Order flow and prices

Ekkehart Boehmer and Julie Wu*

Mays Business School
Texas A&M University
College Station, TX 77843-4218

August 12, 2007

Abstract

We provide new evidence on the relation between order flow and prices, an issue that is central to asset pricing and market microstructure. We examine proprietary data on a broad panel of NYSE-listed stocks that reveal daily order imbalances by institutions, individuals, and market makers. We can further differentiate regular institutional trades from institutional program trades. Our results indicate that order imbalances from different trader types play distinctly different roles in price formation. Institutions and individuals are contrarians with respect to previous-day returns, but differ in the effect their order imbalances have on contemporaneous returns. Institutional imbalances are positively related to contemporaneous returns, and we provide cross-sectional evidence that this relation is likely to be the result of firm-specific information institutions have. Individuals, specialists, and other market makers appear to provide liquidity to these actively trading institutions. Our results also suggest a special role for institutional program trades. Institutions choose program trades when they have no firm-specific information and can afford to trade passively. As a result, program trades provide liquidity to the market. Finally, both institutional non-program and individual imbalances (information which is not available to market participants) have predictive power for next-day returns.

Keywords: Order flow, price formation, informed trading, trader types

JEL code: G14

* Email addresses: eboehmer@mays.tamu.edu, jwu@mays.tamu.edu. We thank Kerry Back, Stewart Mayhew, Gideon Saar, Joshua Slive, and seminar participants at George Mason University, Texas A&M University, the University of Arizona, the 2006 Workshop on the Microstructure of Foreign Exchange and Equity Markets in Ottawa, and the 2007 American Finance Association Meetings in Chicago for valuable comments.

Order flow and prices

Abstract

We provide new evidence on the relation between order flow and prices, an issue that is central to asset pricing and market microstructure. We examine proprietary data on a broad panel of NYSE-listed stocks that reveal daily order imbalances by institutions, individuals, and market makers. We can further differentiate regular institutional trades from institutional program trades. Our results indicate that order imbalances from different trader types play distinctly different roles in price formation. Institutions and individuals are contrarians with respect to previous-day returns, but differ in the effect their order imbalances have on contemporaneous returns. Institutional imbalances are positively related to contemporaneous returns, and we provide cross-sectional evidence that this relation is likely to be the result of firm-specific information institutions have. Individuals, specialists, and other market makers appear to provide liquidity to these actively trading institutions. Our results also suggest a special role for institutional program trades. Institutions choose program trades when they have no firm-specific information and can afford to trade passively. As a result, program trades provide liquidity to the market. Finally, both institutional non-program and individual imbalances (information which is not available to market participants) have predictive power for next-day returns.

Keywords: Order flow, price formation, informed trading, trader types

JEL code: G14

Order flow and prices

1. Introduction

A central prediction of market microstructure theory is that order flow affects prices. This follows from inventory models, where market makers temporarily adjust prices in response to incoming orders (Garman, 1976; Amihud and Mendelson, 1980; Stoll, 1978; Ho and Stoll, 1981). It also follows from information-based models where some traders have information about future asset value, so their trades lead to permanent price adjustments (Glosten and Milgrom, 1985; Kyle, 1985; Easley and O'Hara, 1987). The prediction that order flow affects prices is robust to competition among informed traders (Holden and Subrahmanyam, 1992), endogenous order sizes (Back and Baruch, 2007), and the consideration of strategic uninformed traders (Admati and Pfleiderer, 1988; Spiegel and Subrahmanyam, 1992).

Empirical research is almost uniformly consistent with this basic prediction. Earlier studies using intraday transactions data generally support both inventory and information effects. For example, Ho and Macris (1984) document that an options specialist adjusts prices in a way that is consistent with inventory models. Hasbrouck (1988, 1991a, 1991b) uses Vector Auto-Regressive models to disentangle (transient) inventory effects from (permanent) information effects. He demonstrates significant information effects on prices and some evidence consistent with inventory adjustments. Recent research based on daily observations also finds that net order flow, the difference between buy and sell volume, affects contemporaneous and next-day returns. Chordia, Roll, and Subrahmanyam (2002) show that the aggregate order imbalance is positively associated with market returns, and Chordia and Subrahmanyam (2004) obtain comparable results in the cross-section of stocks.[1]

While microstructure theory clearly distinguishes among different trader types according to their information and motives for trading, data limitations typically limit empirical tests to analysis that pools all traders. In this paper, we use a unique dataset derived from NYSE audit trail data that allows us to distinguish

---

[1] A related literature focuses on the relation between trading volume and returns, see Campbell, Grossman, and Wang (1993), Chordia, Roll, and Subrahmanyam (2001), Llorente, Michaely, Saar and Wang (2002), Baker and Stein (2004), and Chordia, Huh, and Subrahmanyam (2007). Karpoff (1987) surveys earlier work.

buys and sells from different trader types: individuals, institutions, non-NYSE market makers, and specialists. We further differentiate regular institutional trades, index arbitrage program trades, and other program trades.[2] Traders in these categories are likely to differ in their motives and strategies for trading and, in particular, in the quantity and quality of their private information. Therefore, we expect the relation between order flow and returns to differ across these trader types, and our tests are designed to capture these differences. Understanding how trader-type specific order flow affects prices has important implications for modeling the dynamics of liquidity provision, trader behavior, and market design. Moreover, analyzing these differences allows us to refine inferences from empirical microstructure research that is based on aggregate data.

Our first finding is that, during our sample period, institutions trade as contrarians with respect to prior-day returns. This is consistent with aggregate evidence in Lipson and Puckett (2005) and Chordia, Roll, and Subrahmanyam (2002), but contrary to the Nasdaq evidence in Griffin, Harris, and Topaloglu (2003). We further show that, for the largest size quartile, institutions are momentum traders with respect to market movements on the previous day. We argue that the countervailing effects of idiosyncratic and market returns could explain the differences between our results and those in Griffin et al., whose sample is limited to large firms during a period of substantial negative returns.

Second, we find that institutional imbalances are positively related to contemporaneous returns, controlling for market movements and persistence in imbalances. This suggests that institutional trading is associated with positive price impacts, as predicted by theory, and is consistent with a prevalence of information-based trading. While our daily data limits inferences about information content, we show that the institutional price impact coefficient is positively related to cross-sectional proxies for information asymmetry. In particular, institutional imbalances have a greater effect on contemporaneous returns in stocks with larger effective spreads,

---

[2] Program and index arbitrage program trades are institutional trades but we differentiate these from regular institutional trades. First, the NYSE defines program trades as simultaneous trades in 15 or more stocks worth at least $1 million. In contrast, the typical trade size on the NYSE is about $20,000. Second, trading motives differ. Index arbitrage program trading attempts to profit from temporary discrepancies between derivative and cash markets, whereas regular program trading can be associated with other specific trading strategies. Third, regulatory treatment differs across these order types. Both types of program trade must be reported to the exchange, and NYSE Rule 80A suspends some type of index arbitrage program trades on volatile trading days.

controlling for firm size. This indicates that information is an important driver of the effect that institutional imbalances have on prices, but it is also consistent with an inventory effect: if market makers hold undesirable inventory levels, liquidity would be limited, causing high spreads and larger effects of trading on returns. To disentangle these two explanations, we decompose effective spreads into a temporary component (likely associated with inventory effects) and a permanent component (likely associated with information in order flow). We find that institutional order imbalances have a greater effect on returns when the permanent component is large, even when controlling for inventory effects. Therefore, information appears to play a more prominent role than inventory effects in explaining how institutional trading affects prices. Finally, institutional order imbalances also have a greater price impact in stocks with higher R&D expenditures. Because the outcomes of R&D are very uncertain, firms with high R&D are more difficult to value, and subject to greater information asymmetry. The greater price impacts of institutional order imbalances in these stocks again suggest that trader information plays an important role in price formation.

Third, institutional imbalances have explanatory power for next-day returns. This also suggests that institutional trading is, at least in part, information based. We note that this predictive ability cannot be exploited to generate abnormal trading profits, because information on trader groups is confidential and not even disclosed ex post. No trader (including specialists) can observe the trader type and base his own trading on specific types' order flow.

About one quarter of institutional trading involves program trades, and we document that this order type plays a special role during our sample period. Institutions choose endogenously between a regular order and a program trade. Our priors are that program trades are unlikely to be motivated by firm-specific private information because they involve several different securities. Therefore, we expect the relation between order imbalances and prices to differ between program trades and regular institutional trades. The evidence strongly supports this hypothesis. While program-trade imbalances also tend to be contrarian, they differ from regular institutional imbalances in that they have a negative relation to contemporaneous returns. This suggests that institutions use program trades when they have little information and provide liquidity to other traders by trading passively.[3]

Similar to Kaniel, Saar, and Titman (2007), we also find that individuals trade as contrarians and that their trading can predict next-day returns. Kaniel et al. argue that individuals provide liquidity to institutions and we provide further evidence consistent with their interpretation. Specifically, we show that individuals' order imbalances have a negative effect on contemporaneous returns, consistent with liquidity provision. But individuals provide only 5% of trading volume, so that they alone cannot satisfy the imbalances of informed institutional traders. More specifically, in our sample the dollar value of individuals' order imbalances accounts for less than one fifth of the opposite institutional imbalances. This suggests that the remaining imbalance is filled mainly by other institutional traders who use program trades (which account for about 20% of trading volume), and to some extent by specialists and other market makers.

