Liquidity, Information, and Infrequently Traded Stocks

Author(s): David Easley, Nicholas M. Kiefer, Maureen O'Hara and Joseph B. Paperman
Source: *The Journal of Finance*, Vol. 51, No. 4 (Sep., 1996), pp. 1405-1436
Published by: Wiley for the American Finance Association
Stable URL: http://www.jstor.org/stable/2329399
Accessed: 27-06-2016 06:59 UTC

THE JOURNAL OF FINANCE • VOL. LI, NO. 4 • SEPTEMBER 1996

Liquidity, Information, and Infrequently Traded Stocks

DAVID EASLEY, NICHOLAS M. KIEFER, MAUREEN O'HARA, and JOSEPH B. PAPERMAN*

ABSTRACT

This article investigates whether differences in information-based trading can explain observed differences in spreads for active and infrequently traded stocks. Using a new empirical technique, we estimate the risk of information-based trading for a sample of New York Stock Exchange (NYSE) listed stocks. We use the information in trade data to determine how frequently new information occurs, the composition of trading when it does, and the depth of the market for different volume-decile stocks. Our most important empirical result is that the probability of information-based trading is lower for high volume stocks. Using regressions, we provide evidence of the economic importance of information-based trading on spreads.

DESPITE THE LARGE VOLUMES traded on organized exchanges, many (if not most) listed stocks trade infrequently. On the London Stock Exchange, 50 percent of listed stocks account for only 1.5 percent of trading volume, and over 1000 stocks average less than one trade a day. On the New York Stock Exchange (NYSE), it is common for individual stocks not to trade for days or even weeks at a time, while one stock in London never traded in an eleven-year period. One characteristic of such infrequently-traded stocks is their large bid-ask spreads. In London, spreads for the most active "alpha" stocks average 1 percent, while spreads for the least active "delta" stocks average 11.8 percent.[1] For stocks in

* Easley and Kiefer are from the Department of Economics, Cornell University. O'Hara is from the Johnson Graduate School of Management, Cornell University. Paperman is from the Department of Accounting, University of Washington. We thank an anonymous referee, Yakov Amihud, Joel Hasbrouck, Bruce Lehman, Ananth Madhavan, René Stulz, and seminar participants at Cornell University, the London Business School, Erasmus University, the Western Finance Association Meetings, and the JFI Conference on Market Microstructure, Northwestern University for helpful comments. We thank Colin Moriarity and James Shapiro of the New York Stock Exchange for technical assistance, and the Symposium on Infrequently Traded Stocks held at the London Business School for providing the impetus for this research. Easley and O'Hara gratefully acknowledge research support from Churchill College and the Department of Applied Economics, University of Cambridge, and Kiefer gratefully acknowledges support from the Center for Nonlinear Econometrics, University of Aarhus. This research is supported by National Science Foundation Grant SBR93-20889.

[1] Stocks trading on the London Stock Exchange are divided into four categories, denoted alpha, beta, gamma, and delta. The most active stocks, the alpha stocks, are screen traded by multiple market makers, while the least active stocks generally trade without the benefit of a market maker.

the lower volume deciles on the NYSE, spreads as a percentage of stock price can be 50 percent larger than those of frequently traded stocks.

There are several conjectured explanations for these large spreads. The first involves inventory or liquidity effects. If a stock trades infrequently, the specialist handling the stock may have to maintain an inventory imbalance for a long period. This lack of liquidity may induce a risk averse specialist to set higher spreads to compensate for the exposure.[2] A second explanation is market power. For many inactive stocks, only a single market maker provides liquidity, with few limit order traders willing to post competing orders. This monopoly position may allow the market maker to set larger spreads than would arise in a competitive environment. A third explanation is information-based.[3] Infrequently-traded stocks tend to have greater variability in order flow, with active days interspersed with slow days. If, when shares do trade, it is because of traders acting on private information, then the market maker would face large losses.[4] The large spreads arise, therefore, as the natural consequence of the greater risk of informed trading in illiquid stocks.

The cause of these large spreads is of more than academic interest. The question of how to structure trading in infrequently traded stocks has been widely discussed, with proposals to shift less active stocks to alternative clearing mechanisms debated (and employed) in some markets. In London, for example, the difficulty of trading less active stocks led to the development of the SEATS system, in which a market maker provides an alternative to the screen-based system typically used to trade all but the most active stocks. In Paris, less active stocks recently began trading via morning and afternoon call auctions, replacing the continuous auction used to trade active stocks. On the NYSE, infrequently traded stocks are generally assigned to specialists as part of a portfolio of stocks, with the rationale that the actively traded stocks implicitly subsidize the inactive stocks. Yet, the optimality of any of these trading arrangements is questionable, and much of the confusion stems from our lack of knowledge regarding the differential nature of trading in active versus inactive stocks.

In this article we investigate one aspect of this difference by examining how information-based trading differs between active and inactive stocks. Using a new empirical technique, we estimate the risk of information-based trading for a sample of NYSE stocks. Our analysis uses the information in trade data to estimate the probability of informed trade. This allows us to determine not only whether the probability of informed trading differs across volume deciles,

[2] Predictions of price effects due to inventory are summarized in O'Hara (1995). A related problem is that specialists trading illiquid stocks have fewer trades over which to spread any fixed costs of operation.

[3] The price effects of asymmetric information are analyzed in Kyle (1985), Glosten and Milgrom (1985), and Easley and O'Hara (1987).

[4] Note, however, that it can also be argued that less frequently traded stocks generally face lower risks of information-based trading due to the lack of financial analysts following these stocks. For example, Brennan, Jagadeesh, and Swaminathan (1993) argue that stocks with more financial analysts adjust to information events more quickly than do "neglected" stocks.

but also how the components of informed trading differ. For example, we can determine both how frequently new information (or information events) occurs, and how large a fraction of the order flow is from informed traders when it does. Our estimation approach also allows us to compare the "normal" level of noise trading across volume categories, thereby giving us the ability to assess the depth of the market for different volume-decile stocks.

Our most important empirical result is that the probability of informed trading is lower for high volume stocks. We show that high volume stocks tend to have a higher probability of information events and higher arrival rates of informed traders, but that these are more than offset by the higher arrival rates of uninformed traders. Less active stocks face a greater risk of informed trading, and so their larger spreads are consistent with this information-based explanation. We also show that while high volume stocks differ from medium volume stocks, low and medium volume stocks share many similarities. In particular, our trade-based estimates show that the probability of informed trading does not significantly differ across medium and low volume stocks. This prediction is borne out by spread data, where we show that spreads for low and medium volume stocks do not differ by statistically significant amounts. Using regression results, we also provide evidence of the economic importance of information-based trading on spreads.

From a technical perspective, our estimation of a continuous-time sequential trade model illustrates a new empirical technique for the analysis of problems in finance. Rather than search prices for indirect evidence of informed trading, we *directly* measure the effect of informed trading by estimating the market maker's beliefs. Intuitively, our approach uses the fact that in a market maker's price-setting decision problem, prices are the output, while trades are the input to his learning problem. We use the structure of a continuous time microstructure model to provide the decision rules whereby inferences from the order flow affect beliefs. By analyzing the information in this order flow, we can then measure the extent to which trade flows convey different information for different securities. This trade-based approach, while different in application, complements the work of Hasbrouck (1988; 1991) who examines the information in trade innovations as a vector autoregression.[5]

This article is organized as follows. In the next section, we specify a continuous time sequential trade model, and we develop the likelihood function that we use in our estimation. Section II discusses the data and our sample selection technique. In Section III, we estimate the parameters of our model and calculate the probability of information-based trading for each stock in our sample. In Section IV, we then test the implications of our model by examining actual spread behavior, and provide regression results on the differential effects of volume and information on stock spreads. In Section V, we summa-

[5] Hasbrouck (1991) finds that the persistent price impact of trades is greater for firms with smaller market values than for those with larger market values. Since market values and volume are positively correlated, this result is consistent with our finding of greater risk of informed trade in low volume stocks and thus a greater effect on prices from trades in these stocks.

rize our results and discuss the policy implications of our research for the design of trading mechanisms.

I. The Model

A. Trade Process

In this section, we set out a mixed discrete-and-continuous time, sequential trade model of market making. The model is standard in that trade arises from the actions of a group of potentially informed and uninformed traders, and prices arise from the quotes of a risk neutral, competitive market maker. The model differs from traditional microstructure models in that it explicitly models the arrival rates of traders to the market in a continuous time framework. Because our goal is to estimate empirically the model's parameters, this feature greatly facilitates its estimation for the high-volume stocks of interest in this paper.[6]

Individuals trade a single risky asset and money with a market maker over $i = 1, \ldots, I$ trading days. Within any trading day, time is continuous, and it is indexed by $t \in [0, T]$. The market maker stands ready to buy or sell one unit of the asset at his posted bid and ask prices at any time. Because he is competitive and risk-neutral, these prices are the expected value of the asset conditional on his information at the time of trade.

Prior to the beginning of any trading day, nature determines whether an information event relevant to the value of the asset will occur. Information events are independently distributed and occur with probability α. These events are good news with probability 1-δ, or bad news with probability δ. After the end of trading on any day, and before nature moves again, the full information value of the asset is realized.

Let $(V_i)_{i=1}^{I}$ be the random variables giving the value of the asset at the end of trading days $i = 1, \ldots, I$. These values will naturally be correlated. We do not make any specific assumptions about the correlations as they are not needed for our analysis. We let the value of the asset conditional on good news on day i be $\overline{V}_i$; similarly it is $\underline{V}_i$ conditional on bad news on day i. The value of the asset if no news occurs on day i is denoted V_i^*. We assume, of course, that $\underline{V}_i < V_i^* < \overline{V}_i$.[7]

Trade arises from both informed traders (those who have seen any signal) and uninformed traders. On any day, arrivals of uninformed buyers and uninformed sellers are determined by independent Poisson processes. Unin-

[6] This model is similar to the discrete time trading model developed in Easley, Kiefer, and O'Hara (1993). An important difference is that in this paper trade occurs continuously from a population of potentially asymmetrically informed traders. The discrete time likelihood function developed in our earlier work cannot be computed for data sets with many trades per day. The stocks we examine here include many of the most active stocks on the NYSE and so cannot be analyzed with the earlier approach.

[7] In our empirical work we look at several stocks at once. We assume that the random variables giving asset values are independent across firms. For this reason, the analysis in the text is done for a single firm.

