458.

McInish, T. H., and R. A. Wood, 1992, "An Analysis of Intraday Patterns in Bid-Ask Spreads for NYSE Stocks," *Journal of Finance,* 47, 753-764.

New York Stock Exchange, 1991, *New York Stock Exchange Fact Book,* New York.

Schwartz, R. A., 1988, *Equity Markets: Structure, Trading, and Performance,* Harper & Row, New York

Seppi, D. J., 1990, "Equilibrium Block Trading and Asymmetric Information," *Journal of Finance,* Text 45,73-94.

Smidt, S., 1988, "The Bid Ask Spread and the Cost of Transacting," working paper, Johnson Graduate School of Management, Cornell University.

Stoll H. R., 1989, "Inferring the Components of the Bid-Ask Spread: Theory and Empirical Tests," *Journal of Finance,* 44, 115-134.

Subrahmanyam, A., 1991, "Risk Aversion Market Liquidity and Price Efficiency." *Review of Financial Studies,* 4, 417-441.

Tinic, S. M., 1972, "The Economics of Liquidity Services," *Quarterly Journal of Economics,* 86, 79-97.

Wei, P.-H., 1992, "Intraday Variations in Trading Activity, Price Variability and the Bid-Ask Spread, *Journal of Financial Research*, 15, 265-276.

Wood, R. A., T. H. McInish, and J. K. Ord, 1985, "An Investigation of Transactions Data for NYSE Stocks," *Journal of Finance,* 40, 723-739.