Our analysis is closely related to Chordia and Subrahmanyam (2004) and Griffin, Harris, and Topaloglu (2003). Chordia and Subrahmanyam develop a two-period trading model where a competitive discretionary liquidity trader can split orders between two periods. In addition, a nondiscretionary liquidity trader and a competitive informed trader, who receives a noisy signal before trading, submit orders in the second period. A competitive market maker picks up the imbalance resulting in each trading period. Chordia and Subrahmanyam show that it is optimal for the discretionary liquidity trader to split orders, so that order imbalances are positively autocorrelated over time. Moreover, because market makers can partially predict the second-period order imbalance, the model implies a positive relation between returns and lagged imbalances. Using a sample of (on average) 1,322 NYSE-listed stocks between 1988 and 1998, they estimate security-specific time series regressions and find evidence consistent with these predictions.[4]

---

[3] This is consistent with observations by industry participants as well. In practice, sell-side brokers maintain separate trading desks for program trades that do not expect incoming order flow to be informed.

[4] Chordia, Roll and Subrahmanyam (2002) use a similar approach to study daily order imbalances aggregated across stocks. They document that aggregate imbalances are highly persistent and positively related to contemporaneous market returns. They also find that, in the aggregate, traders exhibit contrarian behavior on daily basis.

Griffin, Harris, and Topaloglu (2003) observe the identity of brokerage firms in Nasdaq 100 stocks for each trade over 210 trading days from May 2000. They classify brokers according to their main clientele, and in this way obtain an approximate classification into institutional and retail for most of the trades. They document that institutional imbalances are persistent over several days. Moreover, institutions are more likely to buy after positive returns on the previous day and their imbalance has a positive contemporaneous relation to returns.

Our proprietary data set allows several additional inferences that complement the results in Chordia and Subrahmanyam and Griffin, Harris, and Topaloglu. First, in contrast to Chordia and Subrahmanyam's analysis of order flow aggregated across all traders, we do not have to infer trade direction or, implicitly, market maker trades using the Lee and Ready (1991) algorithm. Rather, we directly observe buys and sells for each trader type and market-makers. Second, Griffin, Harris, and Topaloglu's sample allows a distinction between institutional and retail trades, but it is limited to the 100 most liquid Nasdaq stocks over a short period.[5] In contrast, our panel is much broader and longer and provides a finer trader-type classification that does not depend on classifying brokerage firms. Finally, our objective is somewhat different. While we also provide some results on the determinants of order imbalances, our main focus is on their consequences for contemporaneous and future prices and on measures of market liquidity.[6]

The rest of this paper is organized as follows. We describe the data, sample selection, and variables in Section 2. Section 3 contains the empirical results and Section 4 concludes.

---

[5] One potential advantage of Griffin, Harris, and Topaloglu's (2003) data set is that it contains trade-by-trade information, which they exploit to look at intraday trading patterns. They find that institutions are positive-feedback traders with respect to five-minute interval returns.

[6] In a broader sense, our analysis is also related to several studies that address differences in order imbalances across U.S. trader types around some specific events. Lee (1992) examines order imbalances around earnings announcements to see if institutional investors react differently from individual investors to the same earnings news using trade sizes as proxies for institutions and individuals. Griffin, Harris, and Topaloglu (2005) study aggregate order imbalances of various types of Nasdaq traders around the "tech bubble." Several papers focus on traders in other countries. Choe, Kho and Stulz (1999) analyze order imbalances to investigate if foreign investors contribute to the Korean stock market crisis in 1997. Grinblatt and Keloharju (2000) investigate the trading behavior of Finnish investors. Lee et al. (2004) examine marketable order imbalances from various trader categories on the Taiwan Stock Exchange.

2. Data and sample construction

We use proprietary data from the New York Stock Exchange that allow us to separately observe buy and sell transactions for different trader types. These data cover all securities traded on the NYSE between January 2000 and April 2004 and are based on the NYSE's Consolidated Audit Trail Data (CAUD), which provide information on nearly all trades executed at the NYSE. CAUD are the result of matching trade reports to the underlying order data – for each trade, they show the executed portion of the underlying buy and sell orders. Each component is identified by an account-type variable that gives some information on trader identity. Providing the account type classification is mandatory for brokers (although it is not audited by the NYSE on a regular basis). Different regulatory requirements include obligations to indicate orders that are part of program trades, index arbitrage program trades, specialist trades, and orders from off-NYSE market makers in the stock. Each of these categories is further divided into proprietary member trades, trades by retail customers, and agency trades.

The data set available for this study contains aggregated buy and sell volume for each day and security for certain combinations of account types, represented by the number of trades, share volume, and dollar volume. We can distinguish the following six account-type categories: individuals, institutions, regular institutional program trades, institutional index arbitrage program trades, non-NYSE market maker proprietary trades, and specialists. NYSE account types have been used in a handful of other papers. For example, using the same data set, Kaniel, Saar, and Titman (2007) investigate retail trading and Boehmer and Kelley (2007) look at the relation between informational efficiency and institutional trading. Boehmer, Jones, and Zhang (2007) analyze differences in the informativeness of short selling across account types.

We match the NYSE data to security information from the Center for Research in Security Prices (CRSP) and obtain daily returns, market capitalization, and consolidated trading volume. Our sample includes only domestic, single-class common stocks. Once a security is delisted or its monthly average price falls below $1 or rises above $999, it is subsequently dropped from the sample. We obtain all primary market prices and quotes from the NYSE Trade and Quote (TAQ) database that satisfy certain criteria.[7] For each stock, we aggregate all trades during the same second that execute at the same price and retain only the last quote for every second if multiple quotes were issued. We require that the monthly average number of daily transactions for a stock be greater than 20. In addition, a stock must have at least 100 trading days to be included in the empirical time-series analysis. This procedure leaves 1,300 different firms over the sample period. Unless noted otherwise, all tests involving daily returns are based on end-of-day quote midpoint returns computed from TAQ. We obtain qualitatively identical results using the standard approach based on close-to-close returns from CRSP, but prefer the midpoint returns to abstract from bid-ask bounce in transaction-price returns.

2.1 Measuring order imbalances

Similar to Chordia, Roll, and Subrahmanyam (2002) and Chordia and Subrahmanyam (2004), we compute three measures of order imbalance for each trader group-stock-day observation: the number of buy transactions less the number of sell transactions of a trader group scaled by the total number of trades, the number of shares bought less the number of shares sold by a trader group scaled by total share volume, and a trader group's dollar volume of buys minus sells scaled by total dollar volume. Scaling order imbalances by a stock's trading activity standardizes the imbalance measures across stocks. We use a volume-based normalization (rather than shares outstanding) for two reasons. First, we believe it is preferable to standardize a flow measure by a flow measure. Second, we wish to abstract from volume effects in order imbalances to better focus on the relative imbalances across different trader groups. We have little theoretical guidance for choosing the time-series dimension for our scaling variables. Throughout our analysis, we use contemporaneous volume, but certain time-series patterns in volume could conceivably affect inferences from these tests. To address this issue, we follow Kaniel et al. (2007) and divide current order imbalances by average annual volume. Repeating our tests using this modified measure leaves our results qualitatively unchanged and therefore these tests are not tabulated.

---

[7] We use trades and quotes only during regular market hours. For trades, we require that TAQ's CORR field is equal to zero, and the COND field is either blank or equal to *, B, E, J, or K. We eliminate trades with non-positive prices or sizes. We also exclude a trade if its price is greater than 150% or less than 50% of the price of the previous trade. We include only quotes that have positive depth for which TAQ's MODE field is equal to 1, 2, 3, 6, 10, or 12. We exclude quotes with non-positive ask or bid prices, or where the bid price is higher than the ask price. We require that the difference between bid and ask be less than 25% of the quote midpoint.

Our measures of order imbalances are similar to those used in Griffin, Harris, and Topaloglu (2003), but differ in important ways from the TAQ-based measures used in Chordia, Roll, and Subrahmanyam (2002) and Chordia and Subrahmanyam (2004). TAQ provides information on executed trades, so by construction there is precisely one share bought for every share sold. Therefore, a direct measure of imbalances between demand and supply is not available – shares bought always equal shares sold. Researchers get around this issue by defining order imbalances in terms of order aggressiveness, which can be inferred from approximate algorithms such as Lee and Ready (1991). Based on this algorithm, a trade executed at a price higher (lower) than the prevailing quote midpoint is classified as a buyer- (seller-) initiated. If the transaction price equals the quote midpoint, it is classified as buyer- (seller-) initiated if the transaction price is above (below) the previous transaction price. This procedure seeks to identify the active side of the trade, that is, the side that is less patient and therefore pays a premium over the quote midpoint. In practice, the active side is likely to be a trader using a marketable order; the passive side could be a limit-order trader or a market maker. Order imbalances based on this algorithm count only the initiating side of the trade and, therefore, provide a measure of the relative impatience of buyers and sellers.