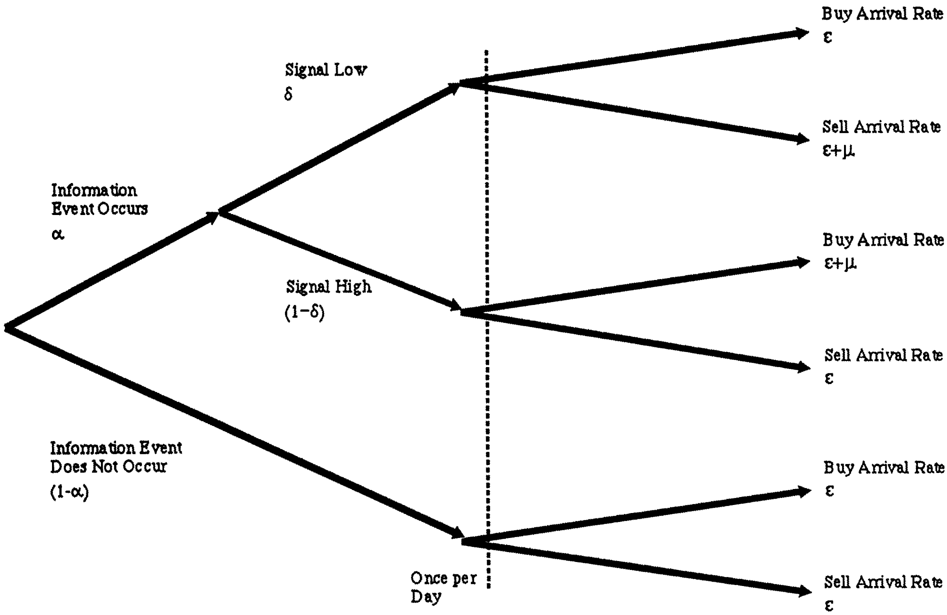

Figure 1. Tree diagram of the trading process. This figure gives the structure of the trading process, where α is the probability of an information event, δ is the probability of a low signal, μ is the rate of informed trade arrival, and ε is the rate of uninformed buy and sell trade arrivals. Nodes to the left of the dotted line occur once per day.

formed buyers and uninformed sellers each arrive at rate ε where this rate is defined per minute of the trading day.[8] On days for which information events have occurred, informed traders also arrive. We assume that all informed traders are risk neutral and competitive. If a trader observes a good signal, then the profit maximizing trade is to buy the stock; conversely, he will sell if he observes a bad signal. We assume that the arrival of news to one trader at a time, and his subsequent arrival at the market, also follows a Poisson process. The arrival rate for this process is μ. All of these arrival process are assumed to be independent.

The tree given in Figure 1 describes this trading process. At the first node of the tree, nature selects whether an information event occurs. If an event occurs, nature then determines if it is good news or bad news. The three nodes (no event, good news, and bad news) before the dotted line in Figure 1 occur only once per day. Then, given the node selected for the day, traders arrive according to the relevant Poisson process. That is, on good event days, the arrival rates are $\varepsilon + \mu$ for buy orders and ε for sell orders. On bad event days,

[8] In a previous version of this article, we allowed uninformed buyers and uninformed sellers to arrive at different rates. In our empirical work we found that typically the arrival rates of uninformed buyers and sellers were not significantly different. None of our conclusions depend on which structure is assumed, so we use the simpler structure.

the arrival rates are ε for buys and $\varepsilon + \mu$ for sells. Finally, on nonevent days, only uninformed traders arrive, and the arrival rate of both buys and sells is ε.

B. Trades and Prices

Each day nature selects one of the three branches of the tree. The market maker knows the probability attached to each branch, and he knows the order arrival process for each of the branches. He does not know, however, which of the three branches has been selected by nature. We assume that the market maker is a Bayesian who uses the arrival of trades, and the rate of trading, to update his beliefs about the occurrence of information events. Because days are independent, we can analyze the evolution of his beliefs separately on each day. Let $P(t) = (P_n(t), P_b(t), P_g(t))$ be the market maker's prior belief about the events "no news" (n), "bad news" (b), and "good news" (g) at time t. So his prior belief at time 0 is $P(0) = (1 - \alpha, \alpha\delta, \alpha(1 - \delta))$.

To determine quotes at time t, the market maker updates his prior conditional on the arrival of an order of the relevant type. For example, the bid at time t, $b(t)$, is the expected value of the asset conditional both on the history of the process prior to the arrival of orders at t (which is captured by the sufficient statistic $P(t)$) and on the fact that someone wants to sell a unit. Let S_t denote the event that a sell order arrives at time t; similarly, B_t is used to represent a buy order at time t. Let $P(t|S_t)$ be the market maker's updated belief vector conditional on the history prior to time t and on the event that a sell order arrives at t.

By Bayes rule, the market maker's posterior probability on no news at time t, if an order to sell arrives at t, is

$$P_n(t|S_t) = \frac{P_n(t)\varepsilon}{\varepsilon + P_b(t)\mu}. \tag{1}$$

Similarly, the posterior probability on bad news is

$$P_b(t|S_t) = \frac{P_b(t)(\varepsilon + \mu)}{\varepsilon + P_b(t)\mu} \tag{2}$$

and the posterior probability on good news is

$$P_g(t|S_t) = \frac{P_g(t)\varepsilon}{\varepsilon + P_b(t)\mu}. \tag{3}$$

At any time t, the zero expected profit bid price, $b(t)$, is the market maker's expected value of the asset conditional on the history prior to t and on S_t. Thus the bid at time t on day i is

$$b(t) = \frac{P_n(t)\varepsilon V_i^* + P_b(t)(\varepsilon + \mu)\underline{V}_i + P_g(t)\varepsilon\overline{V}_i}{\varepsilon + P_b(t)\mu}. \tag{4}$$

Similar calculations show that the ask at time t is

$$a(t) = \frac{P_n(t)\varepsilon V_i^* + P_b(t)\varepsilon \underline{V}_i + P_g(t)(\varepsilon + \mu)\overline{V}_i}{\varepsilon + P_g(t)\mu}. \tag{5}$$

To aid in interpretation of these quotes, it is useful to relate these bid and ask prices to the time t prior expected value of the asset. This expected value of the asset conditional on the history of trade prior to time t is

$$\mathrm{E}[V_i|t] = P_n(t)V_i^* + P_b(t)\underline{V}_i + P_g(t)\overline{V}_i. \tag{6}$$

Substituting equation (6) into the bid and ask equations (4) and (5), respectively, we have

$$b(t) = \mathrm{E}[V_i|t] - \frac{\mu P_b(t)}{\varepsilon + \mu P_b(t)}(\mathrm{E}[V_i|t] - \underline{V}_i) \tag{7}$$

and

$$a(t) = \mathrm{E}[V_i|t] + \frac{\mu P_g(t)}{\varepsilon + \mu P_g(t)}(\overline{V}_i - \mathrm{E}[V_i|t]). \tag{8}$$

These equations demonstrate the explicit role played by arrivals of informed and uninformed traders in affecting trading prices. If there are no informed traders ($\mu = 0$), then trade carries no information, and so the bid and ask are both equal to the prior expected value of the asset. Alternatively, if there are no uninformed traders ($\varepsilon = 0$), then $b(t) = \underline{V}_i$ and $a(t) = \overline{V}_i$ for all t. At these prices no informed traders will trade either, and the market, in effect, shuts down. Generally, both informed and uninformed traders will be in the market, and so the bid is below $\mathrm{E}[V_i|t]$ and the ask is above $\mathrm{E}[V_i|t]$. This spread results from the market maker setting prices to protect himself from losses to informed traders.

The factors influencing the spread are easier to identify if we write the spread explicitly. Let $\Sigma(t) = a(t) - b(t)$ be the spread at time t. Calculation shows that

$$\sum(t) = \frac{\mu P_g(t)}{\varepsilon + \mu P_g(t)}(\overline{V}_i - \mathrm{E}[V_i|t]) + \frac{\mu P_b(t)}{\varepsilon + \mu P_b(t)}(\mathrm{E}[V_i|t] - \underline{V}_i). \tag{9}$$

The spread at time t is the probability that a buy is information-based times the expected loss to an informed buyer, plus a symmetric term for sells.[9] The

[9] Recall that only asymmetric information affects prices in our model, so our spreads also depend only on asymmetric information. Our model thus provides a way to determine information-based differences between spreads. If we cannot reject that information is the same across stock groups, then the differences in spreads must be due to factors other than information such as inventory or market power.

probability that any trade that occurs at time t is information-based is the sum of these probabilities, explicitly

$$\mathrm{PI}(t) = \frac{\mu(1 - P_n(t))}{\mu(1 - P_n(t)) + 2\varepsilon}. \tag{10}$$

This probability depends on the rates of informed and uninformed trading, as well as on the market maker's beliefs regarding the occurrence and composition of information events. So if there is no possibility of news ($P_n(t) = 1$) or if no one trades on private information ($\mu = 0$), then $\mathrm{PI}(t) = 0$ and there is no spread. Alternatively, if all trades are information based ($\varepsilon = 0$), then $\mathrm{PI}(t) =$ 1 and the spread is wide enough ($\overline{V}_i - \underline{V}_i$) to prevent anyone from profiting on private information.

The spread for the opening quotes has a particularly simple form in the natural case in which good and bad events are equally likely. That is, if $\delta = 1 - \delta$ then[10]

$$\sum(0) = \frac{\alpha\mu}{\alpha\mu + 2\varepsilon}\,[\overline{V}_i - \underline{V}_i]. \tag{11}$$

The first term in this equation is the probability that the first trade of the day is information-based. This risk of trading with an informed trader is clearly a crucial factor influencing the size of spreads. If this probability differs between stocks, then our model predicts how initial spreads will differ, and this provides a way to test for information-based differences in spreads.

If, like the market maker, we knew the parameters of the problem, $\theta = (\alpha, \delta, \varepsilon, \mu)$, and observed the order arrival process, then we could compute the stochastic process of bids and asks. This would allow us to directly examine the effect of information on spreads. Although we can observe the order arrival process, we do not know the parameters. These parameters can be estimated, however, from the data on order arrivals. It is to this problem that we now turn.

C. The Likelihood Function

Estimating the parameter vector $\theta = (\alpha, \delta, \varepsilon, \mu)$ is much more complex than just estimating arrival rates from independent Poisson processes. The difficulty arises because we cannot directly observe the arrival of any information events or trades governed by these parameters. Parameters α and δ determine the probabilities of three information events (no news, good news, and bad news), none of which are observable (to us). The remaining parameters refer to arrival rates of uninformed or informed traders. We observe arrivals of orders to buy or sell, but we do not observe which traders are uninformed or informed. Estimation of these parameters thus requires a structural model. Our model

[10] In our empirical work, we find that $\delta = 0.5$ is a good approximation.

provides the structure necessary to extract information on the parameters from the observable variables, buys and sells.

In our model, buys and sells follow one of three Poisson processes on each day. We do not know which process is operating on any day, but we do know that the data reflect the underlying information structure, with more buys expected on days with good events, and more sells on days with bad events.[11] Similarly, on no-event days, there are no informed traders in the market, and so fewer trades arrive. These rates and probabilities are determined by a mixture model in which the weights on the three possible components (i.e., the three branches of the tree reflecting no news, good news, and bad news) reflect their probability of occurrence in the data. The next step in our analysis is to construct this mixture model.

