Who's Informed? An Analysis of Stock Ownership and Informed Trading [1]

Patrick J. Dennis
McIntire School of Commerce, University of Virginia
Monroe Hall
Charlottesville, VA 22903
Phone: 804-924-4050
E-mail: pjd9v@virginia.edu

James P. Weston
Jones Graduate School of Management, Rice University
6100 Main Street, Houston, TX 77005
Phone: 713-348-4480
E-mail: westonj@rice.edu

This version: March 2001

[1]This version of the paper is preliminary. Please do not quote without permission. Comments are very welcome. Address e-mail correspondence to Patrick Dennis at pjd9v@virginia.edu or James Weston at westonj@rice.edu.

Who's Informed? An Analysis of Stock Ownership and Informed Trading

Abstract

We examine the relationship between ownership structure and informed trading. We attempt to reconcile some puzzling results in recent empirical literature about the impact of ownership on informed trading by using a comprehensive set of proxies for informed trading and a recent sample of firms from all three exchanges. As proxies for informed trading, we use four measures: (1) The relative spread, (2) the adverse selection component of the spread, constructed as in Huang and Stoll (1997), (3) the price impact of a trade, constructed as in Foster and Viswanathan (1993) and Hasbrouck (1991), and (4) the probability of informed trading constructed as in Easley, Kiefer, O'Hara and Paperman (1996). We relate these measures of information to cross-sectional differences in firm ownership composition. We control for effects such as size, volume, volatility and price, and we account for the simultaneous effect that institutional ownership and the measure of informed trading can have on each other. We find that both institutional ownership and ownership by insiders are positively related to the intensity of information-based trading. Our results are robust to a variety of estimation techniques and proxies for informed trading. Overall, our results suggest that individual investors are less informed relative to institutions and insiders. These findings are consistent with economies of scale in information acquisition and aggregation, and recent research that indicates that market makers move prices in response to trades by institutions.

1 Introduction

This paper empirically examines the relationship between ownership structure and informed trading. It is known that institutions trade more frequently than individual investors.[1] One explanation for this is that institutions are better informed than individuals and that the more frequent trading is based on that information. On the other hand, institutions may simply churn their clients accounts. Since portfolio managers earn a living through active management, they may have an incentive to trade even if they have no private information.[2] In addition to institutions, insiders may also have an impact on informed trading. In contrast to institutions who must invest scarce resources to acquire private information, insiders may have private information readily available. In this study we relate the information content of trading to the structure of both institutional and insider ownership. We find both types of ownership to be positively related to the information content of trades.

Compared to individuals, institutions have greater resources for gathering and processing information. Institutions may exploit economies of scale in information acquisition creating an informational advantage that leads to more frequent trading. Since their marginal costs of gathering and processing information are lower than for individuals, this may reduce the information asymmetry for firms with a large percentage of institutional ownership. Supporting this hypothesis, Szewczyk, Tsetsekos and Varma (1992) find that firms with relatively high levels of institutional ownership have a smaller price reaction to the announcement of equity offerings. Alangar, Bathala and Rao (1999) find an identical result for dividend change announcements. Finally, Bartov, Radhakrishnan, and Krinsky (2000) find that firms with high levels of institutional ownership have lower levels of post-earnings announcement drift in the stock price. Overall, these studies are consistent with the notion that institutions are better informed than other owners.

[1] See, for example, Gompers and Metrick (2001).
[2] See, for example, Dow and Gorton (1997) or Trueman (1988).

While the existing research on the microstructure effects of ownership structure tends to focus on liquidity, it does provide some evidence about the relationship between the information content of trades and ownership structure. Sarin, Shastri and Shastri (2000) examine a sample of NYSE/AMEX stocks and find that higher institutional ownership results in wider bid-ask spreads and generally lower adverse selection costs.[3] Jennings, Schnatterly and Sequin (1997) examine a sample of NASDAQ stocks and find that higher institutional ownership results in narrower bid-ask spreads and lower adverse selection costs.[4] These results about adverse selection costs are puzzling given recent findings in Sugato (2001). He uses the TORQ database and finds evidence that market makers have some indication about which trades are initiated by institutions and move prices more for these trades than for others, indicating that market makers believe that institutions are better informed.

If institutions and insiders are better informed than individuals, we expect the information content of trades to be larger, not smaller, for stocks that have a relatively high percentage of institutional ownership. We attempt to resolve this puzzle by using a more comprehensive set of proxies based on recent theoretical advances for the information content of trades. Specifically, we use the adverse selection component of the spread (Huang and Stoll (1997)), the price impact of a trade (Foster and Viswanathan (1993) and Hasbrouck (1991)) and the probability of informed trading (Easley, Kiefer, O'Hara and Paperman (1996)). We then relate each of these measures to cross-sectional patterns in ownership structure. Futhermore, we include a sampling of firms from all three exchanges, and use a more recent and larger sample of data.[5] Contrary to existing research, we find that informed trading is directly related to the percentage of the firm's shares

[3]They extract the adverse selection component of the spread following George, Kaul and Nimalendran (1991), Glosten and Harris (1988) and Madhaven and Smidt (1991). The George, Kaul and Nimalendran measure is not related to institutional ownership, while the Glosten and Harris and Madhaven and Smidt measures are negatively related to institutional ownership.

[4]They extract the adverse selection component of the bid-ask spread following George, Kaul, and Nimalendran (1991).

[5]Sarin, Shastri and Shastri (2000) use an 8-month long sample from 1985, we use five quarters from 1997 and 1998.

held by institutions and insiders. Our findings are consistent with the evidence in Sugato (2001) and the hypothesis that institutions and insiders are better informed than individual investors. Further, our results suggest that the positive relationship between share turnover and institutional ownership may be driven by informed trading rather than churning.

The layout of the paper is as follows. Section 2 describes the construction of the measures of the informativeness of trading, Section 3 describes the data, Section 4 discusses the univariate results, Section 5 describes the multivariate results, Section 6 discusses the differences by exchange type, and Section 7 concludes the paper.