Defining order imbalances in terms of trader aggressiveness has two disadvantages. First, the Lee and Ready (1991) algorithm is known to be somewhat inaccurate. Lee and Radhakrishna (2000) show that 40% of NYSE trades cannot be classified at all, and 7% of the remaining trades are not classified correctly. Second, we need to assume that all traders who intend to achieve a certain portfolio position use marketable orders. While this assumption is relatively innocuous on a trade-by-trade basis, it becomes problematic when traders have longer-term horizons and use different order types to achieve their trading targets. Evidence suggests that traders do indeed use complex strategies to achieve trading objectives. In an experimental study, Bloomfield, O'Hara and Saar (2005) find that traders switch among order types based on the value of their information. Complementing Bloomfield et al's experimental results, Anand, Chakravarty, and Martell (2005) document similar order-switching behavior among informed traders based on TORQ data. Finally, Kaniel and Liu (2006) show that informed traders may prefer to use limit orders depending on the horizon of their information.

This type of order switching affects inferences from TAQ-based order imbalances. To illustrate this point, suppose a portfolio manager decides to accumulate 100,000 IBM shares by the end of the day. To achieve this position, his strategy need not be limited to marketable orders. For example, he might purchase 50,000 shares in the morning using limit orders, which lowers execution costs. Later in the day, to make sure he achieves his target, he may execute the remaining shares using market orders. In this case, TAQ order imbalances would be exactly zero – the 50,000 share morning trades would be classified as seller initiated, and the 50,000 share afternoon trades as buyer initiated. In this example, TAQ-based imbalances clearly obscure the actual institutional imbalance of 100,000 shares because they do not capture the consequences of complex strategies.

In this paper, we use a different approach that is not sensitive to order choice or to misclassification associated with trade-signing algorithms. While our data are also trade-based, in the sense that aggregate demand equals aggregate supply, this is not true within individual trader types. For each trading day and each security, we observe imbalances that reflect the entire buying and selling activity for each trader type, including the specialist. For example, suppose retail buyers purchase N shares from institutions. While the aggregate imbalance is zero, we would observe a retail imbalance of N and an institutional imbalance of $-N$. In line with the evidence in Kaniel and Liu (2006), this approach allows both market and limit orders to affect prices.

2.2 Characteristics of order imbalances

We summarize trading activity and order imbalances for our sample in Table 1. For each trader type, we compute cross-sectional averages of time-series means. Panel A shows that institutions account for the bulk of trading: regular institutional share volume averages 56% of total volume, and program/index arbitrage program trading account for 18.8% and 1.6%, respectively. Retail traders account for 5% of volume, other market makers for 0.7%, and specialists for about 18%. These averages are almost identical in terms of dollar trading volume. Looking at the number of trades, we see that institutional trades tend to be larger than the average, while program trades are somewhat smaller. Consistent with Madhavan and Sofianos (1998), we note that specialists do not always take the opposite side of externally initiated trades, which would imply a participation rate of 50%. Thus, a substantial fraction of trading involves only external market participants.

Panel B of Table 1 reports mean levels of order imbalances for each trader type. Institutions are net buyers over the sample period, whether using regular or program trades (the negative imbalance in terms of transactions indicates that institutional buys tend to be larger than institutional sells). The three remaining groups are net sellers. Panel C of Table 1 presents the mean order imbalances scaled by the corresponding measure of total trading volume of a stock. Again, we observe that institutions are net buyers in terms of share and dollar volume, regardless of order type. One difference to the levels in Panel B is that specialists are net buyers based on scaled order imbalances. This could be due to relatively high buying activity from specialists for less actively traded stocks. If the public tries to sell these less liquid stocks, specialists are more likely to step in to provide liquidity by buying from an outside trader. Consistent with a policy that seeks to minimize inventory, we note that specialists' average imbalance is small relative to those of other traders.

2.3 Cross correlations of order imbalances among trader groups

Table 2 shows imbalance correlations across trader groups. We compute the time-series correlation for each stock and then average across stocks. The three imbalance measures generally provide comparable results, and we note a couple of interesting observations. First, with the exception of index arbitrage trades, specialists' imbalances are negatively correlated with those of each other group. This is what we would expect if their trading is mainly passive, that is, specialists engage in market making activity and provide liquidity when orders arrive. Second, institutions tend to trade in the opposite direction as individuals do. This is consistent with Kaniel, Saar, and Titman's (2007) interpretation that individuals provide liquidity to institutions, although the simple correlations do not reveal whether institutions or retail are the more active side. Third, institutions appear to use regular trades and program trades as substitutes. This suggests that program trades serve a specific purpose – we will return to this issue later on.

The table also shows the correlation between imbalances and contemporaneous returns. Consistently across different measures, specialist imbalances are strongly and negatively correlated with returns. This is again an expected consequence of market making – as other traders buy, for example, they drive up price and specialists sell in the course of liquidity provision. Again consistent with Kaniel, Saar, and Titman's interpretation, individuals also seem to provide liquidity in that their imbalances are negatively correlated with returns. Most interesting are the three institutional order types. Focusing on one of the volume measures in Panel B or C, regular institutional trades and index arbitrage trades are moving with the market. In contrast, program trades are moving against the market. This suggests that institutions use regular orders when they are trading actively. Index arbitrage trades attempt to exploit potentially short-lived price discrepancies between derivative and cash markets; therefore, they are also active trades that move price in the direction of trading. In contrast, institutions appear to use program trades primarily when they are trading passively and therefore program trades seem to provide liquidity. Of course, the correlation evidence presented here is only suggestive and we address each of these issues more rigorously below.

2.4 Persistence of order imbalances

Chordia and Subrahmanyam (2004) report that TAQ-based order imbalances are highly persistent on a daily basis. They suggest that this is because traders split orders to minimize price impact. Order splitting is typically attributed to large traders, such as institutions (Keim and Madhavan, 1995; Chan and Lakonishok, 1995). Table 3 shows evidence consistent with this claim: regular institutional trades and program trades are highly persistent. Individual trades, however, show even stronger persistence, consistent with the Nasdaq evidence in Griffin, Harris, and Topaloglu (2003). We measure the weakest persistence for index arbitrage trades; this makes sense if these trades rely on short-lived intraday price discrepancies between cash markets and derivative markets. Specialists are the only trader type with negatively autocorrelated (volume-based) imbalances. This is consistent with inventory management – when specialists accumulate a long inventory position, for example, they are more likely to sell on the subsequent day.

3. Empirical results

Microstructure theory suggests that informed traders impact stock prices (Kyle, 1985; Glosten and Milgrom, 1985). We can reasonably expect different market participants to be differentially informed and to have different trading motives, and therefore their orders to have different effects on prices. While several studies examine institutional influence on returns (see, for example, Keim and Madhavan, 1995; Chan and Lakonishok, 1995; Griffin, Harris, and Topaloglu, 2003), only a few studies look at retail trading (see Jones and Lipson, 2004;

Kaniel, Saar, and Titman, 2007), and little is known about how recent program trading and specialist activity (see Hendershott and Seasholes, 2006) are related to returns.[8]

In this paper, we analyze the dynamic relation between order imbalances and returns for different trader types. The nature of our data allows us to complement Chordia and Subrahmanyam's (2004) analysis by looking at each trader type separately. First, we test how past price changes affect imbalances. These tests permit inferences on the determinants of order imbalances. Second, we estimate the price impact of imbalances from regressions of returns on contemporaneous imbalances. We make inferences about which traders demand and which traders supply liquidity and, by implication, their private information. Purchases on positive-return days are likely to demand liquidity and be information-motivated, while purchases on negative-return days are likely to supply liquidity (and vice versa for sales). Third, we estimate simple predictive regressions that relate returns to imbalances on the previous day. These tests allow additional inferences about the information of different trader types.

Data availability dictates that we conduct most of our analysis at the daily horizon. We note that this may complicate inferences about the causal relation between order flow and returns, because we cannot precisely observe whether an order reacts to returns, say, during the past hour; or whether prices move as a result of recent order flow. Unfortunately, this issue could not be fully resolved even if one had transactions data with trader type information. The average order size during our sample period is on the order of 500 shares even for institutional traders. This indicates that trading programs are sliced into small pieces to achieve lower execution costs. Because publicly available transactions data do not identify which executed trades belong to the same underlying trading program, using intraday data does not help in identifying the causal relation between order flow and returns. For example, suppose that a trading desk initiates a 100,000 share buy order in response to a price decline on the previous day. As this order is executed in smaller slices over the day, yesterday's negative return may partially revert and create the appearance of intraday momentum trading. Because one cannot identify the individual pieces of this order in the data, a researcher cannot identify when the order was initiated. As a result, it is not possible to

---

[8] Two academic studies on program trading by Harris, Sofianos, and Shapiro (1994) and Hasbrouck (1996) use samples dated around 1990.

uncover the true (contrarian) relation. For this reason, it may indeed be advantageous to use a daily horizon, because many trading programs are designed to complete within a trading day.