We first consider the likelihood of order arrivals on a day of known type. Suppose we consider the likelihood function on a bad-event day. The sell orders arrive at a rate $(\mu + \varepsilon)$, reflecting that both informed and uninformed traders will be selling. The buy orders arrive at rate ε, since only uninformed traders buy when there has been a bad information event. The exact distribution of these statistics in our model is independent Poisson. Thus, the likelihood of observing any sequence of orders that contains B buys and S sells on a bad-event day of total time T is given by

$$e^{-\varepsilon T} \frac{(\varepsilon T)^B}{B!} e^{-(\mu+\varepsilon)T} \frac{[(\mu + \varepsilon)T]^S}{S!}. \tag{12}$$

Similarly, on a no-event day, the likelihood of observing any sequence of orders that contains B buys and S sells is

$$e^{-\varepsilon T} \frac{(\varepsilon T)^B}{B!} e^{-\varepsilon T} \frac{(\varepsilon T)^S}{S!}. \tag{13}$$

Finally, on a good event day, this likelihood is

$$e^{-(\mu+\varepsilon)T} \frac{[(\mu + \varepsilon)T]^B}{B!} e^{-\varepsilon T} \frac{(\varepsilon T)^S}{S!}. \tag{14}$$

It is evident from equations (12), (13), and (14) that the number of buys and sells (B, S) is a sufficient statistic for the data given T. Thus, to estimate the order arrival rates of the buy and sell processes, we need only consider the total number of buys, B, and the total number of sells, S, on any day.

The likelihood of observing B buys and S sells on a day of unknown type is the weighted average of equations (12), (13), and (14) using the probabilities of each type of day occurring. These probabilities of a no-event day, a bad-event

[11] Note that, unlike in a Kyle (1985) framework, trades in our model are not aggregated, so it is the composition and total number of trades that determines beliefs and, thus, prices.

day, and a good-event day are, respectively, given by $1 - \alpha$, $\alpha\delta$, and $\alpha(1 - \delta)$, and so the likelihood is

$$L((B, S)|\theta) = (1 - \alpha)*e^{-\varepsilon T}\frac{(\varepsilon T)^B}{B!}\, e^{-\varepsilon T}\frac{(\varepsilon T)^S}{S!} + \alpha\delta*e^{-\varepsilon T}\frac{(\varepsilon T)^B}{B!}\, e^{-(\mu+\varepsilon)T}\frac{[(\mu + \varepsilon)T]^S}{S!} + \alpha(1 - \delta)*e^{-(\mu+\varepsilon)T}\frac{[(\mu + \varepsilon)T]^B}{B!}\, e^{-\varepsilon T}\frac{(\varepsilon T)^S}{S!} \quad (15)$$

For any given day, the maximum likelihood estimator of the information event parameters α and δ will be either 0 or 1, reflecting that information events occur only once a day. Over multiple days, however, these parameters can be estimated from the daily numbers of buys and sells. Thus, intra-day data allows us to estimate the trader selection probabilities in our model, and inter-day data to estimate the information event parameters.[12] Because days are independent, the likelihood of observing the data $M = (B_i, S_i)_{i=1}^{I}$ over I days is just the product of the daily likelihoods,[13]

$$L(M|\theta) = \prod_{i=1}^{I} L(\theta|B_i, S_i). \quad (16)$$

To estimate the parameter vector θ from any data set M, we maximize the likelihood defined in equation (16). This provides direct estimates of the rate of informed and uninformed trading in a particular stock, as well as of the information event structure surrounding that stock.

D. Volume, Information-based Trading, and Spreads

Our interest is in determining whether differences in spreads can be explained by differences in the risk of information-based trading. Our model tells us how to estimate these probabilities of informed trading, and predicts how these probabilities, in turn, affect spreads. The hypothesis we wish to test is that differences in spreads across volume deciles are due to differences in the underlying risk of information-based trading.

Our use of a structural model means that testing this hypothesis involves a joint test of the implications of the model and the model itself. This joint test is structured as follows. First, we do statistical tests of our estimated probabilities to determine if they do, in fact, differ based on the stock's general trading activity. This allows us to test the hypothesis that information-based

[12] This is just intended to provide some intuition about how the estimation works. Of course, we actually use the entire data set to determine the joint parameter vector.

[13] Easley, Kiefer, and O'Hara (1993) tested the independence of information events across days, and found that they could not reject the independence assumption.

trading differs across stocks, and to investigate whether informed trading can explain the actual behavior of spreads. Second, we investigate whether our estimated values actually relate to information-based trading by testing the predictive ability of our model. In particular, note that our parameter values are estimated from trade data, but that our spreads are derived from price data. If our trade-based estimates predict correctly the behavior of spreads, then this can be viewed as confirming evidence of the underlying model. We further test the model by regressing spreads on our estimated probabilities of information-based trading.

II. The Data

We now turn to the estimation of our model and, in particular, the determination of the risk of information-based trading for individual stocks. For each stock in our sample, we need to estimate the parameters of the trade process and then determine how these parameters relate to spreads. There are two difficulties that arise in implementing this procedure. First, if a stock trades too infrequently, there may not be sufficient data to reliably estimate the underlying trade process. Second, as our model makes clear, spreads are affected both by trade parameters and by the range of possible trading prices. The relation between spreads and price level need not be monotonic, and failure to adjust for this may introduce a bias in our analysis. We address these difficulties directly in our sample selection criteria.

A. Sample Selection

Our data are for a random sample of stocks that trade on the NYSE. In forming our sample, we eliminate from consideration all preferred stock, stock rights and warrants, stock funds, and ADRs.[14] To address the trade frequency issue raised above, we rank all qualifying stocks by total 1990 trading volume based on data provided by the NYSE. The sample is then divided into deciles based on trading volume, where the first decile contains the most actively traded stocks. Trading volume decreases dramatically across deciles, so to obtain stocks with different volumes, but with enough activity to make estimation possible, we focus on stocks in the first, fifth, and eighth volume deciles.

To eliminate confounding effects due to stock price levels, we construct a matched sample of stocks having the same share price but differing levels of trading volume.[15] Average closing prices from the CRSP (Center for Research in Security Prices) database are calculated for each stock in our volume deciles for the period October 1, 1990 to December 23, 1990. Every stock from the first

[14] This was done to remove any unique securities whose bid-ask spreads might reflect idiosyncratic factors, and not the volume-related properties we seek to investigate.

[15] Notice that simply using percentage spreads does not completely overcome the difficulty because of the possible nonlinear relation of prices and spreads. Our matched sample approach removes this difficulty.

and fifth decile is then ranked in order of average closing price, and adjacent pairs of stocks from different volume deciles are matched. This procedure yields 75 pairs of stocks, from which 30 pairs are randomly selected. We then select the 30 stocks from the eighth decile whose average closing prices are nearest to the matched pairs' prices from the first and fifth deciles. This results in a total sample of 90 stocks.

The list of selected stocks, their average prices, and total 1990 annual volume are provided in the Appendix (see Table A.I.). By construction, volume decreases as we move from the first to the fifth, and then to the eighth deciles. For the first decile, average annual mean volume equals approximately 147 million shares, while for the fifth decile it is 13.8 million shares, and only 3.7 million for the eighth decile. The scale of trading thus differs dramatically across deciles. The average price of the stocks in each decile is approximately 24 dollars, and again, by construction, we cannot reject the hypothesis that the means and variances of the three price distributions in our sample deciles are the same. Average price for the eighth decile is somewhat lower due to a low number of higher priced stocks to match with pairs from the first and fifth deciles.

B. Trade Data

Trade data for the 90 stocks in our sample is taken from the ISSM database for the period October 1 to December 23, 1990. Previous research (Easley, Kiefer, and O'Hara (1993)) has shown that a sixty day trading window is sufficient to allow reasonably precise estimation of the parameters. It is also short enough that the stationarity built into our trade model is not too unreasonable.

To compute the likelihood function given in equation (16), we need the number of buys and sells on each day for each of our stocks. We can determine these numbers by using the ISSM data. First, we know that large trades sometimes have multiple participants on one side of the trade. Reporting conventions may treat such a transaction as multiple trades, when we would want to say that only one trade arrived.[16] To mitigate this problem, all trades occurring within five seconds of each other at the same price, with no intervening quote revisions, are collapsed into one trade. Second, trades are classified into buys and sells using the technique developed by Lee and Ready (1990). Trades at prices above the midpoint of the bid and ask are called buys; those below the midpoint are called sells. The rationale for this classification is that trades originating from buyers are most likely to be executed at or near the ask, while sell orders trade at or near the bid. This scheme classifies all trades except those that occur at the midpoint of the bid and ask. These trades are classified using the "tick test." Trades executed at a price higher than the previous trade are called buys, and those executed at a lower price are called

[16] For a discussion of this timing problem, see Hasbrouck (1988). Combining trades within short intervals (i.e., five seconds) is standard in the literature.

sells. If the trade goes off at the midpoint, and is at the same price as the last trade, then its price is compared to the next most recent trade. This is continued until the trade is classified. This procedure undoubtedly misclassifies some trades, but it is standard, and it has been shown to work reasonably well.

III. Estimation

The next step in our analysis is to estimate the parameters of the structural model. Recall that the trade process depends on four parameters: α, the probability of an information event; δ, the probability that the information is bad news; μ, the arrival rate of traders who know the new information if it exists (i.e., the informed traders); and ε, the arrival rate of uninformed traders. These parameters, in turn, determine the probability of information-based trading in a stock. It is this probability and its relation to spreads that we wish to investigate.

A. Parameter Estimates

We estimate the parameters of the trade process for each stock in our sample by maximizing the likelihood function conditional on the stock's trade data as described in the previous section. The two probability parameters α and δ were restricted to $(0, 1)$ by a logit transform of unrestricted parameters, and the two rate parameters ε and μ were restricted to $(0, \infty)$ by a logarithmic transform. We then maximize over the unrestricted parameters using the quadratic hill-climbing algorithm GRADX from the GQOPT package. Standard errors for the economic parameter estimates are calculated from the asymptotic distribution of the transformed parameters using the delta method.[17]

Parameter estimates and their standard errors for each stock in the sample are provided in the Appendix (see Table A.II.). The standard errors show that the model can be estimated quite precisely. For the parameters as a whole, the arrival rate variables are estimated with great accuracy, reflecting the precision that arises with the large number of transactions in our data set. The information parameters α and δ have larger standard errors, but are still estimated with reasonable precision.

Table I provides the means of our estimated parameters by volume deciles. In the analysis which follows, we examine these mean effects in more detail, but we note at this point that the ranking of our estimates is consistent across deciles, and that the results reveal important differences in the probability of information-based trading. The variability in the estimates, however, suggests that mean effects may conceal important aspects of parameter behavior. A truer picture may emerge, therefore, from examining and testing the cumulative distributions of our estimated variables.

[17] For discussion of the delta method, see Goldberger (1991), p 102.

Table I

Summary Parameter Estimate Statistics by Decile

This table presents means, medians, and sample standard deviations of parameter estimates by volume decile for the 90 stocks in our sample. The parameter μ is the arrival rate of informed traders, ε is the arrival rate of uninformed traders, α is the probability of an information event, and δ is the probability that new information is bad news. The parameter Prob (Inf) is a composite variable measuring the probability of information-based trade.