2 Methodology

In order to test the relationship between ownership structure and informed trading, we must first estimate the amount of information-based trading. However, distinguishing informative trades from liquidity trades poses some serious difficulties. Ideally, we would identify trades executed for institutions or insiders and compare the impact of those trades to retail trades executed for individuals. Unfortunately, since no data exist that classify the identity of traders,[6] we must rely on proxies for the intensity of information-based trading. In this section, we outline the five different measures of informed trading that we use in the paper. Each of these measures relies, in some way, on the sensitivity of price changes to the direction and/or the size of the order flow or order imbalance. Our five measures of information-based trading are:

1. The quoted bid-ask spread: S

   This is the most primitive measure of information-based trading. Theoretical models of Kyle (1985), Glosten and Harris (1988) and Amihud and Mendelson (1986) argue that the spread

[6] With the exception of the TORQ database, which only covers the time period from November 1990 to January 1991.

is a function of the amount of information asymmetry in a market. In these models, financial intermediaries face a standard "lemons-market" problem. Given that informed investors and liquidity traders are not easily distinguishable, intermediaries drive a wedge between the bid and ask price. This spread enables dealers or specialists to earn a profit from liquidity traders which compensates them for providing liquidity to informed traders. The size of the spread, therefore, is a proxy for the proportion of informed traders in the market.

2. The adverse selection component of the bid-ask spread: HS

   The quoted bid-ask spread measures not only the amount of information asymmetry, but also the fixed cost of making a market and the order processing costs. In order to measure the percentage of the spread that stems solely from information costs, we estimate the components of the spread using the empirical model of Huang and Stoll (1997), which is a generalization of Glosten and Harris' (1988) trade indicator model. They derive a simple model that allows a one-step decomposition of the information component as a percentage of the spread. The remaining spread stems from order processing costs and market maker rents.

   The model identifies these components by measuring how prices change as a function of the direction of the last trade. We define an indicator variable, Q_t, which takes on the values $\{-1, 0, 1\}$ based on the direction of the last trade. That is, define P_t and M_t as the transaction price and midpoint of the spread at time t, respectively. Q_t is then defined as: $Q_t = -1$ if $P_t < M_t$ (indicates a sell order) ; $Q_t = 0$ if $P_t = M_t$; $Q_t = 1$ if $P_t > M_t$ (indicates a buy order). S is the quoted spread and ε_t represents the random (iid) public information shock at time t. The regression equation is then specified by:

$$\Delta P_t = \frac{S}{2}\Delta Q_t + \alpha\frac{S}{2}Q_{t-1} + \varepsilon_t$$

   where α measures the proportion of the half spread, S, that stems from information costs. The remaining proportion of the spread $(1 - \alpha)$ is due to order processing costs and market

maker rents. This specification is slightly different from Huang and Stoll (1997) who assume inventory costs are also captured by α. That is, we assume that inventory costs, as a proportion of the spread, are equal to zero. To the extent that inventory costs do exist, they will be captured in our estimate of α. However recent empirical evidence suggests that inventory costs are likely to be close to zero.[7]

To understand the intuition behind this model, consider the limiting cases. If $\alpha = 0$ then previous trades provide no information. As a result, there should be no reason for the midpoint of the spread to change. In this case, orders simply bounce between a fixed bid and ask as the true value of the security follows a martingale sequence. On the other hand, if $\alpha = 1$, then the last trade signals to the dealer that the trade was fully informative. As a result, the market maker moves the midpoint of the spread to the last transaction price. That is, the dealer moves the spread to straddle the last bid (following a sell order) or ask (following a buy order). For value of α between 0 and 1, the amount by which the midpoint of the spread moves in reaction to the last trade measures the amount of the spread attributable to this component. We call this measure HS.

3. The adverse selection cost: $HS2$

   This measure is a simple transformation of the adverse selection component of the of the spread from percentage terms into dollar costs. To construct this measure, we simply multiply our estimate of α by the average spread S. This measure is therefore just the product of our first two measures and is given by:

$$HS2 = \alpha * S.$$

4. The Hasbrouck-Foster-Visvanathan price impact of a trade: HFV

[7] See Madhavan and Sofianos (1998).

Our fourth measure is based on the models of Foster and Viswanathan (1993) and Hasbrouck (1991). Given that orders are often serially correlated (i.e. buy orders tend to follow buy orders; sells follow sells), some portion of the order flow may be expected. The VAR model of Hasbrouck (1991) accounts for this by allowing the informativeness of a trade to depend only on the unexpected portion of the order flow.

Let V_t be the signed trade volume corresponding to the time t price change ΔP_t, and let Q_t be the trade indicator variable described above. Following Brennan and Subramanyam (1996), we estimate the following model using five lags:

$$\begin{aligned} V_t &= \theta_0 + \sum_{i=1}^{5} \beta_i \Delta P_{t-i} + \sum_{i=1}^{5} \gamma_i V_{t-i} + \tau_t \\ \Delta P_t &= \phi_0 + \phi_1 \Delta Q_t + HFV \tau_t + \varepsilon_t. \end{aligned}$$

The first regression in this framework finds the portion of the signed order flow, V_t, that cannot be explained by past order flow or price changes. We use the residuals from this regression, τ, as a proxy for the unexpected portion of the order flow in the second regression on price changes. In this setup, HFV (the coefficient on the unexpected order flow) measures the impact of a trade on future price changes in a similar manner to α. However, this measure accounts for serial correlation in order flow.