On the technical side, our hypothesis tests could be affected by autocorrelated and cross-correlated errors. To alleviate problems associated with correlation over time, we estimate time-series regressions for each stock and conduct inferences on the cross-section means of estimated coefficients. The cross-sectional correlations in most regression specifications turn out to be quite small, but we correct all test statistics for cross-sectional correlation using the procedure described in Chordia and Subrahmanyam (2004).[9]

From here on, we report only results based on dollar-volume imbalances, which we believe best capture the essence of the argument in Kyle (1985) and Glosten and Milgrom (1985) that order imbalances are related to returns. We have repeated all regressions using scaled imbalances defined in terms of transactions and share volume. Our results do not qualitatively change across measures and we note differences where applicable.

3.1 Determinants of order imbalances

To determine how order imbalances on day t depend on past returns, we estimate the following time-series regression for each trader type:

$$OIB_{it} = \alpha_i + \sum_{k=1}^{5} \beta_{ik} R^*_{i,t-k} + \sum_{k=1}^{5} \gamma_{ik} R_{m,t-k} + \sum_{k=1}^{5} \delta_{ik} OIB_{i,t-k} + \varepsilon_{it} \qquad (1)$$

where OIB is the scaled trader-type specific dollar imbalance, R_{m} is the equally-weighted close-to-close midpoint return across all sample stocks, and R_i^* is the residual from a time-series regression of R_i, the close-to-close midpoint returns for stock i, on R_{m}. We employ close-to-close midpoint returns to mitigate the effect of bid-ask bounce on returns, although we obtain qualitatively identical results using returns based on closing prices from CRSP. Decomposing a stock's total returns into market and idiosyncratic returns allows us to separately assess each component's effect on order imbalances. We include lagged order imbalances to control for persistence in this variable.

---

[9] Wilcoxon tests on cross-sectional median coefficients produce similar results throughout our tests and are therefore not tabulated.

We first estimate a restricted variant of Equation (1) that replaces the R_i^* and R_m by the respective 5-day return preceding day *t*. Panel A of Table 4 presents cross-sectional mean coefficient for the restricted model and Panel B presents the unrestricted model. Consistent with Table 3, both regressions show that specialists' order imbalances tend to be negatively autocorrelated and those of all other trader types are positively autocorrelated.

We find that institutions trade as contrarians relative to past returns. In fact, comparing the magnitude of coefficients, institutions show the strongest contrarian response among all trader types when they use regular trades. Contrarian behavior with respect to security-specific past returns is less pronounced when institutions use program trades, and it is not visible when they engage in index arbitrage. Our results contrast to Griffin, Harris, and Topaloglu's (2003) findings that institutions are trend chasers at daily horizons. But our results are consistent with Lipson and Puckett (2005), who study pension fund order imbalances on volatile days and find that pension funds are contrarian traders. They are also consistent with the evidence in Chordia, Roll and Subrahmanyam (2002), who find that aggregate trade-based order imbalances are contrarian. In Panel B, we show that the contrarian behavior is primarily driven by returns on the previous two days.

Similar to Kaniel, Saar, and Titman (2007), we find that individuals trade as contrarians relative to a stock's returns during the previous week. Only specialists trade in the direction of previous-week returns, apparently in response to the contrarian demand by the other trader types.

While regular institutional and individual imbalances are not sensitive to market returns, we find that institutional program and index arbitrage imbalances are contrarian with respect to market returns. In fact, these imbalances are more sensitive to market returns than to idiosyncratic returns. This is a notable and intuitive result, suggesting that institutions use program trades to respond to market movements.

Panel C demonstrates that these return effects are generally present across size quartiles (size quartiles are based on the time-series average market value of equity). Only trading in the largest firms (quartile 4) has a somewhat different relation to past returns. In this quartile, institutions and individuals are still contrarian with respect to idiosyncratic returns, but they are momentum traders with respect to market returns. Indeed, the counteracting influences of market and security returns in the top size quartile could potentially reconcile the differences between our results and those in Griffin, Harris, and Topaloglu, because their analysis does not allow market returns to affect order imbalances directly.[10] Their sample consists of all Nasdaq 100 stocks between May 1, 2000 and February 28, 2001. During this period, the Nasdaq 100 index declined by 50.7%. If institutional trading decisions depend on market returns as in Panel C, this pronounced decline should prompt large negative institutional imbalances for large stocks. Put differently, we would expect a positive correlation between security returns and imbalances during this period. It is, therefore, possible that the trend-chasing behavior documented in Griffin, Harris, and Topaloglu is driven by selling due to these large market-wide price moves and not a response to security-specific returns.

Finally, we note that our results are related to the cross-sectional institutional momentum patterns documented at the quarterly horizon (see Grinblatt, Titman, and Wermers, 1995; Sias, 2005). One could view these longer-horizon results as describing institutional investment decisions, while our results describe institutional trading decisions. This distinction is important for two reasons. First, the models used for long-term investment decisions could differ from those applied to daily trading decisions; so their relation to prior returns need not be the same. Second, information on institutional holdings is only available with quarterly frequency. While this is sufficient to characterize institutional investment decisions, our findings illustrate that higher-frequency information is desirable when analyzing institutional trading decisions.

3.2 Price impact: order imbalances and contemporaneous stock returns

In this section, we ask how daily order imbalances are related to contemporaneous returns. We use this analysis to make inferences about differences in how imbalances affect subsequent price changes and, by implication, about differences in informedness and liquidity provision across trader types. In general, traders with short-lived information need to trade actively so their orders execute before their information gets otherwise impounded into prices. Impatient, active traders tend to move prices in the direction of the order. For example, a

---

[10] Griffin, Harris, and Topaloglu control for market movements by regressing order imbalances on excess returns, defined as security returns net of market returns. When we repeat this approach on our data, the coefficient on excess returns are very similar to those reported in Table 4. In particular, institutional imbalances are still significantly negatively related to past (excess) returns. Therefore, allowing the coefficient on market returns to vary does not cause the different results. We also obtain similar results when we include unadjusted security returns (and omit market returns).

market buy order should lead to a price increase. In contrast, patient traders can afford to trade passively. For example, a limit buy order that is priced below the current ask price only executes once prices have sufficiently declined. Upon execution, the active part of this trade (a sell order) has exerted downward pressure on price. In this case, the buyer provides liquidity to the market. In general, we expect a stronger positive relation between imbalances and returns for trader types who, on average during a trading day, are more informed and trade aggressively because their information is short-lived; we expect a negative relation for trader types who, on average, trade passively and supply liquidity to the market.

An important caveat to this analysis is that we cannot directly establish a causal relation between order imbalances and returns at daily horizons. We believe, however, that this is not a serious concern for our analysis. First, microstructure theory suggests a causal link from order flow to prices, and this holds in both inventory models (Stoll, 1978; Ho and Stoll, 1981) and asymmetric-information models (Glosten and Milgrom, 1985; Kyle, 1985). Importantly, these canonical models show that price innovations are a function of order flow innovations and not the other way around. Second, there is little empirical evidence that suggests reverse causality. In fact, using NYSE system order data for November 2002, Jones and Lipson (2004) show that institutions do not chase intraday short-term trends. This alleviates concerns about reverse causality in our sample, because Jones and Lipson's sample (period) is representative of ours.[11]

To examine how prices are related to contemporaneous order imbalances, we estimate the following regression model:

$$R_{it} = \alpha_i + \beta_i R_{mt} + \sum_{k=0}^{4} \gamma_{ik} OIB_{i,t-k} + \varepsilon_{it} \qquad (2)$$

where the variables are as defined in Equation (1). We believe it is important to control for market returns, because we would like to capture return movements that are idiosyncratic to the order imbalances we examine.

[11] As we discuss on page 12, it is not possible to resolve this issue empirically even with intraday transactions data.

We also repeat the estimation with a different risk adjustment and use the three Fama-French factors instead of market returns. The results are qualitatively identical and therefore not reported.[12]

3.2.1 Basic price impact results

Table 5 contains results on the relation between order imbalances and contemporaneous price changes. We report cross-sectional averages of the security-specific time-series coefficients. Controlling for persistence in order imbalances and market returns, the coefficient on contemporaneous institutional imbalances is positive. Thus, institutional buying is associated with a greater price increase than implied by the simple market-model adjustment, and institutional selling is associated with a greater price decline. This positive relation is consistent with information-based trading, but it could also be a temporary response to uninformed imbalances. In the latter case, however, we would expect a quick reversal that we do not observe in the data. Specifically, the positive coefficient of OIB_{t-1} implies that price impacts do not revert on the next trading day. This suggests that information-driven order imbalances drive the price impacts we document.