Parameter	First Decile	Fifth Decile	Eighth Decile
Number in Sample	30	30	30
μ			
Mean	0.131970	0.030148	0.015696
Median	0.104864	0.027596	0.014122
Std. dev.	0.079314	0.013238	0.008607
ε			
Mean	0.175742	0.023970	0.009614
Median	0.136797	0.022917	0.008925
Std. dev.	0.141192	0.013158	0.005093
α			
Mean	0.500294	0.433952	0.356320
Median	0.477761	0.448613	0.363841
Std. dev.	0.141192	0.170253	0.173540
δ			
Mean	0.349078	0.444393	0.501787
Median	0.360357	0.418164	0.455418
Std. dev.	0.227188	0.238763	0.318183
Prob(Inf)			
Mean	0.163919	0.207788	0.220245
Median	0.154193	0.205858	0.196712
Std. dev.	0.043794	0.064794	0.121155

To compare these distributions, we use nonparametric statistics, specifically the Kruskal-Wallis test and the Mann-Whitney test (also called the Wilcoxon rank sum test).[18] The Kruskal-Wallis test determines whether the three population distribution functions are identical. In particular, we can test whether one of the three populations differs from the other populations. These test statistics are given in Table II, Panel A. The Wilcoxon test allows us to compare two samples with a directional hypothesis. That is, we can test whether the values for one sample tend to be higher or lower than for the second sample. These test statistics are given in Table II, Panel B.

We first consider the estimates of the information event parameter, α, which is the probability of an information event occurring before the start of a trading

[18] It would be difficult to use classical statistics to perform tests on the distributions of our parameters. The restriction to non-negative numbers for our rate parameters μ and ε, and the restriction to (0, 1) for the probability parameters α and δ, obviously violate the normality required for most standard statistical tests. Because we have no basis for making assumptions about the distributions of our parameters, deriving specific parametric tests is also problematic. There are methods for building empirical distributions, such as bootstrapping, but our limited sample size does not allow us to use these methods.

Table II

Nonparametric Tests

The Kruskal-Wallis statistic is used to test the null hypothesis that parameter values for all three volume samples are drawn from identical populations versus the alternative hypothesis that at least one of the populations tends to furnish greater observed values than other populations. The parameter μ is the arrival rate of informed traders, ε is the arrival rate of uninformed traders, α is the probability of an information event, and δ is the probability that new information is bad news. The parameter Prob (Inf) is a composite variable measuring the probability of information-based trade.

Panel A: Kruskal-Wallis Tests on Parameters

Parameter	Test Statistic
μ	66.279
ε	69.859
α	10.853
δ	4.236
Prob(Inf)	8.027

Critical value for $\alpha = 0.05$ is 5.991.

The Mann-Whitney statistic is used to test the null hypothesis that two samples are drawn from identical populations against the alternative that one population tends to yield higher values. The parameter μ is the arrival rate of informed traders, ε is the arrival rate of uninformed traders, α is the probability of an information event, and δ is the probability that new information is bad news. The parameter Prob (Inf) is a composite variable measuring the probability of information-based trade.

Panel B: Mann-Whitney Tests on Parameters

	Pairwise Comparisons (n = 30, m = 30)		
Parameter	1 to 5	1 to 8	5 to 8
μ	6.402	6.623	4.480
ε	6.505	6.653	5.071
α	1.390	3.326	1.789
δ	−1.508	−1.937	−0.547
Prob(Inf)	−2.883	−1.952	0.192

The test statistic is normally distributed and the critical value for $\alpha = 0.05$ is ± 1.6449.

day. Table I shows that the mean α is 0.500 for decile 1, 0.434 for decile 5, and 0.356 for decile 8. Thus, the probability of information events is highest for our active stocks, and declines for our less active stocks. A similar, but more complex pattern, is revealed by the cumulative distributions of α. The distributions of α tend to differ across volume deciles, with that of the most active stocks higher than that of the medium volume stocks, and generally higher still than for the low volume stocks.[19]

[19] One interesting feature of our estimates is their variability. For the most active stocks, α ranges from a low of 0.21 to a high of 0.72. Similarly, for the least active stocks, α ranges from 0.14

We would like to know the statistical properties of these differences, and, in particular, if the distribution of α for our active stocks lies above that for our inactive stocks. The Kruskal-Wallis test in Table II shows that it does, with a test statistic of 10.853, well above the 0.05 confidence level of 5.991. The hypothesis that the three distributions are the same is thus strongly rejected. The Mann-Whitney test shows that low volume stocks have a significantly lower probability of information events than do high or medium volume stocks. However, the difference between medium and high volume stocks is not as significant.[20]

The second information parameter in our model is δ, which is the probability that new information is bad news. There is no theoretical reason to expect the probability of bad news to differ by volume deciles, and hence the estimation of this parameter provides a simple check on the reasonableness of our model. Table I shows that the means for the three deciles are similar, being 0.349 for the first decile, 0.444 for the fifth decile, and 0.502 for the eighth decile. The cumulative distributions of δ exhibit multiple crossings, suggesting that there are no significant differences in this probability between deciles. The Kruskal-Wallis test confirms this, with a test statistic of 4.236 (well below the 0.05 cutoff level of 5.991 needed to reject the hypothesis of no difference in the three distributions). Thus, as expected, there is no significant evidence of differences in the direction of information across volume deciles.

We now consider the parameters relating to the arrival rates of uninformed and informed traders. The arrival rate of uninformed traders is our estimated parameter ε, while the informed arrival rate is our estimated parameter μ. Table I shows dramatic differences in uninformed arrival rates across volume deciles. For uninformed traders, the rate falls from 0.176 for decile 1 to 0.024 for decile 5, and to 0.010 for decile 8. Comparing distributions of these variables (see Table II), the Kruskal-Wallis test soundly rejects that the distributions are identical (test statistic 69.9), while the Mann-Whitney test reveals the expected result that the ε distribution for decile 1 stocks differs significantly from that for decile 5 or decile 8 stocks. These tests also reveal that the distributions of uninformed trader arrival rates also differ between deciles 5 and 8, with the arrival rate significantly lower for the least active stocks.

The behavior of informed order arrivals exhibits similar behavior. The estimated mean value of μ is 0.132 for decile 1, 0.030 for decile 5, and it is 0.016 for decile 8 stocks. Thus, the informed arrival rate is higher for more active stocks. The Kruskal-Wallis tests and the Wilcoxon tests in Table II strongly confirm the rank ordering of decile 1 exceeding decile 5 and decile 8, and decile

to 0.70. Hence, there are infrequently traded stocks that often have new information, and there are actively traded stocks for which very little new occurs.

[20] The hypotheses that the α distributions for deciles 1 and 8, or deciles 5 and 8, are the same are rejected at the 0.05 level. The hypothesis that the α distribution for deciles 1 and 5 are the same can be rejected at the 0.10 level, but not at the 0.05 level.

5 exceeding decile 8. Thus, higher volume stocks have higher arrival rates of informed traders.

These results demonstrate that liquid stocks have higher arrival rates of uninformed traders and higher arrival rates of informed traders. These higher arrival rates are consistent with the stocks' higher volume, but they do not necessarily explain observed spread behavior. That is, one might expect that a higher rate of informed trading would lead to higher spreads, and not to lower ones. This reasoning, however, misses the complex link that exists between spreads and information. What matters for the spread is the *overall* risk of information-based trading, and, as our model demonstrates, this depends on the interaction of our information event and arrival rate probabilities. Having estimated the parameter values, we can now determine this probability of informed trade for the stocks in our sample, and examine how it differs between frequently and infrequently traded stocks.

B. The Probability of Informed Trade

The probability of informed trade is a composite variable reflecting the various parameters characterizing the trade process. This probability is given by our model in equation (10), where we showed that the probability of informed trade is

$$\mathrm{PI} = \frac{\alpha\mu}{\alpha\mu + 2\varepsilon} \tag{17}$$

for the market maker's initial beliefs. As is apparent from equation (17), the probability of information-based trading depends on the arrival rates of traders (both informed and uninformed) and on the probability that new information exists. Consequently, it is the interaction of our estimated parameters that matters for determining the effect of information-based trading on spreads.

We calculated this probability for each stock in our sample. The mean values are reported in Table I (individual estimates are given in the last column of Table A.II. in the Appendix). The data reveal an intriguing result: The risk of informed trade is clearly *lowest* for the active stocks. The stocks in our first volume decile have on average lower probabilities of informed trade than do the stocks in less active deciles. The mean results show that the probability of information-based trades is approximately 0.164 for the stocks in our most active sample, but it rises to 0.208 for the fifth decile stocks, and to 0.220 for the eighth decile.

Figure 2 depicts the cumulative distributions of PI for each volume decile. The Figure shows a crossing point between the first decile distribution and the others, but overall the decile 1 cumulative distribution lies to the left of the distributions for deciles 5 and 8. The Kruskal-Wallis statistic strongly rejects the hypothesis that the distributions are the same. The test statistic is 8.027, which is above the 0.05 critical level of 5.991. Pairwise testing using the

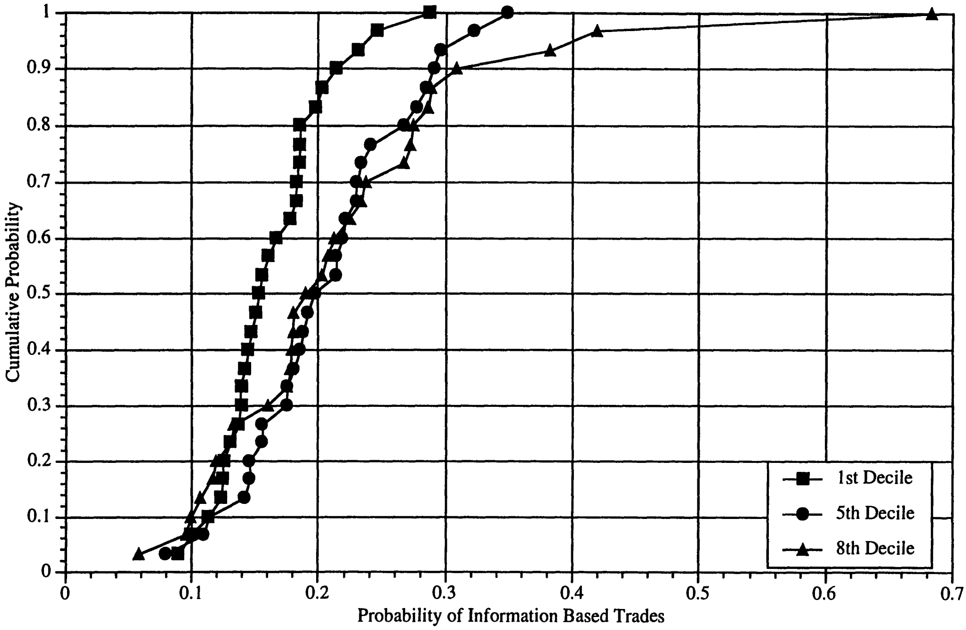

Figure 2. Cumulative distribution of the probability of information-based trades by volume decile. This graph shows the cumulative probability distribution of the probability that a trade comes from an informed trader, P(Inf|Trade), for each of the three 1990 volume deciles in our sample.