5. The probability of informed trading: PI

Our final measure of information costs differs from the previous four. Rather that estimate the size of the spread or the sensitivity of prices to order flow, we estimate the probability of informed trading using the sequential trade model of Easley, Kiefer, O'Hara and Paperman (EKOP) (1996). The EKOP model relies on the total number of buy and sell trades during a day to identify informed trading. In this model, information events may occur only before

the start of trading on each day.[8] The probability of an information event occurring is given by α. Given that an information event has occurred, the probability of *bad* news occurs with probability δ while good news has probability $(1-\delta)$. In this model, there are two types of stylized traders: informed traders who know the true value of the asset and uninformed traders who trade purely for liquidity purposes. The key feature of this model is that the arrival of these two types of traders are governed by independent poisson processes. Regardless of information events, the arrival rate of uninformed traders is ε. Informed traders, on the other hand, will arrive to the market only if an information event has occurred, and then only on one side of the market with arrival rate μ. The probability of informed trading is then given by:

$$PI = \frac{\alpha\mu}{\alpha\mu + 2\varepsilon}.$$

Since the distribution of the number of buy and sell trades is governed by poisson processes, the parameters of the model may be estimated via maximum likelihood. The likelihood of observing B buys and S sells over some time interval T on a given day can therefore be summarized as the sum of three weighted poisson processes where the weights are given by the probability of an information even, α, and the probability of the bad news, δ:

$$\begin{aligned} L(B,S|\alpha,\delta,\varepsilon,\mu) &= (1-\alpha)\left[e^{-(\varepsilon T)}\frac{(\varepsilon T)^B}{B!}\right]\left[e^{-(\varepsilon T)}\frac{(\varepsilon T)^S}{S!}\right] \\ &\quad +\alpha(1-\delta)\left[e^{-((\varepsilon+\mu)T)}\frac{((\varepsilon+\mu)T)^B}{B!}\right]\left[e^{-(\varepsilon T)}\frac{(\varepsilon T)^S}{S!}\right] \\ &\quad +\alpha\delta\left[e^{-(\varepsilon T)}\frac{(\varepsilon T)^B}{B!}\right]\left[e^{-((\varepsilon+\mu)T)}\frac{((\varepsilon+\mu)T)^S}{S!}\right]. \end{aligned}$$

Assuming that the days are independent, the likelihood of observing a sequence of buys and sells $X = (B_i, S_i)_{i=1}^N$ over N days is simply:

$$L(X|\alpha,\delta,\varepsilon,\mu) = \prod_{i=1}^{N} L_i(B_i,S_i|\alpha,\delta,\varepsilon,\mu).$$

[8] This is in contrast to the Glosten and Harris (1988) model presented above, where information events may occur at any time.

We use constrained maximum likelihood to estimate the parameters of the likelihood function. The probabilities, α and δ are constrained to lie in the interval $(0, 1)$ and the arrival rates, ε and μ are constrained to be between zero and ∞. The probability of informed trading, PI, is then constructed using our estimates of α, ε, and μ. Standard errors for PI are computed by the delta method.

3 Data

Our sample consists of all domestic common stock securities listed on the NYSE, AMEX and NASDAQ. From this universe of stocks, we select only those firms for which data can be found on both the CRSP tapes and the TAQ database. Further, we require that there be at least 500 trades in each security per quarter to ensure accurate estimates of our information measures. In addition, a security must be actively traded for at least 30 trading days during a quarter to be included in the sample. ¿From these filters, we retain roughly 5,500 firms per quarter from Q4 1997 to Q4 1998. Table 1 provides a description of our sample over time, by exchange. Roughly one third of our firms are from the NYSE, 7% from AMEX and 60% from the NASDAQ. Since there are significant differences in the market microstructure of these exchanges, we allow our results to vary by exchange. In addition, Section 6 explores the sensitivity of our results to exchange type.

The data for our tests come from three sources. First, we collect descriptive statistics from CRSP. For each firm in our sample, we construct quarterly measures of price (P), market capitalization ($SIZE$), volume (VOL), share turnover ($TURN$), and volatility ($VOLAT$). Quarterly price is measured as the average daily closing price during the quarter. Quarterly market capitalization is the average daily market capitalization (price times shares outstanding) during the quarter. Quarterly volume is taken to be the total number of shares transacted over the quarter. Share turnover is total quarterly volume divided by shares outstanding. The volatility of returns is constructed for each firm-quarter as the standard deviation of daily returns over the quarter.

Our second data source is the TAQ database provided by the NYSE. These files contain records for all trades and quotes for the NYSE, AMEX, and NASDAQ. Quarterly measures of the quoted spread are calculated as the difference between the ask and bid price over all quotes during standard trading hours for each firm-quarter. Relative spreads are calculated as average quoted spreads divided by price midpoints. The transactions from the TAQ database are also used to construct the various information measures using the methodologies described in Section 2.

The third source of data for this study is Disclosure Incorporated's Compact D database. The percentage of shares held by institutions ($INST$) and and insiders ($INSIDE$) was collected from this database, and the data was matched to CRSP and TAQ by the firm's CUSIP number.

Table 2 describes our sample of firms. Given the large number of firms in our study, there is considerable variation in size, price, and volume over our sample. There is also significant left skewness in each of these variables (means are much larger than medians). As a result, we use log transformations of the variables in all our test to mitigate the influence of outliers.

Our five measures of the information content of trading are consistent with previous studies. The average quoted spread in our sample is 26.15 cents, or roughly 2/8ths. The average relative spread implies that the quoted spread is, on average, three percent of the price. Our measure of the adverse selection component of the spread is always greater than zero and generally varies between 15 and 60 percent of the spread. Our average estimate of HS is 34 percent of the spread and is roughly consistent with Huang and Stoll (1997). The adverse selection cost (in dollar terms) also has a reasonable magnitude and varies between 1.1 and 16.5 cents per trade with an average of 6.4 cents. The HFV measure is also always greater than zero, suggesting price changes are positively related to the direction and size of order flow. Our average HFV estimate of 3.85 implies that an unexpected 1,000 share buy order would lead, on average, to a 3.85 cent price increase (or 0.19 percent of a 20 dollar stock). The probability of informed trading is 0.23 on average, suggesting

that about one in every five traders is informed. Last, our measures of ownership structure are indicate that institutions own 31 percent of the firm's equity and insiders own 9.79 percent of the firm's equity on average.

Given that we use five different proxies for the information content of trades, we also examine the relationship between our various measures. Table 3 provides a correlation table for our five information proxies. All four information measures (HS, $HS2$, HFV, and PI) are positive and significantly related to the level of the quoted spread. However, all measures except for the PI are negatively related to the relative spread. This result may arise from the fact that the relative spread is constructed with price in the denominator. All of our other estimated measures (HS, $HS2$, HFV, and PI) are positively and significantly related to each other with the exception of the PI and HS which have a negative correlation. In general, the PI has a low correlation with our other measures of information-based trading. This finding is, perhaps, not surprising considering that the PI is estimated using only daily order imbalances while the other three estimated measures (HS, $HS2$, and HFV) are constructed using intraday price sensitivities. Overall, our five measures are related to each other in a roughly consistent manner, yet appear uncorrelated enough to provide independent proxies for information-based trading.