The mean effect of institutional imbalances on returns depends on the order type institutions use. While regular institutional trading has a positive contemporaneous effect on prices, institutional program trade imbalances have a negative effect, and index arbitrage imbalances have no effect. These observations make economic sense. It is unlikely that program trades of either type are motivated by private information about individual securities, because they involve simultaneous orders in at least fifteen different securities.[13] Buy-side institutions often use program trades to change the scale of their portfolio. For example, when institutions experience inflows or outflows, they do not immediately change the position in every stock in their portfolio. Instead, they prefer to trade those stocks where they obtain the best price. This could be achieved by a passive trading strategy that places a set of limit orders for a range of stocks (which, for sufficient size and at least fifteen stocks, would be classified as a program strategy). Depending on which orders execute, the institutions can then

---

[12] Market returns are correlated with imbalances, as shown in Table 4, so model (2) is subject to multicollinearity. As a robustness check, we estimate a regression without market adjustment and another regression of excess returns (over R_m) on current and lagged order imbalances and obtain qualitatively identical results in both cases. Therefore, we present results from (2) and do not impose the restriction that the coefficient on R_m equal one.

[13] Strategies such as "pairs trading," where traders attempt to arbitrage price differences between two similar securities often involve informed traders. But pairs trading would not generally be classified as program trading.

cancel the remaining orders and/or resubmit new ones to remain close to its desired target portfolio. Such a strategy would supply liquidity to the market, consistent with a negative price impact for program trades.

Individuals, specialists, and other market makers' imbalances have a significantly negative association with contemporaneous returns. These trader types appear to provide liquidity to the active institutional traders. The negative price impact for specialists and other market makers is what we expect from bona fide market making activities and is in line with the results in Hendershott and Seasholes (2006). The negative price impact for individuals suggests that they buy when prices are falling and sell when prices are rising. As suggested by Kaniel, Saar and Titman (2007), this trading (against the market) pattern makes individuals natural liquidity providers to institutions. They note that liquidity provision from individuals does not necessarily mean that individuals intentionally trade like market makers to profit from making two-sided markets. Instead, it may be the case that individuals happen to take the other side of the market when institutional trades start moving prices. It is important that the negative contemporaneous relation with returns does not imply that individuals and market makers lose because they may receive compensation for providing liquidity (Kaniel, Saar and Titman, 2007).

Panel B of Table 5 shows average price impact coefficients for the same model, but computed separately for each size quartile. Coefficients for individuals and market makers are largely consistent with those in Panel A, but the disaggregation reveals some differences in institutional price impacts across size quartiles. Institutional imbalances move with prices in the smallest quartile whether they arise from regular or program trades. It is well known that institutions concentrate their holdings in larger firms, so institutional trading in the smallest quartile could be dominated by information-based active traders. In the largest quartile, information-based institutional traders, who have positive price impact, may co-exist with passive institutional traders, including program traders, who have a negative coefficient. We have no good explanation for why institutional price impacts are zero in the middle quartiles – perhaps the effect of informed order flow is dominated by the effect of liquidity-motivated orders. Overall, we note that only individuals and market makers consistently do *not* move prices in the direction of their imbalances.

3.2.2 Price impacts and information asymmetry

Keim and Madhavan (1995) document considerable heterogeneity in trading styles based on past price movements. Some institutions pursue trend-chasing strategies while others tend to adopt contrarian strategies; thus, the overall effect of institutional trading strategies on contemporaneous prices could differ substantially across institutions. Unfortunately, we do not have information on trader identity beyond the account types and cannot differentiate between institutions likely to trade frequently on private information (perhaps hedge funds and other active traders) and others (such as index funds). This naturally makes it difficult to isolate information-based trading by looking at institutions as a group.

One way to address the heterogeneity across institutions empirically is to relate firm-specific price impact coefficients to cross-sectional variation in security characteristics. In particular, if the positive coefficients arise because of information-based trading, we would expect them to be larger for firms that are characterized by greater information asymmetry. Following Llorente, Michaely, Saar, and Wang (2002), we regress the price impact coefficients (the coefficient on *OIB (t)* in Panel A of Table 5) on proxies for information asymmetry.

We measure information asymmetry in two different ways. First, we calculate intraday proxies for information asymmetry. Higher information asymmetry is typically associated with greater relative effective spreads, defined as twice the absolute difference between the execution price and the quote midpoint prevailing when the trade was reported (see Bessembinder, 2003). However, effective spreads also increase for non-informational reasons. We attempt to separate the two effects by decomposing effective spreads into a temporary component (realized spreads) and a permanent component (the trade-to-trade price impact). Realized spreads represent short-term effects (for example, due to inventory management). They are unrelated to information and are measured as the price change from the trade price to the quote midpoint five minutes after the trade (multiplied by -1 for sell-signed trades). Trade-by-trade price impacts, the permanent effects, provide an estimate of the degree of informed trading in a security. They are defined as the change in quote midpoints from just before a trade to five minutes afterwards. We compute daily equally-weighted averages of these variables and use their time-series averages as regressors. If information asymmetries are driving institutional price impact coefficients, we expect that they are positively related to the permanent component of spreads in the cross-section of stocks.

Second, we use a firm's R&D expenditures and the value of intangible assets to proxy for information asymmetry. Firms with high R&D or a large proportion of intangible assets are more difficult to value, and they are subject to greater information asymmetry. We obtain data on R&D investments (scaled by sales) and intangible assets (scaled by total assets) for the year ending in December 1999, the most recent reporting date that precedes our sample period. Several firms have missing values for these variables and it is not always clear whether the data item is truly missing or actually zero. In our analysis, we assume that all missing items represent zeroes but add a dummy variable that is one for observations with non-missing R&D data, and zero otherwise.

The estimates in Table 6 are broadly consistent with an information-based explanation for institutional price impact coefficients. Coefficients on effective spreads are positive in Panel A, controlling for firm size. Thus, the price impact of institutional order imbalances increases with information asymmetry associated with larger effective spreads. When we decompose spreads into temporary and permanent components, only the permanent component is related to institutional price impacts. Therefore, informed trading activity appears to be the main driver of institutional price impacts.[14] Panel B presents a similar picture. While price impacts are unrelated to intangibles, we find significantly larger price impacts in firms with higher R&D expenditures. This suggests that the greater information asymmetry associated with R&D investments makes prices more sensitive to the information in institutional trading decisions. Taken together, both results support the interpretation that the positive contemporaneous relation between institutional trading and price changes is mainly attributable to the information revealed by the institution's order flow.

3.3 Predictability: order imbalances and future stock returns

A different way to evaluate the average information advantage of particular trader groups is to estimate return movements on the day following an order imbalance. Chordia and Subrahmanyam (2004) show that trade-

---

[14] Our inferences from Panel A in Table 6 could be tautological. We can only interpret effective spreads as proxies for information asymmetry if we *assume* that informed trading moves prices in the appropriate direction. Therefore, given the price impact coefficients estimated in Table 5, a negative coefficient on effective spreads in Table 6 would be difficult to explain. Nevertheless, we find it helpful to demonstrate that these presumed relations exist in the data. In unreported tests we regress price impact coefficients on effective spreads (and their components) measured at the beginning of our sample period. This avoids the endogeneity issue and we obtain qualitatively similar results (RES and the permanent component have significantly positive coefficients, although the temporary component also becomes significantly positive). This alternative specification suggests that the relation in Table 6 Panel A is not substantially driven by a tautology.

based order imbalances predict next-day returns. In this section, we investigate which trader types are driving this predictability.

While return predictability is consistent with private information, a positive relation between current imbalances and future returns could also arise if traders split their order across days and the resulting autocorrelation in imbalances is not immediately reflected in prices. Evidence in Chordia, Roll, and Subrahmanyam (2005) shows, however, that at least for large stocks this is not the case – this type of information is rapidly impounded into prices. Despite predictability in imbalances over several days, they find little evidence of predictability in returns for intervals longer than about 30 minutes. Therefore, it is unlikely that order splitting alone could drive a positive relation between imbalances and subsequent returns at the daily horizon. We also note that any predictability based on trader-group specific imbalances could not be exploited by market participants, because information on group-specific order flow is not publicly disclosed.[15]

We estimate the following predictive regressions:

$$R_{i,t} = \alpha_i + \beta_i R_{mt} + \sum_{k=1}^{5} \gamma_{ik} OIB_{i,t-k} + \varepsilon_{it} \ , \tag{3}$$

where the variables are as defined in Equation (1). Similar to the forecast regressions used in Chordia and Subrahmanyam (2004), we regress a stock's return on five lags of a trader type's imbalances and the market return. We obtain qualitatively similar results when we add lagged security returns as explanatory variables, or when we use excess returns over market as the dependent variables (and omit market return on the right hand side).