Wilcoxon rank sum test reveals that we can reject that the fifth decile lies below the first decile, and similarly for the relation of the first and eighth deciles. All of these test statistics are given in Table II.

What may be equally important is what our results do not show. Our estimates reveal that the probability of informed trading does *not* differ between stocks in the fifth and eighth deciles. The Mann-Whitney test statistic of 0.1920 is clearly insignificant, dictating that the risk of informed trading is the same across both deciles. The multiple crossings of the distributions in Figure 2 also vividly show the virtually identical behavior of the probability of informed trade for the fifth and eighth volume deciles.

These results provide strong evidence that the differential behavior of spreads across volume deciles can be at least partially explained by asymmetric information. Our estimated probabilities of informed trade show that, for the stocks in our sample, the risk of informed trading is lower for active stocks, and that it is essentially the same for medium and low volume stocks. Given these estimates, our model's pricing equations yield two predictions: First, the spreads for active stocks will be lower than spreads for the less active stocks. Second, our model predicts that spreads of less frequently traded stocks will be the same. Thus, we expect to find no difference in the spreads between decile 5 stocks and decile 8 stocks.

Table III

Spread Summary Statistics by Volume Decile

Descriptive statistics are presented in this table for the 90 stocks included in our sample reported by 1990 volume decile. Average spread is the time weighted mean of the New York Stock Exchange (NYSE) quoted spread from the Institute for the Study of Security Markets (ISSM) database for the period of 10/1/90 to 12/23/90. Average percentage spread is the time-weighted mean of the NYSE quoted spread divided by the midpoint of the quote from the ISSM database for the period of 10/1/90 to 12/23/90.

	First Decile	Fifth Decile	Eighth Decile
Number in sample	30	30	30
Average spread			
Mean	0.1763	0.2549	0.2708
Median	0.1717	0.2581	0.2802
Std. dev.	0.0243	0.0588	0.0585
Average % spread			
Mean	1.4140	1.9158	1.9824
Median	0.7123	1.1446	1.1688
Std. dev.	1.6961	1.9379	2.0211

IV. Spreads and Information-Based Trading

We now turn to testing the economic significance and validity of our model's predictions. The predictions of our model can be tested in two ways. First, we examine actual spread behavior for the 90 stocks in our three volume deciles. This provides a direct test of the model's predictions regarding differential spread behavior across volume deciles due to asymmetric information. As noted in the Introduction, however, spreads may also be influenced by factors such as inventory and market power. While our model does not incorporate these factors directly, we can test how well our information-based estimates do in explaining overall spread behavior. This provides a de facto test on the economic validity of our model. To investigate this, we use regression analysis to consider how well our estimated variables do in predicting spread behavior.

We first consider the relation of spreads and volume for the stocks in our sample. Table III provides summary statistics on spreads and percentage spreads by volume deciles. The data reveal the expected result that the mean average spread decreases with trading activity. The average first decile spread is 0.18, and this increases to 0.25 for the fifth decile, and to 0.27 for the eighth. Statistical testing shows that we can reject the hypothesis that the first decile has the same mean average-spread as the fifth and eighth decile. We cannot reject the hypothesis that the mean-average spreads of the fifth and eighth deciles are the same.[21]

[21] Testing that the first and either fifth or eighth deciles have the same mean average-spread is complicated because we can reject the hypothesis that they have the same variance in average-spread. Both large sample tests and small sample, normally distributed tests, however, suggest that these mean average-spreads are different. Ignoring the difference in variances, the t values for the hypotheses that the mean average-spreads are the same for the first and fifth deciles and

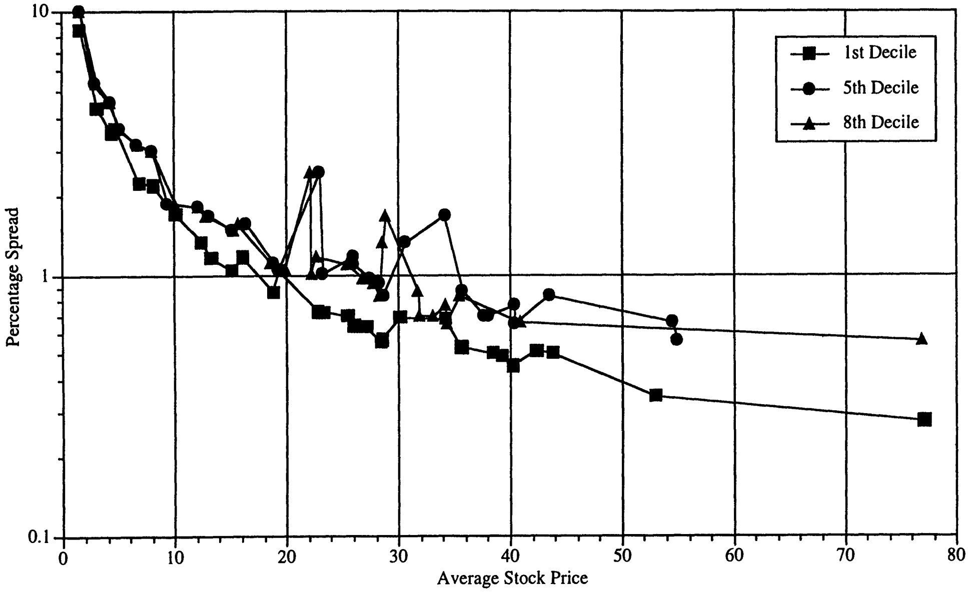

Figure 3. Percentage spread and average stock price by volume decile. This graph shows the relationship between average stock price and average quoted spread as a percentage of the spread midpoint for each of the three 1990 volume deciles in our sample. Average price is calculated as the mean of the CRSP daily file closing price for 10/1/90 to 12/23/90. Average percentage spread is the average of the opening quoted spread divided by the midpoint of the bid/ask quote for the same period as given by the ISSM database.

Better information about the distribution of percentage spreads can be obtained from Figure 3, which provides a plot of percentage spread against price for each decile. The curve for the first decile is below that for the fifth decile at all points in our sample. The curve for the fifth decile is generally, but not always, below that for the eighth decile.[22] To compare the percentage spread curves, we use the Wilcoxon signed ranks test. Here it is appropriate to compare percentage spreads for our matched pairs of stocks.

We strongly reject the null hypothesis that percentage spreads tend to be lower in the fifth decile than in the first decile (the test statistic is 4.741, well above the 0.05 critical level of 1.6449). We similarly reject the hypothesis that percentage spreads tend to be lower in the eighth decile than in the first decile

for the first and eighth deciles are 5.18 and 6.25, respectively. We cannot reject the hypothesis that the fifth and eighth deciles have the same variance in average-spread, and so, using t-tests, we cannot reject the hypothesis that the fifth and eighth deciles have the same mean average-spread with a t value of 0.55.

[22] The data also reveal that regardless of the volume decile, lower price stocks have higher average percentage spreads. This effect is largely concentrated for stock prices under twenty dollars, as above that level spreads appear to approximately level off. That high volume stocks face such a nonlinear spread is surprising, in that it suggests that many traders pay large costs to transact in those securities. Since firms can, to a large extent, influence the level of their stock price, this also raises the question of why firms allow such costs to persist.

(the test statistic is 4.576). Again, however, we cannot reject the hypothesis that percentage spreads in the eighth decile are lower than in the fifth decile (the test statistic is 1.471).[23]

These data provide strong evidence in support of our asymmetric information explanation for spread behavior. In particular, our model predicts that due to their lower estimated risk of information-based trading, the spreads of first decile stocks will be below those of fifth and eighth decile stocks. This prediction is strongly supported by the data. Second, our model predicts that because our estimated risk of informed trading is the same for stocks in the fifth and eighth deciles, their spreads should also be the same. This, too, is supported by the data. That spread behavior is not monotonically related to liquidity is a new, and we believe, unique finding of this study. These results suggest that differences in information-based trading play a major role in explaining differences in spread behavior between active and inactive stocks.

Our analysis thus far has focused on relating the stock-specific estimates of our model to spread behavior. As we have shown, these results are statistically significant, and appear to predict well the behavior of actual spreads. An additional test is to consider how well our estimated variables do in predicting overall spread behavior. In particular, our model provides an estimate of the probability of information-based trading in each stock. If our estimate correctly identifies this variable, then the estimated parameter should affect spreads in predictable ways. This can be investigated via regression analysis. Such testing provides a simple de facto check on the validity of our estimation approach, and it allows us to determine the economic significance of our results.

To investigate these influences, we return to the opening spread derived in equation (11). The simplifying approximation used in that equation of equally likely good news and bad news is consistent with our empirical findings so we can write the opening spread on day i as[24]

$$\sum = [\overline{V}_i - \underline{V}_i]\mathrm{PI}. \tag{18}$$

The bracketed term is the price range of the asset, while the second term is the probability that the opening trade comes from an informed trader. If we assume that the price range is a linear function of the stock price, denoted V, then our spread can be re-expressed as

$$\sum = \beta_1 \cdot V \cdot \mathrm{PI}, \tag{19}$$

[23] We also investigate the behavior of opening percentage spreads. Again, the hypotheses that the first decile spreads are greater than those in either the fifth or eighth decile are strongly rejected. The test statistic comparing the fifth and eighth decile percentage spreads falls to 0.319, suggesting that there is no statistically significant difference between spreads in these two deciles.

[24] Of course, all of the variables in this equation are stock-specific. In this equation, and in the regressions to follow, asset values and parameters differ across stocks and it is this variability that we are exploring. These values should all be viewed as indexed by stock. We have not included the index in the text in order to keep the notation simple.

where β_1 is the constant in our linear relationship.[25] This gives the intuitive relationship that spreads depend positively on the probability of informed trade.

Of course, the spread in any stock may be influenced by other factors not in our model. For example, average daily dollar volume may also affect spreads if inventory effects matter.[26] Our model considers only asymmetric information, so we have no explicit prediction about this effect. Intuitively, however, the inventory cost to the market maker should be positively related to the distance a trade takes the market maker from his desired inventory investment, and negatively related to dollar volume. This negative relation arises because large trading volumes allow the market maker to move back to desired inventory levels more quickly.[27]

This discussion suggests the estimating equation

$$\sum = \beta_0 + \beta_1 \cdot V \cdot \mathrm{PI} + \beta_2 \cdot \mathrm{Vol} + \eta \tag{20}$$

where Vol is average daily dollar volume, η is the error term, and β_0 is a constant. The constant term is added because the model ignores any costs the market maker incurs outside of losses to informed traders. Since a competitive market maker will have to recoup any fixed costs of operation from traders, the spread will be higher than that predicted by the model, and this is accounted for in the regressions by the constant term β_0. If our model accurately estimates the probability of informed trade, we would expect β_1 to be positive. Although we do not have a model of inventory effects, we would expect β_2 to be negative.

We ran this regression over the 90 stocks in our sample. The average spread is calculated from an average of the daily opening spread for each stock during the sample period of October 1–December 23, 1990.[28] The stock price, V, is the CRSP end-of-day price averaged over the sample period. The parameter values for the probability of informed trade, PI, are our estimated values given in the Appendix. The volume variable, Vol, is average daily dollar volume, computed as the average daily number of shares traded times the midpoint of the opening spread, over the sample period.