4 Univariate Results

Table 4 presents a univariate description of the relationship between our measures of information-based trading and ownership structure. The upper panel partitions institutional ownership and the lower panel partitions insider ownership by quintiles. In the upper panel, both the quoted spread and the relative spread are negatively related to institutional ownership, which is consistent with the notion that institutions are attracted to more liquid stocks. The HS and HFV measures of the informativeness of trade vary positively with institutional ownership, which would seem to indicate that a given trade in a stock with high levels of institutional ownership would be more informed.

The PI measure, however, tells a different story. Stocks that have high levels of institutional ownership have a lower probability that any given trade is informed. While the conclusions from different measures seem to be conflicting, these results are only univariate. It could be the case that the PI measure is low for stocks with high levels of institutional ownership simply because these stocks are larger and have much higher volume than the lower quintiles.

The univariate patterns for insider ownership are different than those for institutional ownership. The relative spread, the adverse selection cost ($HS2$), and the PI measures are all increasing with increased insider ownership, while the adverse selection component (HS) is decreasing with increased insider ownership. Again, the same concerns that apply to the univariate institutional ownership analysis apply here as well. For example, insider ownership is negatively related to both the size and turnover of the firm. It could be the case that the PI measure is high for stocks with high levels of insider ownership simply because these stocks have much lower volume than the other quintiles. To address these problems, in Section 5 we examine these questions in a multivariate setting.

5 Multivariate Results

In the previous section, our univariate tests produced mixed results on the relationship between ownership structure and information-based trading. However, there may be systematic differences in firm size and liquidity that vary with both ownership and information-based trading. In this section, we test the relationship between ownership and information, controlling for differences in other factors that may affect information. There are two relationships that we have to concern ourselves with. First, our hypothesis is that higher levels of institutional ownership are positively related to our measures of informed trading. Simultaneously, our measures of informed trading could influence the level of institutional ownership. For example, it is known that institutions prefer firms that are more liquid and have lower bid-ask spreads. Due to this, we use a simultaneous

equation model to account for the effect that ownership structure and our measures of informed trading have on each other. We include control variables that account for firm size and liquidity. The simultaneous equation model is:

$$\begin{aligned} INFO_{i,t} &= a_1 INST_{i,t} + a_2 INSIDE_{i,t} + a_3 VOL_{i,t} + a_4 TURN_{i,t} \qquad (1) \\ &\quad + a_5 VOLAT_{i,t} + a_6 NASDAQ_{i,t} + a_7 AMEX_{i,t} + a_8 NYSE_{i,t}, \\ INST_{i,t} &= b_0 + b_1 SIZE_{i,t} + b_2 VOL_{i,t} + b_3 TURN_{i,t} + b_4 \frac{1}{P_{i,t}} + b_5 INFO_{i,t}. \qquad (2) \end{aligned}$$

We include volume in (1) to control for liquidity. Even if a stock has low volume, it could still be considered liquid if the volume was high relative to the shares outstanding. To account for this, we include turnover in addition to volume. Volatility is included since this may also affect the spreads that the market maker sets. A dummy variable is included for the exchange where the firm's stock trades. Since the sum of the three exchange dummies equals one for each observation, no constant term is needed in (1). Recent research by Gompers and Metrick (2001) and Falkenstein (1996) provides guidance about which control variables to include in the second equation. Since institutions prefer equity from large firms with relatively high prices and high liquidity, we include size, volume, turnover and price in (2).

The correlations between the independent variables are in Table 5. As we would expect, the correlation between institutional ownership, size, volume price and turnover is large and positive. Also, stocks with high institutional ownership tend to have lower volatilities.[9] The correlations between insider ownership and institutional ownership size are negative, indicating that those firms that have a high percentage of insider ownership tend to be smaller and not heavily owned by institutions. Furthermore, the size of the firm, volume and turnover are all highly positively correlated.

We estimate the coefficients in the simultaneous equation model comprised of equations (1) and (2) by two-stage least squares. Table 6 presents the results from the estimation. Consistent

[9]These results are consistent with Gompers and Metrick (2001).

with the results of the univariate analysis, the estimated coefficient on $INST$ in the regression where the relative spread is dependent variable is negative and significant. We interpret this as the preference of institutions for more liquid stocks. This result is contrary to the findings of Sharin, Shastri and Shastri (2000) who find that the relationship between institutional ownership and spreads is positive. Since their sample consists of 8 months of data from 1984, this results could be driven, in part, by the different sample periods. While there is a negative relationship between institutional ownership and the relative spread, the opposite is true for our measures of adverse selection and price impact. Controlling for firm size, volatility, share turnover, volume, and exchange type, there is still an economically and statistically significant positive relationship between ownership by institutions and the adverse selection component of the spread (HS and $HS2$) and price impact, (HFV), which is consistent with the evidence in Sugato (2001). It is puzzling that institutional ownership is negatively related to the probability of informed trading. While this is consistent with the univariate results, is opposite what we expected. In results discussed later in the paper, this only appears to be the case for NASDAQ stocks - there is no significant relationship between institutional ownership and the probability of informed trading for stocks listed on the NYSE/AMEX.

Similar to the results for the estimated coefficients on institutional ownership, we find that all three measures of information based trading - the adverse selection component of the spread, price impact, and probability of informed trading are positive and significant, as we would expect. While our results suggest that both types of ownership are positively related to the adverse selection component of and price impact on the bid-ask spread, institutional ownership has an economic impact that is an order of magnitude larger. Since institutions generate a much higher volume of trade than insiders, this result is consistent with risk-averse market makers wanting to protect themselves from incurring larger dollar losses trading against informed traders.