Table 7 reports the results. While Griffin, Harris, and Topaloglu (2003) find no predictability on Nasdaq, we show that non-program institutional imbalances predict next-day returns in our sample. This suggests, largely consistent with the estimates reported in Chordia and Subrahmanyam (2004), that institutions have some information about future returns on NYSE-listed stocks. It is also consistent with Boehmer, Jones, and Zhang (2007), who show that institutional shorting activity is more informative than shorting by other trader types. But

[15] The specialist can observe whether an order is part of a program trade, but cannot see any of the other account types.

institutional imbalances only contain information when they result from regular trades. When institutions decide to use program trades, their imbalances are not informative. This corroborates our earlier argument that institutions use program trades primarily for liquidity-motivated trading.[16]

Consistent with Kaniel, Saar, and Titman (2007), we also observe that individuals' imbalances have predictive ability. Their imbalances are, on average, informative about returns during the next few days. Specialists do not appear to have private information or cannot trade to exploit it – their market making function implies that they buy in declining markets and sell in rising markets to satisfy the trading demand of other market participants. As a result, their imbalances are negatively related to next-day returns. In Panel B, we compute separate coefficients for each size quartile and find largely similar results. In particular, institutional and individual imbalances tend to predict next-day returns, while program and specialist imbalances do not.

A somewhat puzzling aspect of our findings is that regular institutional imbalances and retail imbalances both predict next-day returns. Because individuals have negative contemporaneous price impacts (Table 5), they appear to trade rather passively and it is surprising that their imbalances predict returns. In an attempt to resolve this apparent inconsistency, we modify Equation (3) by decomposing next-day returns into a non-trading (overnight) period and a trading period. Specifically, we compute overnight returns from the closing midpoint on day t to the opening midpoint on day t+1, and compute the trading-day return from the opening midpoint on day t+1 to the close on that day. Market returns, an independent variable, are decomposed accordingly. This procedure allows a more detailed look at the dynamics of the order imbalance-return relation.

Table 8 reveals interesting contrasts between the predictive power of regular institutional and individual order imbalances. The negative effect of -0.0005 that regular institutional imbalances have on overnight returns

---

[16] Our result that program trades are negatively related to contemporaneous and future returns differ from earlier findings by Harris, Sofianos, and Shapiro (1994) and Hasbrouck (1996). Based on intraday data, they both argue that program trades contain information. The differences could be due to different samples and different periods. Harris et al. use aggregate information on program trades from 1989 to 1990, and Hasbrouck uses program trades on a small sample of firms over three months from November 1990. It is likely that trading strategies have since changed, especially the use of limit-order strategies. Specifically, specialists had no obligation to display limit orders immediately during the periods these studies analyze, which probably made limit orders less attractive to traders. But without limit orders, a main argument for the liquidity-supplying nature of today's program trades does not apply. During our 2000-2004 sample period, limit orders are the dominant order type and their usage has increased after the NYSE started to display its order book publicly in 2002 (see Boehmer, Saar, and Yu, 2005).

(Panel A) suggests a slight overnight return reversal. But these reversals are very small – compared to the same-day price impact coefficient in Table 5 (0.0028), the overnight reversal is an order of magnitude smaller. Moreover, the positive and large coefficient of 0.0027 in Panel B shows that yesterday's institutional order imbalances can predict the subsequent trading-day return. Consistent with our previous analysis, these dynamics again suggest that regular institutional imbalances are largely driven by information. As a result, they have a positive price impact and predictive power for next-day returns.

Individual order imbalances, on the other hand, are associated with quite different return patterns. Their negative contemporaneous price impact (Table 5) is also slightly reversed overnight, but individual imbalances have no effect on returns during the next trading period. The more striking finding is that the apparent predictive power of individual imbalances that we observe in Table 7 (with a coefficient of 0.0075) accrues entirely overnight (with a coefficient of 0.0076 in Panel A of Table 8). This suggests that individual imbalances are strongly related to temporary price changes, consistent with Kaniel, Saar, and Titman's (2007) interpretation that individuals provide liquidity to institutions.

3.4 Additional robustness checks

We perform additional tests to check the robustness of our main results. First, our main analysis treats order flow across trader types as independent, because we rely on separate models for each type. This approach could be misleading if the dependent variables (order imbalances or returns) are correlated across trader types. To address this issue, we re-estimate the models in Tables 4, 5, and 7 using seemingly unrelated regression (SUR) models. An additional issue arises when explaining order imbalances (Table 4) because order imbalances in a stock sum to zero across trader types. We add this identity to the system to assure that predicted values satisfy this restriction. This makes the modified system singular and we arbitrarily omit the equation describing order imbalances of "other market makers" to identify it. Inferences from these SUR models are qualitatively identical to those in Tables 4, 5, and 7 and are not tabulated.[17]

---

[17] This estimation approach is similar to a classic consumer demand model where expenditure-share equations are restricted to sum up to unity.

Second, our models in Table 4 (determinants of order imbalances) and Table 7 (determinants of future returns) can be viewed, with slight modifications, as components of a vector autoregressive (VAR) model. This perspective captures the dynamic relations among returns and order imbalances and explicitly allows for lagged endogenous effects. Specifically, for each trader type, we estimate the following system of equations:

$$\begin{aligned} OIB_{i,t} &= \alpha_i + \sum_{k=1}^{5} \beta_{ik} OIB_{i,t-k} + \sum_{k=1}^{5} \gamma_{ik} R_{i,t-k} + \varepsilon_{i,t} \\ R_{i,t} &= \alpha_i + \sum_{k=1}^{5} \lambda_{ik} OIB_{i,t-k} + \sum_{k=1}^{5} \eta_{ik} R_{i,t-k} + \delta_{i,t} \end{aligned} \tag{4}$$

where $OIB_{i,t}$ is the scaled dollar imbalance of stock i on date t and $R_{i,t}$ is the quote midpoint return net of market returns (we obtain similar results with raw returns). In another test, we estimate the dynamic structural equations of the above VAR model for each trader type that includes contemporaneous *OIB* in the return equation (Table 5). Finally, we examine both the coefficients and the cumulative impulse responses to orthogonalized shocks to these systems. For each of these tests, inferences about the behavior of order imbalances and the ability of different trader types to predict returns remain unchanged.

Third, our price-impact analysis in Table 5 seeks to estimate the price impact that order imbalances of different trader types have. This analysis relates two variables that are measured over different horizons: we compute order imbalances from executed orders during official trading hours, but measure returns from close to close. As a result, we could mistakenly interpret trader reactions to overnight returns as price impacts. To address this concern, we repeat the price-impact analysis in Table 5 using same-day returns computed during trading hours only (measured from the opening midpoint to the closing midpoint). The results are almost identical to the ones using close-to-close returns: regular institutional trades have positive price impact, while regular program trades, individuals and specialists appear to trade against the market. These results show that the relation between imbalances and returns arises during trading hours, which corroborates our earlier inferences that institutions are likely to possess short-lived information and their trading impacts contemporaneous returns.

4. Conclusions

Microstructure theory predicts that order flow affects prices (Kyle, 1985; Glosten and Milgrom, 1985). While this prediction is empirically well supported, we know little about which traders drive this relation. Trading strategies and information differ across traders and, therefore, we also expect that the relation between order flow and prices differs across traders. We provide new evidence on this issue using proprietary NYSE data on daily order imbalances for different trader groups. For all common stocks between 2000 and April 2004, we observe buys and sells for institutions, individuals, and market makers, and can further distinguish regular institutional trades from institutional program and index-arbitrage program trades. Institutions account for 77% of total share volume during this period, individuals for 5%, and specialists for about 18%. Thus, institutions clearly are the most important trader group.

First, we document that institutions are contrarians with respect to returns on the previous day. This finding contrasts to evidence based on quarterly holdings, which suggests that institutions are momentum traders at longer horizons (see Sias, 2005). These results are not mutually inconsistent; but because momentum trading would arguably be most destabilizing at shorter horizons, our results appear to alleviate such concerns about institutional trading behavior. We further show that individuals are contrarians as well, consistent with Kaniel, Saar, and Titman (2007). In fact, only specialists trade as if they are momentum traders on a daily basis – but this is a plausible result of bona fide market-making activity. A positive-return day is typically characterized by positive order imbalances and market makers may need to short to satisfy this demand. When returns reverse on the next day, they can purchase shares to rebalance their inventory.

Second, we document that order imbalances from different trader types play distinctly different roles in price formation. While institutions and individuals are contrarians, they differ in the effect their order imbalances have on contemporaneous returns. Institutional imbalances are positively related to contemporaneous returns, and we provide cross-sectional evidence that this relation is likely to be the result of firm-specific information institutions have. In contrast, the imbalances of individuals, specialists, and institutional program traders are negatively related to contemporaneous returns. This suggests that these trader types provide liquidity to actively trading institutions. Moreover, this result suggests a special role for institutional program trades. Institutions appear to choose regular trades when they have firm-specific information, but they choose program trades when they do not and can, therefore, afford to trade passively. As a result, program trades provide liquidity to the market.