The results for our estimation are given in the first column of Table IV. Perhaps the first thing to note is that all coefficients for our estimated vari-

[25] The implication of this assumption is that the standard deviation of information-based prices, $(\overline{V}_i, \underline{V}_i)$ for asset i, is linear in the stock price. Alternative structures could be tested, but we see no reason to prefer any particular alternative.

[26] We also consider average daily volume rather than dollar volume. Either variable leads to similar conclusions.

[27] Of course, it would be better to estimate an inventory model that was derived from a closed form solution to a risk-averse market maker's decision problem. We do not have such a model.

[28] We use opening spreads for consistency with our theoretical model, which predicts how such spreads will differ with respect to our estimated information-based trading probability. Spreads throughout the day will change in response to the specific order flow and hence make comparisons more difficult.

Table IV

Regression Results

This table presents the results from estimating the linear regression given by $\Sigma = \beta_0 + \beta_1 \cdot V \cdot PI + \beta_2 \cdot \text{Vol} + \eta$. The dependent variable Σ is the average quoted opening spread calculated from information on the Institute for the Study of Security Markets (ISSM) database for the period of 10/1/90 to 12/23/90. The average price, V, is obtained by averaging the Center for Research in Security Prices (CRSP) daily prices for this same time period. The dollar volume, Vol, is the average daily number of shares traded times V. The probability of informed trade, PI, is reported in the Appendix. The regressions use Ordinary Least Squares (OLS); t-statistics are reported in parenthesis.

	General Model	Restriction to $\beta_1 = 0$	Restriction to $\beta_2 = 0$
Intercept	0.2114	0.2885	0.2034
	(20.453)	(31.954)	(18.030)
V*PI	0.0193		0.0178
	(9.463)		(7.982)
Vol	−1.035E-11	−6.879E-12	
	(−4.572)	(−2.175)	
Adj. R^2	0.5216	0.0402	0.4134
F-Value	49.518	4.730	63.720

ables have the predicted signs.[29] Most important for us is that the coefficient on the probability of informed trade is positive and statistically significant, dictating that the greater the probability of informed trade, the larger are spreads. This regression explains a significant portion of the variance of spreads with an adjusted R^2 of 52.16 percent and an associated F-Value of 49.5.

It is interesting to explore the explanatory power of each variable in our regression individually. We first examine the importance of volume alone in determining spreads by restricting the coefficient on V*PI to equal 0. These results are given in the second column of Table IV. Now the dollar volume enters negatively as expected, but it is of marginal significance. Moreover, the R^2 of this restricted regression is only 4.02 percent, suggesting that volume alone does not have much explanatory power. We next consider the importance of our information variable when the coefficient on volume is restricted to 0. These results are given in the third column of Table IV. The probability of informed trading again has a positive coefficient and it is strongly significant. The R^2 of the regression is 41.34 percent with an F-Value of 63.7.[30] As this regression is directly suggested by our model, we interpret these results as

[29] The specific coefficient estimates should be interpreted with caution, however, since the regressors are stochastic. The error in variables will bias the reported results for the restricted regressions toward coefficient values and F-values of zero. The direction of bias for the general regression is indeterminate depending on the correlation of the errors for the two independent variables. The level of bias is proportional to the standard error of noise divided by the standard error in the true value of the underlying independent variables.

[30] We also explore the value of price in predicting spread behavior. Our regression results suggest that price adds little explanatory value beyond that conveyed by PI times V.

strong evidence in favor of our approach. At least within our sample, the probability of informed trade is a better predictor of spread than is volume. This regression evidence, combined with our earlier analysis, suggests that differences in spreads between volume deciles are at least partially explained by differences in information-based trading.

V. Conclusions and Policy Implications

For a sample of NYSE stocks, we have investigated the differential behavior of active and infrequently traded stocks. Using a new empirical technique, we use trade data to estimate the probability of information-based trade for each stock in our sample. Our analysis reveals that the risk of information-based trading is lower for active stocks than it is for infrequently traded securities. We also demonstrate that the risk of information-based trading does not differ between our medium and low volume stocks, yielding the prediction that spreads for these stocks should also not differ. We then test the predictions of our model using price data, and found strong support for our model.

Our results provide a number of insights into market behavior. Our finding of higher information-based trading in low volume stocks suggests that the large spreads in such stocks are not merely the result of market power by market makers, or difficulties in risk-bearing due to inventory. Less active stocks are riskier because they are subject to more information-based trading, a result consistent with Amihud and Mendleson's (1986) finding that average risk-adjusted returns increased significantly with bid-ask spreads.[31] We conjecture that differences in the probability of informed trading may explain other anomalies found in the literature, and we plan to investigate this in future research.

One implication of our results is that private information is more important for infrequently traded stocks. Although information events happen more rarely in these stocks, when new information occurs it has a greater impact on trading. Such trading impacts are also investigated in a vector-autoregression by Hasbrouck (1991). He concludes that, for stocks with small market values, trade innovations had greater persistent price impacts, which he interprets as arising from greater informational asymmetries. Our work explains this finding by directly showing that low volume/small capitalization stocks have a higher probability of informed trade. The greater price effects found by Hasbrouck are thus a natural response to the greater risk of information-based trading in such stocks.

One intriguing implication of the empirical results concerns the relation between the composition of trade and liquidity. Although high volume stocks

[31] See also Brennan and Subrahmanyam (1994), who examine the link between the price impact of trades and expected returns using a Fama-French factor model. An issue raised by this line of research is why this information risk is not diversifiable. One possible explanation is that stocks must be bought and sold individually, meaning that one cannot diversify away the bid-ask spread. A similar effect arises with respect to taxes, as diversification there can also not remove tax liabilities for individual securities.

tend to have (marginally) higher probabilities of information events and higher arrival rates of informed traders, these are more than offset by the higher arrival rate of uninformed traders. This result highlights the crucial role of market depth, which is usually defined as the size or scale of noninformation-linked trading. From the perspective of the market maker, the less active stocks are riskier, since there is a higher probability that any trade comes from an informed trader. The problem with less active stocks, therefore, is not that there are too many informed traders, but that there are too few uninformed ones.

From a policy perspective, this result may explain the almost universal failure of screen trading for inactive stocks. Screen trading relies on limit orders to provide liquidity, but placing a limit order is equivalent to writing a "free option" (see O'Hara (1995) for discussion of the free option problem). The greater the risk of informed trading, the more valuable is this option, and consequently traders demand greater compensation (i.e., a wider spread) to place a limit order. Such wide spreads can curtail the rate of uninformed trading, however, further exacerbating the underlying problem. The recent switch by the London Stock Exchange from screen trading to a market maker-assisted system is a response to this problem, and our results suggest that it should reduce the trading costs for inactive stocks.

Pagano and Roell (1996) suggest that trading costs can also be reduced by increasing the transparency of the trading mechanism. Their analysis concludes that greater transparency reduces the ability of informed traders to profit from their information, and thus reduces the losses of uninformed traders. Our finding of greater information-based trading in inactive stocks suggests that, at least for these stocks, greater transparency would be desirable. One way to achieve this is to employ different trading mechanisms for active and inactive stocks. The newly introduced Tradepoint electronic trading system in London features just such an approach. While active stocks will trade continuously throughout the trading day, inactive stocks will trade periodically through call auctions.[32] Our findings here suggest that switching the trading of illiquid stocks to such a call market, as has also recently been done on the Paris Bourse, should result in an improvement of overall trader welfare.[33]

[32] Tradepoint features an "instant auction" mechanism for active stocks and a "periodic auction" for less frequently traded stocks. Although both systems involve electronic order clearing, the rules underlying the two mechanisms differ. Initial market reaction has been particularly enthusiastic regarding the periodic auction mechanism, reflecting perhaps the greater benefits that may arise in the trading of these less active stocks.

[33] A second development in Paris for the trading of less active stocks are market makers employed by the issuing firm. These market makers are paid by the issuing firm and are restricted by company policy to set fixed bid-ask spreads. Given that these inactive stocks are already subject to greater informed trading and that the firm may not have the optimal incentives from the perspective of uninformed traders, this trading approach seems unlikely to yield much reduction in trading costs.

APPENDIX

Table A.I.

NYSE Stocks Included in Sample

This table presents the stocks included in the testing of our model. Thirty pairs of stocks from the first (highest) 1990 volume decile and the fifth volume decile were selected matched on stock price. An additional sample of thirty stocks from the eighth volume decile with similar prices were added. Average price is calculated as the mean CRSP daily closing price from 10/1/90 to 12/23/90. Total stock volume for 1990 was provided by the NYSE.

Ticker	Company Name	CUSIP	Average Price	1990 Volume
	Panel A: First (Highest) Volume Decile Stocks			
CCI	Citicorp	17303410	13.24	409,462,200
WX	Westinghouse Electric Corp.	96040210	27.16	311,279,900
GTE	G T E Corp.	36232010	28.46	298,377,500
F	Ford Motor Co	34537010	28.54	276,229,000
AMD	Advanced Micro Devices Inc.	00790310	4.39	262,823,800
EK	Eastman Kodak Co	27746110	40.20	227,904,900
TOY	Toys R Us	89233510	22.84	204,891,100
UIS	Unisys Corp.	90921410	3.11	199,081,000
PFE	Pfizer Inc.	71708110	77.04	179,367,300
ADM	Archer Daniels Midland Co	03948310	23.25	178,436,550
S	Sears Roebuck & Co	81238710	25.47	163,519,300
AN	Amoco Corp.	03190510	53.00	139,615,500
CHL	Chemical Banking Corp.	16372210	12.38	131,502,700
PN	Pan Am Corp.	69775710	1.52	123,274,200
P	Phillips Petroleum Co	71850710	25.99	121,590,000
DI	Dresser Industries Inc.	26159710	19.34	117,323,200
MXS	Maxus Energy Corp.	57773010	10.09	112,338,700
AL	Alcan Aluminum Ltd.	01371610	18.77	102,497,700
BUD	Anheuser Busch Cos. Inc.	03522910	39.29	102,059,300
PDG	Placer Dome Inc.	72590610	15.14	101,317,310
GT	Goodyear Tire & Rubber Co	38255010	16.03	98,489,100
GPS	Gap Inc.	36476010	34.16	92,811,000
BMG	Battle Mountain Gold Co	07159310	6.84	89,663,700
BR	Burlington Resources Inc.	12201410	42.35	87,124,600
PCI	Paramount Communications Inc.	69921610	38.40	85,805,400
DGN	Data General Corp.	23768810	4.72	84,421,800
CBU	Commodore International Ltd.	20266010	8.06	84,319,400
AET	Aetna Life & Casualty Co	00814010	35.62	81,330,700
BN	Borden Inc.	09959910	30.12	79,998,600
DE	Deere & Co	24419910	43.73	79,562,000
	First Volume Decile Mean Values		24.98	154,213,915
	Panel B: Fifth Volume Decile Stocks			
MUR	Murphy Oil Corp.	62671710	40.35	16,000,300
CF	C F & I Steel Corp.	12518510	15.16	15,826,500
APM	Applied Magnetics Corp.	03821310	7.98	15,564,200
G	Greyhound Dial Corp.	25247010	23.23	15,462,200