Generally, the control variables have the expected signs and significance. The coefficient on

volume is negative and significant, indicating that more liquid stocks have lower costs of informed trading. The coefficient on turnover is positive and significant in the relative spread regression, which is puzzling. However, it since turnover is highly correlated with volume it is difficult to interpret this result. Also, the sign and significance of the estimated coefficients on the ownership variables, which are the central variables of interest, do not change when when turnover is dropped from the model. The coefficients on volatility are all positive and significant, indicating that more volatile firms have higher information costs associated with them. Though we estimated the models using the logarithm of volume, turnover and volatility, the signs and significance of the coefficients are the same when the original variables are used.

The second panel in Table 6 contains the estimated coefficients from equation (2) in the model. Here, institutional ownership is a function of the measure of informed trading, firm size, volume, turnover, and the reciprocal of price. The estimated coefficient on the relative spread is negative, which is consistent with existing evidence that institutions prefer stocks that have lower spreads. While the estimated coefficient on percentage of the spread that represents the adverse selection cost (HS) is is positive, the coefficient on the dollar cost of the adverse selection ($HS2$) is negative. This implies that institutions avoid these stocks. One possible explanation is that they do so in an attempt to lower their effective cost of trading. On the other hand, the estimated coefficient on price impact (HFV) and probability of informed trading (PI) are all positive and significant, which implies that institutions favor these stocks. While this is counter intuitive, there are many factors that go into a money managers decision to purchase a stock besides our measures of informed trading. Omission of these factors could bieas this this result. The purpose of this study is not to model these factors, but rather to understand the impact of ownership structure on informed trading.

Size and turnover are positive and significant, indicating that institutions prefer large firms with high turnover. Volume is negative an significant in all the models, which is opposite what one would

expect. Again, this result appears to be driven by the correlation with turnover. When the model is re-estimated without turnover, volume is positive and significant, as one would expect. Since we are primarily concerned with the estimated coefficents on the ownership variables in equation (1), the inclusion of highly correlated control variables will not bias our estimates of these coefficients.

6 Differences by exchange type

In the previous sections, we show that the information content of trades increases with the proportion of institutional and insider ownership. However, the nature of information-based trading may vary systematically by the type of exchange an issue is traded on. For example, recent evidence suggests that dealer markets like the NASDAQ may be more anonymous than auction markets.[10] In this section, we test whether the relationship between information and ownership structure differs by exchange type.

In Table 6 the coefficient on the NASDAQ dummy variable is significant and more positive than the NYSE dummy in the spread and $HS2$ regressions. This result is consistent with earlier work comparing transactions costs on dealer and auction markets (See, e.g. Huang and Stoll (1995)). However, NASDAQ stocks have, on average, a smaller adverse selection component of the spread and probability of informed trading when compared to the NYSE stocks. These results are also consistent with previous studies comparing the components of the spread between dealer and auction markets (see Affleck-Graves, Hedge, and Miller (1994)) The AMEX, also an auction market, has mixed results compared to the NYSE with a smaller spread and probability of informed trading but a larger HS, HS2 and HFV.

Table 7 provides a summary of our estimation of the simultaneous equation model of equation 1 by exchange type. Here we include a constant but do not include exchange dummies. We perform

[10]For example, see Garfinkel and Nimalendran (1998) or Heidle and Huang (2000).

the regression separately for firms listed on an auction market (like the NYSE or AMEX) and firms listed on a dealer market (NASDAQ). Table 7 presents the coefficients on the ownership variables only.[11] Overall, we find that the relationship between ownership and informed trading is consistent across different exchange types. The only difference between the exchanges appears to be in the probability of informed trading. Insider ownership is positively related to the probability of informed trading on the NYSE-AMEX, but not the NASDAQ. In contrast, institutional ownership is not related to the probability of informed trading on the NYSE-AMEX, but on the NASDAQ there is a negative relationship. The coefficients on ownership for the HS, PI, and HFV regressions are smaller for the NASDAQ sample while the impact of ownership on our HS2 measures are larger for NASDAQ stocks. We interpret these results as evidence that individual investors are less informed than institutions or insiders, regardless of the trading environment.

7 Conclusion

This paper tests the relationship between ownership structure and the information content of equity trading. We find that information-based trading is positively and significantly related to the amount of both institutional and inside ownership. Our results are robust to a variety of estimation techniques and proxies for informed trading.

Overall, we find strong evidence that both institutions and insiders are better informed (Or, at least, that their trades have a greater impact on prices and order imbalances) than individual investors. These results are consistent with the hypothesis that there are economies of scale in information acquisition and aggregation. Further, the positive relation between informed trading and institutional ownership suggests that the previously documented relationship between institutional ownership and share turnover may be driven by information-based trading rather than churning.

[11] The coefficients on the control variables are qualitatively similar to those presented in Table 6.

References

[1] Affleck-Graves, John, S. Hedge, and R. Miller, 1994, Trading Mechanisms and the Components of the Bid-Ask Spread, *The Journal of Finance* 49, 1471-1488.

[2] Alangar, S., C. Bathala and R. Rao, 1999, The Effect of Institutional Interest on the Information Content of Dividend-Change Announcements, *The Journal of Financial Research* 22, 429-448.

[3] Amihud, Yakov, and Haim Mendelson, 1986, Asset pricing and the bid-ask spread, *Journal of Financial Economics* 17, 223-249.

[4] Bartov, E. , S. Radhakrishnan and I. Krinsky, 2000, Investor sophistication and patterns in stock returns after earnings announcements. *The Accounting Review* 75, 43-63.

[5] Brennan, Michael J. and Avanidhar Subrahmanyam, 1996, Market Microstructure and Asset Pricing: On the Compensation for Illiquidity in Stock Returns, *Journal of Financial Economics* 41, 441-64.

[6] Dow, J., G. Gorton, 1997, Noise Trading, Delegated Portfolio Management, and Economic Welfare. *Journal of Political Economy* 105, 1024-50.

[7] Easley, David, Nicholas M. Kiefer, Maureen O'Hara, and Joseph B. Paperman, 1996, Liquidity, information, and infrequently traded stocks, *Journal of Finance* 51, 1405-1436.

[8] Falkenstein, E. Preferences for Stock Characteristics as Revealed by Mutual Fund Portfolio Holdings. *Journal of Finance*, V. 51, N. 1. (1996), 111-136.