Third, both institutional non-program and individual imbalances have predictive power for next-day returns. In contrast, specialist and program trade imbalances are negatively related to next-day returns. These results do not imply that profitable trading strategies exist, because trader-type information is not publicly (or even privately) disseminated. They do suggest, however, that institutional order imbalances often arise from private information. But because these imbalances also move prices contemporaneously, institutional trading profits appear to be bounded. This scenario is broadly consistent with evidence that institutions have some stock-picking ability (see, for example, Daniel, Grinblatt, Titman, and Wermers, 1997) and that institutions improve the informational efficiency of share prices (see Boehmer and Kelley, 2007). Moreover, our results also suggest that institutions use program trades when they do not have private information. This makes intuitive sense, because by packaging orders into baskets institutional traders can signal to the market that they are uninformed, which should result in lower execution costs.

During our sample period, institutional non-program trades generate 56% of share volume in the average stock. Our results imply that this portion of trading activity tends to be more informed than other trades. Therefore, institutional trading appears to drive the generally positive relation between order flow and prices. Individuals provide 5% of volume and trade against the market, consistent with Kaniel, Saar, and Titman's (2007) interpretation that they provide liquidity to institutions. Their order volume is far too small, however, to satisfy institutional imbalances. We find that the remainder of these imbalances is filled by market makers and, in particular, by other institutions who are apparently not privately informed and use program trades.

References

Admati, Anat and Paul Pfleiderer, 1988, A theory of intraday spread patterns: Volume and price variability, *Review of Financial Studies* 1, 3-40.

Anand, Amber, Sugato Chakravarty, and Terrence Martell, 2005, Empirical evidence on the evolution of liquidity: Choice of market vs. limit orders by informed and uninformed traders, *Journal of Financial Markets* 8, 289-309

Amihud, Yakov and Haim Mendelson, 1980, Dealership market: Market making with inventory, *Journal of Financial Economics* 8, 31-53.

Back, Kerry and Shmuel Baruch, 2007, Working orders in limit-order markets and floor exchanges, *Journal of Finance*, forthcoming.

Baker, Malcolm and Jeremy Stein, 2004, Market liquidity as a sentiment indicator*, Journal of Financial Markets* 7, 271-299.

Bessembinder, Hendrik, 2003, Issues in assessing trade execution costs, *Journal of Financial Markets* 6, 233-257.

Bloomfield, Robert, Maureen O'Hara, and Gideon Saar, 2005, The "make or take" decision in an electronic market: evidence on the evolution of liquidity, *Journal of Financial Economics* 75, 165-199.

Boehmer, Ekkehart, Charles Jones, and Xiaoyan Zhang, 2007, Which shorts are informed? *Journal of Finance*, forthcoming.

Boehmer, Ekkehart and Eric Kelley, 2007, Institutional investors and market efficiency, Working paper, Texas A&M University.

Boehmer, Ekkehart, Gideon Saar, and Lei Yu, 2005, Lifting the veil: An analysis of pre-trade transparency at the NYSE, *Journal of Finance* 60, 783-815.

Campbell, John, Sanford Grossman, and Jiang Wang, 1993, Trading volume and serial correlation in stock returns, *Quarterly Journal of Economics* 108, 905-939.

Chan, Louis K.C. and Josef Lakonishok, 1995, The behavior of stock prices around institutional trades, *Journal of Finance* 50, 1147-1174.

Choe, Hyuk, Bong-chan Kho, and Rene Stulz, 1999, Do foreign investors destabilize stock markets? The Korean experience in 1997, *Journal of Financial Economics* 54, 227-264.

Chordia, Tarun, Sahn-wook Huh, and Avanidhar Subrahmanyam, 2007, The cross-section of expected trading activity, *Review of Financial Studies* 20, 709-740.

Chordia, Tarun, Richard Roll, and Avanidhar Subrahmanyam, 2001, Market liquidity and trading activity, *Journal of Finance* 56, 501-530.

Chordia, Tarun, Richard Roll, and Avanidhar Subrahmanyam, 2002, Order imbalance, liquidity and market returns, *Journal of Financial Economics* 65, 111-130.

Chordia, Tarun, Richard Roll, and Avanidhar Subrahmanyam, 2005, Evidence on the speed of convergence to market efficiency, *Journal of Financial Economics* 76, 271-292.

Chordia, Tarun and Avanidhar Subrahmanyam, 2004, Order imbalance and individual stock returns: Theory and evidence, *Journal of Financial Economics* 72, 485–518.

Daniel, Kent, Mark Grinblatt, Sheridan Titman, and Russ Wermers, 1997, Measuring mutual fund performance with characteristic-based benchmarks, *Journal of Finance* 52, 1035-1058.

Easley, David and Maureen O'Hara, 1987, Price, trade size, and information in securities markets, *Journal of Financial Economics* 19, 69-90.

Garman, Mark, 1976, Market microstructure, *Journal of Financial Economics* 3, 257-275.

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

Griffin, John, Jeffery Harris and Selim Topaloglu, 2003, The dynamics of institutional and individual trading, *Journal of Finance* 58, 2285-2320.

Griffin, John, Jeffery Harris, and Selim Topaloglu, 2005, Who drove and bust the tech bubble? Working paper, University of Texas at Austin.

Grinblatt, Mark and Matti Keloharju, 2000, The investment behavior and performance of various investor types: a study of Finland's unique data set, *Journal of Financial Economics* 55, 43-67

Grinblatt, Mark, Sheridan Titman, and Russ Wermers, 1995, Momentum investment strategies, portfolio performance, and herding: A study of mutual fund behavior, *American Economic Review* 85, 1088-1105.

Harris, Lawrence, George Sofianos, and James Shapiro, 1994, Program trading and intraday volatility, *Review of Financial Studies* 7, 653-685.

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

Hasbrouck, Joel, 1991a, Measuring the information content of stock trades, *Journal of Finance* 46, 178-208.

Hasbrouck, Joel, 1991b, The summary informativeness of stock trades: An econometric analysis, *Review of Financial Studies* 4, 571-595.

Hasbrouck, Joel, 1996, Order characteristics and stock price evolution: an application to program trading, *Journal of Financial Economics* 41, 129-149.

Herdershott, Terrence and Mark Seasholes, 2006, Market maker inventories and stock prices, *American Economic Review Papers and Proceedings*, forthcoming.

Ho, Thomas and Richard Macris, 1984, Dealer bid-ask quotes and transaction prices: An empirical study of some AMEX options, *Journal of Finance* 39, 23-45.

Ho, Thomas and Hans R. Stoll, 1981, Optimal dealer pricing under transactions and return uncertainty, *Journal of Financial Economics* 9, 47-73.

Holden, Craig W. and Avanidhar Subrahmanyam, 1992, Long-lived private information and imperfect competition, *Journal of Finance* 47, 247-270.

Jones, Charles and Marc Lipson, 2004, Are retail orders different? Working paper, Columbia University.

Kaniel, Ron and Hong Liu, 2006, So what orders do informed traders use? *Journal of Business* 79, 1867-1913.

Kaniel, Ron, Gideon Saar, and Sheridan Titman, 2007. Individual investor trading and stock returns, *Journal of Finance*, forthcoming.

Karpoff, Jonathan, 1987, The relation between price changes and trading volume: A survey, *Journal of Financial and Quantitative Analysis* 22, 109-126.

Keim, Donald and Ananth Madhavan, 1995, Anatomy of the trading process: Empirical evidence on the behavior of institutional traders, *Journal of Financial Economics* 37, 371-398.

Kyle, Albert, 1985, Continuous auctions and insider trading, *Econometrica* 53, 1315-1336.

Lee, Charles, 1992, Earnings news and small traders: An intraday analysis, *Journal of Accounting and Economics* 15, 265-302.

Lee, Charles and Balkrishna Radhakrishna, 2000, Inferring Investor Behavior: Evidence from TORQ, *Journal of Financial Markets* 3, 83-111.

Lee, Charles and Marc Ready, 1991, Inferring trade direction from intraday data, *Journal of Finance* 46, 733-746.

Lee, Yi-Tsung, Yu-Jane Liu, Richard Roll, and Avanidhar Subrahmanyam, 2004, Order imbalances and market efficiency: Evidence from the Taiwan Stock Exchange, *Journal of Financial and Quantitative Analysis* 39, 327-341.

Lipson, Marc and Andy Puckett, 2005, Volatile markets and pension fund trading, Working paper, University of Georgia.

Llorente, Guillermo, Roni Michaely, Gideon Saar, and Jiang Wang, 2002, Dynamic Volume-Return Relation of Individual Stocks, *Review of Financial Studies* 15, 1005-1047.