Table A.I.—*Continued*

Ticker	Company Name	CUSIP	Average Price	1990 Volume
	Panel B—*Continued*			
SFD	Smiths Food & Drug Ctrs Inc.	83238810	27.53	15,348,300
WOA	Worldcorp Inc.	98190410	4.25	15,157,000
LNC	Lincoln National Corp. In	53418710	38.00	14,820,100
UCU	Utilicorp United Inc.	91800510	19.23	14,701,800
BNL	Beneficial Corp.	08172110	40.41	14,514,100
FMC	F M C Corp.	30249130	28.24	14,338,700
AME	Ametek Inc.	03110510	9.38	14,029,100
HAD	Hadson Corp.	40501810	1.60	14,003,500
FQA	Fuqua Industries Inc.	36102810	12.10	13,884,600
ATM	Anthem Electronics Inc.	03673210	16.43	13,700,500
SCG	Scana Corp.	80589810	34.15	13,649,800
FLO	Flowers Industries Inc.	34349610	13.10	13,528,600
FSS	Firth Sterling Inc.	33799190	28.68	13,417,500
CUM	Cummins Engine Inc.	23102110	35.71	13,400,500
SNG	Southern New England Telecom	84348510	30.68	13,389,600
CCK	Crown Cork & Seal Inc.	22825510	54.94	13,066,800
MCL	Moore Corp. Ltd.	61578510	22.95	13,042,900
MAI	Microwave Association Inc.	55261810	5.11	12,261,900
OGE	Oklahoma Gas & Elec Co	67885810	37.61	12,018,600
LOC	Loctite Corp.	54013710	54.51	11,859,000
RGS	Rochester Gas & Elec. Corp.	77136710	18.84	11,704,600
CPY	C P I Corp.	12590210	26.01	11,640,300
EFU	Eastern Gas & Fuel Assoc.	27637F10	25.98	11,635,200
GAL	Galoob Lewis Toys Inc.	36409110	2.86	11,620,600
RNB	Republic New York Corp.	76071910	43.48	11,426,800
RLC	Rollins Truck Leasing Corp.	77574110	6.67	10,979,700
	Fifth Volume Decile Mean Values		24.17	13,533,110
	Panel C: Eighth Volume Decile Stocks			
CZM	Calmat Co	13127110	22.63	4,669,000
HB	Hillenbrand Ind. Inc.	43157310	35.49	4,585,000
LMS	Lamson & Sessions Co	51369610	4.30	4,555,900
SEE	Sealed Air Corp.	81211510	22.21	4,519,700
AJG	Gallagher Arthur J & Co	36357610	22.13	4,395,700
AVA	Audio Video Affiliates Inc.	05090310	3.07	4,381,300
FED	Firstfed Financial Corp.	33790710	15.63	4,324,900
AGL	Angelica Corp.	03466310	28.85	4,319,700
UIC	United Industrial Corp.	91067110	8.01	4,227,200
RXN	Rexene Corp.	76168210	1.59	4,112,400
SAR	Santa Anita Realty Enterprises	80120920	25.47	3,973,800
PNY	Piedmont Natural Gas Inc.	72018610	28.35	3,768,200
BDG	Bandag Inc.	05981510	76.88	3,670,000
CER	Cilcorp Inc.	17179410	33.10	3,619,800
CYC	Cyclops Industries Inc.	23252810	12.06	3,571,000
DSO	De Soto Inc.	25059510	34.23	3,557,100
NC	Nacco Industries Inc.	62957910	28.60	3,481,300

Table A.I.—*Continued*

Ticker	Company Name	CUSIP	Average Price	1990 Volume
		Panel C—*Continued*		
LOG	Rayonier Timberlands L P	75507810	19.87	3,400,600
BRY	Berry Petroleum Co	08578910	15.29	3,331,100
ESL	Esterline Technologies Corp.	29742510	6.80	3,317,000
SWN	Southwestern Energy Co	84546710	31.72	3,236,600
WMK	Weis Markets Inc.	94884910	27.79	3,216,000
CRS	Carpenter Technology Corp.	14428510	40.89	3,167,400
PEO	Petroleum & Resources Corp.	71654910	26.81	3,144,000
MSA	Medusa Corp.	58507230	12.78	3,129,300
CMI	Club Med Inc.	18947010	18.54	3,107,300
CRI	Core Industries Inc.	21867510	5.05	3,060,900
CNL	Central Louisiana Elec. Inc.	15389760	34.29	3,056,200
TII	Thomas Industries Inc.	88442510	10.10	3,015,500
CES	Commonwealth Energy Sys	20280010	31.91	2,989,600
	Eighth Volume Decile Mean Values		22.82	3,696,783

Table A.II.

Continuous Time Trading Model Parameter Estimates

This table presents the parameters estimated using our model. The parameter μ is the arrival rate of informed traders, ε is the arrival rate of uninformed traders, α is the probability of an information event, and δ is the probability that new information is bad news. The parameter Prob (Inf) is a composite variable measuring the probability of information-based trade. Maximum likelihood estimation is performed using the hill-climbing algorithm GRADX in the GQOPT statistical package. Standard errors are given parenthesis below the parameter estimates.

Ticker	μ	ε	α	δ	Prob(Inf)
	Panel A: Estimates for stocks in the first (highest) 1990 volume decile				
ADM	0.148779	0.253751	0.477802	0.191793	0.122862
	(0.006977)	(0.002852)	(0.067681)	(0.077792)	(0.015503)
AET	0.101146	0.164674	0.414596	0.373757	0.112945
	(0.006695)	(0.002421)	(0.074097)	(0.103345)	(0.017136)
AL	0.055838	0.066926	0.536405	0.503276	0.182851
	(0.003862)	(0.001621)	(0.078417)	(0.096772)	(0.021197)
AMD	0.083882	0.074381	0.532888	0.284067	0.231051
	(0.004368)	(0.001741)	(0.073270)	(0.083662)	(0.023656)
AN	0.093385	0.192306	0.583953	0.340066	0.124179
	(0.005508)	(0.002626)	(0.071341)	(0.083955)	(0.013724)
BMG	0.106477	0.082259	0.248948	0.159629	0.138763
	(0.007582)	(0.001623)	(0.061586)	(0.107645)	(0.028324)
BN	0.069492	0.103059	0.477719	0.512118	0.138719
	(0.004889)	(0.001957)	(0.077220)	(0.104378)	(0.018720)
BR	0.184790	0.115503	0.215878	0.599507	0.147258
	(0.008006)	(0.001675)	(0.053243)	(0.140098)	(0.031358)
BUD	0.063244	0.130105	0.681873	0.362643	0.142167
	(0.004567)	(0.003001)	(0.107731)	(0.086296)	(0.017277)
CBU	0.070096	0.042113	0.326844	0.399000	0.213843
	(0.004836)	(0.001148)	(0.066218)	(0.117989)	(0.033069)

Table A.II.—*Continued*

Ticker	μ	ε	α	δ	Prob(Inf)
			Panel A—*Continued*		
CCI	0.224647	0.416855	0.627419	0.358072	0.144613
	(0.007790)	(0.003859)	(0.064891)	(0.079827)	(0.013105)
CHL	0.114256	0.130489	0.456736	0.584995	0.166639
	(0.005624)	(0.002045)	(0.067981)	(0.096391)	(0.020758)
DE	0.103251	0.167461	0.359186	0.133000	0.099692
	(0.007699)	(0.002416)	(0.073557)	(0.079636)	(0.017213)
DGN	0.033160	0.032406	0.637148	0.486356	0.245841
	(0.002614)	(0.001248)	(0.088598)	(0.094280)	(0.024250)
DI	0.142132	0.143106	0.436768	0.056801	0.178239
	(0.006148)	(0.002082)	(0.066102)	(0.049985)	(0.022343)
EK	0.163630	0.339637	0.410295	0.570004	0.089946
	(0.008890)	(0.003320)	(0.069020)	(0.104349)	(0.013605)
F	0.370661	0.539220	0.660926	0.251728	0.185111
	(0.009082)	(0.004557)	(0.062850)	(0.069386)	(0.014483)
GPS	0.100973	0.097646	0.440241	0.135349	0.185416
	(0.005220)	(0.001773)	(0.068166)	(0.070289)	(0.023345)
GT	0.102726	0.109990	0.305960	0.113204	0.125014
	(0.006685)	(0.001806)	(0.064026)	(0.078136)	(0.022522)
GTE	0.157482	0.287883	0.658892	0.060782	0.152699
	(0.006490)	(0.003388)	(0.067006)	(0.040651)	(0.013382)
MXS	0.044475	0.063784	0.707127	0.719619	0.197774
	(0.003235)	(0.001740)	(0.083287)	(0.082570)	(0.018451)
P	0.064405	0.122475	0.727368	0.571204	0.160543
	(0.004081)	(0.002377)	(0.074478)	(0.081241)	(0.014595)
PCI	0.172666	0.201344	0.520275	0.128743	0.182396
	(0.006534)	(0.002590)	(0.067050)	(0.060567)	(0.019336)
PDG	0.070785	0.082074	0.410499	0.422886	0.150395
	(0.005077)	(0.001768)	(0.076262)	(0.108068)	(0.022452)
PFE	0.150390	0.252900	0.504745	0.373575	0.130493
	(0.007980)	(0.003501)	(0.079167)	(0.091643)	(0.016483)
PN	0.092726	0.074329	0.647703	0.790634	0.287752
	(0.004096)	(0.001865)	(0.071175)	(0.066883)	(0.021673)
S	0.127108	0.202135	0.721529	0.774888	0.184910
	(0.005659)	(0.003185)	(0.071523)	(0.067537)	(0.014379)
TOY	0.210267	0.257082	0.450901	0.037199	0.155687
	(0.010518)	(0.003540)	(0.074983)	(0.036645)	(0.020086)
UIS	0.348571	0.265456	0.388478	0.177433	0.203223
	(0.009896)	(0.002853)	(0.063946)	(0.080395)	(0.026597)
WX	0.187642	0.260912	0.439855	0.000000	0.136567
	(0.007424)	(0.002752)	(0.065155)	(0.000000)	(0.017789)
		Panel B: Estimates for stocks in the fifth 1990 volume decile			
AME	0.029402	0.024354	0.506179	0.419513	0.234039
	(0.002725)	(0.001089)	(0.095184)	(0.104744)	(0.030015)
APM	0.025787	0.014025	0.417028	0.000000	0.277132
	(0.002383)	(0.000652)	(0.071217)	(0.000057)	(0.034987)