[9] Foster, Douglas F., and S. Viswanathan, 1993, Variations in trading volume, return volatility, and trading costs: Evidence on Recent Price Formation Models, *Journal of Finance* 48, 187-211.

[10] Garfinkel, Jon A., and M. Nimalendran, 1998, Market Structure and Trader Anonymity: An Analysis of Insider Trading, Working Paper.

[11] George, T., Gautam Kaul, and M. Nimalendran, 1991, Estimation of the Bid-Ask Spread and its Components: A New Approach. *Review of Financial Studies* 4, 623-656.

[12] Glosten, Lawrence R., and Lawrence E. Harris, 1988, Estimating the components of the bid/ask spread, *Journal of Financial Economics* 21, 123-42.

[13] Gompers, P.A. and A. Metrick, 2001, Institutional Investors and Equity Prices, *Quarterly Journal of Economics* 116, 229-259.

[14] Hasbrouck, Joel, 1991, Measuring the information content of stock trades, *Journal of Finance* 46, 179-207.

[15] Heidle, Hans, and Roger D. Huang, 1999, Information-Based Trading in Dealer and Auction Markets: An Analysis of Exchange Listings, Working paper, Vanderbilt University.

[16] Huang, Roger D., and Hans Stoll, 1997, The components of the bid-ask spread: a general approach, *Review of Financial Studies* 10, 995-1034.

[17] Huang, Roger D., and Hans Stoll, 1996, Dealer versus Auction Markets: A Paired Comparison of Execution Costs on NASDAQ and the NYSE, *Journal of Financial Economics* 41, 313-57.

[18] Jennings, W., K. Schnatterly, and Paul Seguin, 2000, Institutional Ownership, Information, and Liquidity. *Working Paper.*

[19] Kothare, M. and P. Laux, 1995, Trading Costs and the Trading Systems for NASDAQ Stocks, *Financial Analysts Journal*, 42-53.

[20] Kyle, Albert, 1985, Continuous auctions and insider trading, *Econometrica* 53, 1315-35.

[21] Madhavan, Anath, and George Sofianos, 1998, An Empirical Analysis of NYSE Specialist Trading, *Journal of Financial Economics* 48, 189-210.

[22] Madhavan, Anath, and Seymour Smidt, 1993, An Analysis of Changes in Specialist Inventories and Quotations, *Journal of Finance* 48, 1595-1628.

[23] Sarin, A., Shastri, K., and Shastri K., Nov 2000, Ownership Structure and Stock Market Liquidity, Working Paper, University of Pittsburg.

[24] Chakravarty, Sugato, 2001, Stealth Trading: Which Traders' Trades Move Stock Prices?, *Journal of Financial Economics* Forthcoming.

[25] Szewczyk, S., G. Tsetsekos and R. Varma, 1992, Institutional Ownership and the Liquidity of Common Stock Offerings, *The Financial Review* 27, 211-225.

[26] Trueman, B., 1988, A Theory of Noise Trading in Securities Markets. *Journal of Finance* 43, 83-95.

Table 1:

Sample Description

This table contains a description of the sample by date and exchange listing. The sample is selected from the universe of NASDAQ, NYSE and AMEX stocks for which data could be found from CRSP, TAQ, and Disclosure.

Number of firms in the sample	NYSE	AMEX	NASDAQ	Total
Q4 1997	1,719	394	3,629	5,742
Q1 1998	1,648	362	3,164	5,174
Q2 1998	1,759	414	3,448	5,621
Q3 1998	1,774	401	3,530	5,705
Q4 1998	1,744	400	3,227	5,371
Total	8,644	1,971	16,998	27,613

Table 2:

Summary Statistics

This table contains summary statistics for our sample. Reported statistics are computed as the cross-sectional averages over all firm-quarter observations. Market value is the average quarterly price (over all daily closing prices) times the number of shares outstanding. Price is the average daily closing price over each quarter. Quarterly volume is computed as the total number of shares traded each quarter. Share turnover is quarterly volume divided by shares outstanding. The standard deviation of returns is constructed for each firm-quarter as the standard deviation of daily returns over the quarter. The quoted spread is taken as the difference between the posted ask and bid price over all transactions for each stock-quarter as reported on the TAQ database. The relative spread is the quoted spread divided by price. The adverse selection component of the spread is estimated using the methodology of Huang and Stoll (1997). The adverse selection cost is the adverse selection component of the spread multiplied by the average quoted spread. The Hasbrouck-Foster-Visvanathan measure is constructed following Hasbrouck (1991) (with five lags). The probability of informed trading is estimated via maximum likelihood following Easley, Kiefer, O'Hara, and Paperman (1995). Institutional ownership data are collected from Compact Disclosure.

Descriptive Variables	Obs	Mean	Std.	5%	median	95%
Market value, Millions ($SIZE$)	27,613	1,669	8,668	8	135	5,975
Price, Dollars (P)	27,613	18.87	19.94	1.28	13.61	53.48
Volume, Millions of Shares (VOL)	27,613	8.80	29.10	0.14	1.66	38.00
Share Turnover ($TURN$)	27,613	0.18	0.18	0.03	0.13	0.50
Standard Deviation of returns ($VOLAT$)	27,613	0.04	0.03	0.01	0.03	0.10
Information variables						
Quoted Spread (cents)	27,613	26.15	20.99	8.28	19.29	68.58
Relative Spread (%)	27,613	3.19	3.64	0.32	2.11	9.83
Adverse Selection Component of the spread (HS)	27,607	33.97	14.62	15.48	30.75	61.98
Adverse Selection Cost ($HS2$)	27,607	6.39	5.68	1.13	4.93	16.50
Hasbrouck-Foster -Visvanathan measure (HFV)	26,351	3.85	3.18	0.55	3.11	9.50
Probability of informed trading (PI)	27,613	22.91	8.17	11.14	22.38	36.80
Ownership variables						
Percentage Institutional Ownership ($INST$)	27,512	31.06	25.86	0.00	25.59	78.98
Percentage Insider Ownership ($INSIDE$)	27,437	9.79	17.41	0.00	0.35	51.30

Table 3:

Correlations between between information measures

This table contains the correlation coefficients between the proxies for the information content of trades that are used in the regressions. The quoted spread is taken as the difference between the posted ask and bid price over all transactions for each stock-quarter as reported on the TAQ database. The relative spread is the quoted spread divided by price. The adverse selection component of the spread is estimated using the methodology of Huang and Stoll (1997). The adverse selection cost is the adverse selection component of the spread multiplied by the average quoted spread. The Hasbrouck-Foster-Visvanathan measure is constructed following Hasbrouck (1991) (with five lags). The probability of informed trading is estimated via maximum likelihood following Easley, Kiefer, O'Hara, and Paperman (1995). P-values are reported in parenthesis.

	Quoted Spread	Relative Spread	Information Component of the spread (HS)	Information Cost ($HS2$)	Price Impact (HFV)
Relative Spread	0.0013				
	(0.830)				
Information Component of the spread (HS)	0.0692	-0.3470			
	(0.000)	(0.000)			
Information Cost ($HS2$)	0.8214	-0.1072	0.5122		
	(0.000)	(0.000)	(0.000)		
Price Impact (HFV)	0.5804	-0.1732	0.6205	0.7979	
	(0.000)	(0.000)	(0.000)	(0.000)	
Probability of informed trading (PI)	0.0595	0.3336	-0.1126	0.0111	0.0338
	(0.000)	(0.000)	(0.000)	(0.066)	(0.000)

Table 4:

Univariate Relationships Between Information and Ownership

This table contains estimates of our five measures of information-based trading by quintile of institutional ownership and quartile of insider ownership. Observations that have zero insider or institutional ownership have their own category and are excluded from the quintiles. There are 1,971 firm-quarter observations that have no institutional ownership and 5,108 firm-quarter observations in each institutional ownership quintile. There are 12,559 firm-quarter observations that have no insider ownership and 2,976 firm-quarter observations in each institutional ownership quintile. The quoted spread is taken as the difference between the posted ask and bid price over all transactions for each stock-quarter as reported on the TAQ database. The relative spread is the quoted spread divided by price. The adverse selection component of the spread is estimated using the methodology of Huang and Stoll (1997). The adverse selection cost is the adverse selection component of the spread multiplied by the average quoted spread. The Hasbrouck-Foster-Visvanathan measure is constructed following Hasbrouck (1991) (with five lags). The probability of informed trading is estimated via maximum likelihood following Easley, Kiefer, O'Hara, and Paperman (1995).

Quintile of Institutional Ownership	Institutional Ownership	Quoted Spread	Log of Relative Spread	HS	HS2	HFV	PI
$INST = 0$	0.000	0.235	-3.428	0.303	0.054	0.033	0.245
1	0.034	0.264	-3.100	0.279	0.058	0.031	0.258
2	0.146	0.299	-3.545	0.319	0.071	0.039	0.242
3	0.289	0.286	-3.920	0.347	0.070	0.041	0.233
4	0.477	0.254	-4.432	0.368	0.066	0.040	0.214
5	0.727	0.215	-4.965	0.399	0.059	0.043	0.192

Quintile of Insider Ownership	Insider Ownership	Quoted Spread	Log of Relative Spread	HS	HS2	HFV	PI
$INSIDE = 0$	0.000	0.258	-3.932	0.342	0.063	0.038	0.229
1	0.005	0.205	-4.716	0.364	0.053	0.037	0.197
2	0.034	0.274	-4.164	0.353	0.068	0.041	0.223
3	0.104	0.290	-3.833	0.324	0.069	0.038	0.233
4	0.232	0.268	-3.673	0.319	0.064	0.036	0.241
5	0.526	0.281	-3.512	0.327	0.068	0.040	0.252

Table 5:

Correlations between independent variables

This table contains the correlation coefficients between the independent variables that are used in the pooled and fixed effects regressions Market value is the average quarterly price (over all daily closing prices) times the number of shares outstanding. Share turnover is quarterly volume (total number of shares traded each quarter) divided by shares outstanding. Volatility is constructed for each firm-quarter as the standard deviation of daily returns over the quarter. All correlations are significant at the 1% level.

	Institutional Ownership ($INST$)	Inside Ownership ($INSIDE$)	Log Market Value ($SIZE$)	Log Turnover ($TURN$)	Log Volatility ($VOLAT$)	Log Volume (VOL)
$INSIDE$	-0.1479					
$SIZE$	0.6019	-0.1558				
$TURN$	0.3661	-0.0795	0.2271			
$VOLAT$	-0.3885	0.1556	-0.6225	0.0935		
VOL	0.5183	-0.1410	0.7304	0.7085	-0.1659	
$\frac{1}{P}$	-0.2947	0.045	-0.4479	-0.0445	0.5117	-0.0981

Table 6:

Determinants of information-based trading

This table contains results of regressions of proxies for the information content of trading on explanatory variables. The simultaneous equation model is:

$$\begin{aligned} INFO_{i,t} = \; & a_1 INST_{i,t} + a_2 INSIDE_{i,t} + a_3 VOL_{i,t} + a_4 TURN_{i,t} \\ & + a_5 VOLAT_{i,t} + a_6 NASDAQ_{i,t} + a_7 AMEX_{i,t} + a_8 NYSE_{i,t}, \\ INST_{i,t} = \; & b_0 + b_1 SIZE_{i,t} + b_2 VOL_{i,t} + b_3 TURN_{i,t} + b_4 \tfrac{1}{P_{i,t}} + b_5 INFO_{i,t} \end{aligned}$$

where t indexes the end of calendar quarter for the period from December 31, 1997 to December 31, 1998 and i indexes the firm. The dependent variable, $INFO$, is one of the five measures of the information component of the spread. $LNSPREAD$ is the log of the difference between the posted ask and bid price, divided by the price midpoint, over all transactions for each stock-quarter as reported on the TAQ database. HS is the adverse selection component of the spread is estimated using the methodology of Huang and Stoll (1997). $HS2$, the adverse selection cost, is the adverse selection component of the spread multiplied by the average quoted spread. HFV is the Hasbrouck-Foster-Visvanathan measure constructed following Hasbrouck (1991) (with five lags). PI is the probability of informed trading estimated via maximum likelihood following Easley, Kiefer, O'Hara, and Paperman (1995). $INST$ is the percentage of shares held by institutional owners, $INSIDE$ is the percentage of shares held by company insiders, $TURN$ is the average daily turnover for the firm during the quarter, $VOLAT$ is the logarithm of the standard deviation of the firm's returns over the preceding 90 days, $SIZE$ is the natural logarithm of the market value of equity, $NASDAQ$, $AMEX$, and $NYSE$ are equal to one if the firm trades on the respective exchange and zero otherwise. The t-statistics are shown in the parentheses below coefficient estimates.