Table A.II.—*Continued*

Ticker	μ	ε	α	δ	Prob(Inf)
			Panel B—*Continued*		
ATM	0.052518	0.040572	0.284167	0.449828	0.155347
	(0.004603)	(0.001110)	(0.065827)	(0.131807)	(0.029452)
BNL	0.043709	0.043676	0.545516	0.343334	0.214430
	(0.003097)	(0.001297)	(0.076468)	(0.090289)	(0.023658)
CCK	0.058222	0.041817	0.305313	0.229402	0.175288
	(0.004914)	(0.001147)	(0.067969)	(0.102680)	(0.030568)
CF	0.027573	0.011499	0.142340	0.639729	0.145778
	(0.008942)	(0.000804)	(0.098103)	(0.199258)	(0.055551)
CPY	0.013817	0.011787	0.508369	0.375314	0.229564
	(0.002229)	(0.000798)	(0.138583)	(0.113197)	(0.037616)
CUM	0.031420	0.022955	0.338264	0.190158	0.187980
	(0.003751)	(0.000936)	(0.085893)	(0.100719)	(0.033280)
EFU	0.023795	0.022978	0.703337	0.372136	0.266959
	(0.002209)	(0.001086)	(0.098518)	(0.087797)	(0.025048)
FLO	0.028229	0.027037	0.244766	0.117384	0.113304
	(0.004682)	(0.000964)	(0.084943)	(0.106972)	(0.029055)
FMC	0.045509	0.039111	0.548037	0.355236	0.241759
	(0.003117)	(0.001264)	(0.078239)	(0.091828)	(0.025307)
FQA	0.015822	0.011630	0.401657	0.793062	0.214579
	(0.002356)	(0.000676)	(0.099281)	(0.104370)	(0.036678)
FSS	0.022570	0.022879	0.373503	0.738963	0.155566
	(0.003284)	(0.001041)	(0.112061)	(0.127847)	(0.032415)
G	0.046466	0.064144	0.584984	0.517879	0.174838
	(0.003845)	(0.001925)	(0.103898)	(0.094247)	(0.022158)
GAL	0.027618	0.015445	0.445887	0.380282	0.285021
	(0.002532)	(0.000745)	(0.078816)	(0.101621)	(0.034295)
HAD	0.018629	0.012351	0.630779	0.850933	0.322356
	(0.002030)	(0.000829)	(0.115617)	(0.066679)	(0.031617)
LNC	0.045059	0.046114	0.451339	0.432510	0.180669
	(0.004417)	(0.001568)	(0.101129)	(0.108096)	(0.027197)
LOC	0.021677	0.017982	0.282821	0.322317	0.145646
	(0.003624)	(0.000831)	(0.095430)	(0.162858)	(0.034491)
MAI	0.032574	0.023394	0.589245	0.566179	0.290894
	(0.002534)	(0.001075)	(0.088740)	(0.092545)	(0.028108)
MCL	0.044228	0.025374	0.614056	0.416815	0.348609
	(0.002560)	(0.000985)	(0.070494)	(0.085297)	(0.026391)
MUR	0.022857	0.016454	0.409346	0.471536	0.221373
	(0.002511)	(0.000810)	(0.086472)	(0.117411)	(0.034113)
OGE	0.017847	0.022168	0.590342	0.738052	0.192012
	(0.002873)	(0.001204)	(0.173176)	(0.093683)	(0.032225)
RGS	0.010020	0.011324	0.678250	1.000000	0.230823
	(0.001762)	(0.000678)	(0.126991)	(0.000153)	(0.029868)
RLC	0.010532	0.006155	0.490846	0.623701	0.295762
	(0.001744)	(0.000550)	(0.126452)	(0.122875)	(0.042877)
RNB	0.037702	0.016259	0.242665	0.147423	0.219575
	(0.003757)	(0.000686)	(0.061669)	(0.097743)	(0.042448)
SCG	0.016704	0.025553	0.506819	0.608889	0.142114
	(0.003146)	(0.001587)	(0.226950)	(0.127879)	(0.041086)
SFD	0.023238	0.015634	0.166123	0.382899	0.109894
	(0.005213)	(0.000774)	(0.083777)	(0.218721)	(0.038823)

Table A.II.—*Continued*

Ticker	μ	ε	α	δ	Prob(Inf)
			Panel B—*Continued*		
SNG	0.021666	0.028524	0.647124	0.501842	0.197286
	(0.002693)	(0.001432)	(0.153140)	(0.102556)	(0.029576)
UCU	0.037625	0.027666	0.335320	0.346465	0.185674
	(0.003623)	(0.000976)	(0.074444)	(0.119245)	(0.031631)
WOA	0.020530	0.008922	0.215313	0.185430	0.198536
	(0.003911)	(0.000633)	(0.089562)	(0.137560)	(0.051319)
		Panel C: Estimates for stocks in the eighth 1990 volume decile			
AGL	0.008984	0.007071	0.371322	0.839158	0.190866
	(0.002375)	(0.000630)	(0.176682)	(0.126145)	(0.051378)
AJG	0.003373	0.002284	0.508222	0.859637	0.272857
	(0.001469)	(0.000423)	(0.268244)	(0.205619)	(0.079340)
AVA	0.006915	0.001047	0.655093	0.914514	0.683882
	(0.001015)	(0.000243)	(0.119221)	(0.059288)	(0.033041)
BDG	0.019207	0.015862	0.357461	0.185846	0.177920
	(0.003076)	(0.000849)	(0.111782)	(0.111367)	(0.036493)
BRY	0.019557	0.006102	0.118937	0.458319	0.160084
	(0.004892)	(0.000454)	(0.064227)	(0.218618)	(0.058368)
CER	0.006649	0.007369	0.598488	0.999987	0.212585
	(0.002046)	(0.000554)	(0.217940)	(0.003486)	(0.037743)
CES	0.016155	0.006841	0.052492	0.999996	0.058362
	(0.005867)	(0.000402)	(0.037332)	(0.001942)	(0.036167)
CMI	0.018267	0.007654	0.088339	0.000003	0.095357
	(0.006407)	(0.000487)	(0.064408)	(0.001863)	(0.047321)
CNL	0.010147	0.009515	0.247215	0.147481	0.116469
	(0.003969)	(0.000842)	(0.214888)	(0.276367)	(0.060781)
CRI	0.021944	0.012886	0.430298	0.139395	0.268144
	(0.002755)	(0.000805)	(0.105895)	(0.084559)	(0.038335)
CRS	0.016012	0.009876	0.383833	0.599071	0.237325
	(0.002374)	(0.000655)	(0.103588)	(0.128806)	(0.040743)
CYC	0.006167	0.003663	0.482166	0.572868	0.288693
	(0.001444)	(0.000475)	(0.156464)	(0.145203)	(0.056613)
CZM	0.036994	0.012867	0.094519	0.815155	0.119622
	(0.006382)	(0.000560)	(0.042299)	(0.169005)	(0.045069)
DSO	0.023084	0.014339	0.497891	0.261871	0.286116
	(0.002138)	(0.000738)	(0.080834)	(0.089775)	(0.032657)
ESL	0.014076	0.008713	0.358960	0.508345	0.224778
	(0.002440)	(0.000647)	(0.116908)	(0.132725)	(0.044568)
FED	0.014168	0.004793	0.177081	0.276019	0.207444
	(0.002739)	(0.000388)	(0.066411)	(0.185759)	(0.056666)
HB	0.022139	0.016068	0.368722	0.134139	0.202560
	(0.002660)	(0.000768)	(0.083731)	(0.100751)	(0.034396)
LMS	0.012340	0.006469	0.755898	0.999999	0.418926
	(0.001392)	(0.000533)	(0.098626)	(0.000578)	(0.028315)
LOG	0.014946	0.013405	0.382175	0.194302	0.175633
	(0.004380)	(0.001024)	(0.217059)	(0.112920)	(0.047907)
MSA	0.010725	0.005812	0.410451	0.452518	0.274695
	(0.001995)	(0.000604)	(0.143870)	(0.134393)	(0.051662)
NC	0.040740	0.025552	0.187955	0.258682	0.130312
	(0.009710)	(0.001223)	(0.100471)	(0.157354)	(0.041366)

Table A.II.—*Continued*

Ticker	μ	ε	α	δ	Prob(Inf)
		Panel C—*Continued*			
PEO	0.013508	0.013358	0.303271	0.368338	0.132954
	(0.003054)	(0.000753)	(0.120786)	(0.166132)	(0.037590)
PNY	0.008161	0.007880	0.231668	0.779228	0.107116
	(0.003949)	(0.000642)	(0.202171)	(0.217208)	(0.054113)
RXN	0.008280	0.003786	0.566835	0.999997	0.382688
	(0.001340)	(0.000376)	(0.105314)	(0.001377)	(0.039856)
SAR	0.022476	0.015654	0.303035	0.200276	0.178678
	(0.002881)	(0.000740)	(0.078139)	(0.133841)	(0.035823)
SEE	0.022236	0.009137	0.181370	0.394095	0.180788
	(0.003593)	(0.000525)	(0.063362)	(0.183518)	(0.047317)
SWN	0.009960	0.006042	0.541575	0.541263	0.308601
	(0.001631)	(0.000585)	(0.138887)	(0.116292)	(0.043014)
TII	0.012854	0.010443	0.356582	0.665751	0.179960
	(0.002453)	(0.000700)	(0.122209)	(0.147001)	(0.041801)
UIC	0.007262	0.012682	0.385458	0.087074	0.099395
	(0.003799)	(0.000847)	(0.276004)	(0.210632)	(0.044350)
WMK	0.023548	0.011237	0.292375	0.400265	0.234503
	(0.002886)	(0.000645)	(0.077556)	(0.145157)	(0.043101)

REFERENCES

Amihud, Y., and H. Mendleson, 1986, Asset pricing and the bid-ask spread, *Journal of Financial Economics* 17, 223–249.

Brennan, M. J., Jegadeesh, N., and B. Swaminathan, 1993, Investment analysis and the adjustment of stock prices to common information, *Review of Financial Studies* 6, 799–824.

Brennan, M. J., and A. Subrahmanyam, 1994, Market microstructure and asset pricing: On the compensation for adverse selection in stock returns, Working paper, UCLA.

Easley, D., and M. O'Hara, 1987, Price, trade size, and information in securities markets, *Journal of Financial Economics* 19, 69–90.

Easley, D., Kiefer, N., and M. O'Hara, 1993, One day in the life of a very common stock, Working paper, Cornell University.

Glosten, L., and P. Milgrom, 1985, Bid, ask, and transaction prices in a specialist market with heterogeneously informed traders, *Journal of Financial Economics* 13, 71–100.

Goldberger, A. S., 1991, *A Course in Econometrics* (Harvard University Press, Cambridge, Mass.).

Hasbrouck, J., 1988, Trades, quotes, inventory, and information, *Journal of Financial Economics* 22, 229–252.

Hasbrouck, J., 1991, Measuring the information content of stock trades, *Journal of Finance* 46, 179–207.

Kyle, A., 1985, Continuous auctions and insider trading, *Econometrica* 53, 1315–1336.

Lee, C., and M. Ready, 1991, Inferring trade direction from intraday data, *Journal of Finance* 46, 733–746.

O'Hara, M., *Market Microstructure Theory*, 1995, (Blackwell Publishers, Cambridge, Mass.).

Pagano, M., and A. Roell, 1996, Transparency and liquidity: A comparison of auction and dealer markets with informed trading, *Journal of Finance*, 51, 579–612.