	LNSPREAD	HS	HS2	HFV	PI
INST	-4.356	0.819	0.509	0.271	-0.167
	(-53.0)	(50.9)	(46.0)	(46.5)	(-21.6)
INSIDE	-0.228	0.057	0.029	0.017	0.007
	(-8.2)	(10.6)	(8.9)	(9.0)	(2.5)
VOL	-0.143	-0.074	-0.036	-0.020	-0.019
	(-22.0)	(-57.2)	(-41.6)	(-43.5)	(-30.6)
TURN	0.180	-0.021	-0.026	-0.014	0.017
	(17.5)	(-10.4)	(-20.2)	(-19.8)	(18.8)
VOLAT	0.482	0.065	0.039	0.026	0.009
	(31.8)	(21.9)	(21.4)	(24.6)	(6.0)
NASDAQ	1.488	1.245	0.509	0.299	0.607
	(13.4)	(57.5)	(36.1)	(38.8)	(61.5)
AMEX	1.117	1.540	0.575	0.342	0.575
	(9.6)	(67.5)	(38.1)	(41.8)	(54.7)
NYSE	1.534	1.418	0.476	0.304	0.639
	(13.0)	(61.8)	(32.2)	(37.5)	(61.6)
N	27340	27334	27334	26090	27340

Table 6 (continued)

This table contains the estimated coefficients and t-statistics for the second equation in the simultaneous equation model where the dependent variable is the percentage of institutional ownership ($INST$).

LNSPREAD	-0.067				
	(-12.8)				
HS		0.127			
		(9.8)			
HS2			-0.922		
			(-13.1)		
HFV				0.331	
				(3.1)	
PI					2.113
					(13.8)
SIZE	0.066	0.097	0.139	0.102	0.134
	(19.5)	(58.3)	(47.0)	(39.6)	(48.5)
VOL	-0.045	-0.051	-0.115	-0.057	-0.045
	(-18.6)	(-21.7)	(-24.3)	(-14.6)	(-15.4)
TURN	0.104	0.127	0.170	0.136	0.143
	(28.5)	(44.6)	(43.1)	(38.6)	(40.7)
$\frac{1}{P}$	0.030	0.026	0.047	0.026	0.008
	(10.8)	(8.9)	(11.6)	(7.3)	(1.6)
CONSTANT	0.571	0.769	1.673	0.875	0.095
	(14.1)	(23.0)	(25.7)	(16.5)	(1.3)

Table 7:

Results by exchange type

This table contains results of regressions of proxies for the information content of trading on explanatory variables. Panel A includes only those firms that trade on the NYSE or the AMEX, while Panel B includes only those firms that trade on the NASDAQ. Only the estimated coefficients on $INST$ and $INSIDE$ for the first equation in the system are reported. The estimates for the other coefficients in the model are qualitatively the same as those reported in Table 6. The simultaneous equation model is:

$$\begin{aligned} INFO_{i,t} = \quad & a_1 INST_{i,t} + a_2 INSIDE_{i,t} + a_3 VOL_{i,t} + a_4 TURN_{i,t} \\ & + a_5 VOLAT_{i,t} + a_6 NASDAQ_{i,t} + a_7 AMEX_{i,t} + a_8 NYSE_{i,t}, \\ INST_{i,t} = \quad & b_0 + b_1 SIZE_{i,t} + b_2 VOL_{i,t} + b_3 TURN_{i,t} + b_4 \tfrac{1}{P_{i,t}} + b_5 INFO_{i,t} \end{aligned}$$

where t indexes the end of calendar quarter for the period from December 31, 1997 to December 31, 1998 and i indexes the firm. The dependent variable, $INFO$, is one of the five measures of the information component of the spread. $LNSPREAD$ is the log of the difference between the posted ask and bid price, divided by the price midpoint, over all transactions for each stock-quarter as reported on the TAQ database. HS is the adverse selection component of the spread is estimated using the methodology of Huang and Stoll (1997). $HS2$, the adverse selection cost, is the adverse selection component of the spread multiplied by the average quoted spread. HFV is the Hasbrouck-Foster-Visvanathan measure constructed following Hasbrouck (1991) (with five lags). PI is the probability of informed trading estimated via maximum likelihood following Easley, Kiefer, O'Hara, and Paperman (1995). $INST$ is the percentage of shares held by institutional owners, $INSIDE$ is the percentage of shares held by company insiders, $TURN$ is the average daily turnover for the firm during the quarter, $VOLAT$ is the logarithm of the standard deviation of the firm's returns over the preceding 90 days, $SIZE$ is the natural logarithm of the market value of equity. The t-statistics are shown in the parentheses below coefficient estimates.

Panel A: Results for NYSE-AMEX

	LNSPREAD	HS	HS2	HFV	PI
INST	-6.971	1.414	0.434	0.299	-0.008
	(-28.5)	(28.1)	(19.1)	(24.4)	(-0.6)
INSIDE	-0.617	0.129	0.036	0.033	0.026
	(-6.6)	(6.7)	(5.6)	(7.4)	(4.7)
N	10,504	10,502	10,502	10,190	10,504

Panel B: Results for NASDAQ

	LNSPREAD	HS	HS2	HFV	PI
INST	-3.064	0.620	0.494	0.233	-0.200
	(-47.9)	(48.9)	(45.7)	(42.1)	(-22.9)
INSIDE	-0.115	0.029	0.022	0.010	0.004
	(-5.1)	(6.2)	(6.4)	(5.6)	(1.2)
N	16,836	16,832	16,832	15,900	16,836