Madhavan, Ananth and George Sofianos, 1998, An empirical analysis of NYSE specialist trading, *Journal of Financial Economics* 48, 189-210.

Sias, Richard, 2005, Reconcilable differences: Momentum trading by institutions, Working paper, Washington State University.

Spiegel, Matthew and Avanidhar Subrahmanyam, 1992, Informed speculation and hedging in a noncompetitive securities market, *Review of Financial Studies* 5, 307-330.

Stoll, Hans R., 1978, The supply of dealer services in securities markets, *Journal of Finance* 33, 1133-1151.

Table 1. Summary Statistics.

We present cross-sectional averages of time-series means for 1,300 NYSE common stocks from January 2000 to April 2004. Panel A shows the fraction of trading volume of each trader type. Panel B presents the level of order imbalances (measured as the difference between buys and sells) by trader types. Panel C presents each trader type's order imbalances scaled by the total trading activity of each stock each day.

	Institutions	Regular program trades (institutional)	Index arbitrage program trades (institutional)	Individuals	Specialists	Other market makers
Panel A: Relative trading volume of each trader type						
% of transactions	45.29%	27.03%	3.52%	5.26%	18.00%	0.91%
% of share volume	56.04%	18.77%	1.61%	5.01%	17.89%	0.68%
% of dollar volume	56.04%	18.77%	1.61%	5.01%	17.89%	0.68%
Panel B: Level of order imbalances by trader types						
Order imbalances in number of transactions	-12	9	4	-5	-5	-1
Order imbalances in shares	3,032	5,007	1,034	-4,696	-311	-538
Order imbalances in dollar volume	150,686	190,623	43,317	-205,093	-10,543	-40,025
Panel C: Scaled order imbalances by trader types						
Scaled order imbalances in transactions / number of trades	-1.26%	1.03%	0.38%	-1.47%	-0.23%	-0.21%
Scaled order imbalances in shares / share volume	0.71%	0.85%	0.08%	-1.48%	0.08%	-0.15%
Scaled order imbalances in dollars / dollar volume	0.69%	0.83%	0.07%	-1.49%	0.14%	-0.15%

Table 2. Cross Correlations across Trader Types.

We report cross-sectional averages of time-series correlations. The sample includes 1,300 NYSE common stocks from January 2000 to April 2004. Order imbalances refer to differences between buys and sells. Return is a stock's daily close-to-close quote-midpoint returns. Panel A reports order imbalances measured in transactions standardized by the total number of transactions. Panel B reports order imbalances measured in shares standardized by total share volume. Panel C reports order imbalances measured in dollars standardized by total dollar volume.

	Institutions	Regular program trades (institutional)	Index arbitrage program trades (institutional)	Individuals	Specialists	Other market makers	Return
Panel A: Order imbalances measured in transactions standardized by the total number of transactions							
Institutions	1.00	-0.20	-0.12	-0.08	-0.27	-0.02	-0.02
Regular program trades (institutional)		1.00	0.13	-0.11	-0.43	-0.05	-0.03
Index arbitrage program trades (institutional)			1.00	-0.06	-0.21	-0.05	0.11
Individuals				1.00	-0.10	0.19	-0.05
Specialists					1.00	-0.04	-0.17
Other market makers						1.00	-0.10
Return							1.00
Panel B: Order imbalances measured in shares standardized by total share volume							
Institutions	1.00	-0.24	-0.08	-0.13	-0.21	-0.05	0.01
Regular program trades (institutional)		1.00	0.07	-0.06	-0.05	-0.03	-0.02
Index arbitrage program trades (institutional)			1.00	-0.03	0.00	-0.02	0.09
Individuals				1.00	-0.04	0.10	-0.06
Specialists					1.00	0.02	-0.25
Other market makers						1.00	-0.08
Return							1.00
Panel C: Order imbalances measured in dollars standardized by total dollar volume							
Institutions	1.00	-0.24	-0.08	-0.13	-0.21	-0.05	0.01
Regular program trades (institutional)		1.00	0.07	-0.06	-0.05	-0.03	-0.02
Index arbitrage program trades (institutional)			1.00	-0.03	0.01	-0.02	0.08
Individuals				1.00	-0.04	0.10	-0.06
Specialists					1.00	0.02	-0.25
Other market makers						1.00	-0.08
Return							1.00

Table 3. Persistance of Order Imbalances.
We report cross-sectional averages of time-series autocorrelations. The sample includes 1,300 NYSE common stocks from January 2000 to April 2004. Panel A reports order imbalances measured in transactions standardized by the total number of transactions. Panel B reports order imbalances measured in shares standardized by total share volume. Panel C reports order imbalances measured in dollars standardized by total dollar volume.

	Institutions	Regular program trades	Index arbitrage program trades	Individuals	Specialists	Other market makers
Panel A: Order imbalances measured in transactions standardized by the total number of transactions						
lag1	0.26	0.32	0.09	0.45	0.17	0.21
lag2	0.15	0.20	0.07	0.37	0.10	0.17
lag3	0.11	0.15	0.08	0.33	0.08	0.15
lag4	0.08	0.12	0.06	0.31	0.07	0.14
lag5	0.06	0.08	0.01	0.29	0.04	0.13
Panel B: Order imbalances measured in shares standardized by total share volume.						
lag1	0.21	0.29	0.04	0.27	-0.14	0.14
lag2	0.12	0.18	0.04	0.20	-0.03	0.11
lag3	0.09	0.13	0.05	0.18	-0.01	0.09
lag4	0.07	0.11	0.02	0.16	0.00	0.09
lag5	0.05	0.08	0.01	0.14	0.00	0.08
Panel C: Order imbalances measured in dollars standardized by total dollar volume.						
lag1	0.21	0.29	0.04	0.27	-0.14	0.14
lag2	0.12	0.18	0.04	0.20	-0.03	0.11
lag3	0.09	0.13	0.05	0.18	-0.01	0.09
lag4	0.07	0.11	0.02	0.16	0.00	0.09
lag5	0.05	0.08	0.01	0.14	0.00	0.08

Table 4. Determinants of Order Imbalances.

For each security, we regress dollar order imbalance scaled by total dollar volume, OIB (t), on a stock's lagged residual returns (the residual from regressing stock returns on contemporaneous market returns), lagged market return, Rm (t-k), and trader-type specific lagged order imbalances, OIB (t-k). Security-specific returns are computed based on closing quote midpoint, and market returns are computed as the equally-weighted average of these returns across all sample stocks. We report cross-sectional averages of the time-series regression coefficients. The sample includes 1,300 NYSE common stocks from January 2000 to April 2004. Panel A uses cumulative returns over the previous week, while Panel B uses daily returns over the previous week. Panel C uses the same model as Panel A, but provides separate average coefficients for each size quartile (using the time-series mean market value of equity for each firm). Numbers in boldface indicate that the mean coefficient is significant at the 5% level, using a t-test with an adjustment for cross-sectional correlations.

	Institutions		Regular program trades		Index arbitrage program trades		Individuals		Specialists		Other market makers	
	Mean	t	Mean	t	Mean	t	Mean	t	Mean	t	Mean	t
Panel A: Order imbalances and previous-week cumulative returns												
Intercept	**0.0048**	8.38	**0.0050**	14.57	**0.0007**	9.82	**-0.0076**	-16.23	**0.0005**	2.28	**-0.0008**	-9.92
Residual Ret (t-5, t-1)	**-0.1584**	-14.25	**-0.0947**	-15.01	-0.0012	-0.69	**-0.0822**	-23.91	**0.1589**	22.39	**-0.0064**	-4.22
Rm (t-5,t-1)	-0.0010	-0.07	**-0.1370**	-14.33	**-0.0281**	-9.49	-0.0080	-1.12	**0.1000**	11.84	0.0018	1.26
OIB (t-1)	**0.1790**	97.92	**0.2478**	106.77	**0.0349**	6.51	**0.1911**	82.84	**-0.1646**	-55.50	**0.1045**	32.20
OIB (t-2)	**0.0589**	38.16	**0.0768**	45.44	**0.0261**	8.31	**0.0898**	50.40	**-0.0719**	-31.23	**0.0466**	20.16
OIB (t-3)	**0.0385**	27.39	**0.0443**	27.60	**0.0455**	15.80	**0.0697**	41.04	**-0.0258**	-13.55	**0.0415**	21.44
OIB (t-4)	**0.0236**	16.45	**0.0332**	20.39	**0.0188**	6.73	**0.0562**	33.89	**-0.0057**	-3.22	**0.0355**	18.87
OIB (t-5)	**0.0221**	14.93	**0.0234**	15.01	**0.0067**	2.43	**0.0587**	38.30	**0.0119**	6.88	**0.0347**	20.44
Panel B: Order imbalances and previous-week daily returns												
Intercept	**0.0047**	8.35	**0.0049**	14.66	**0.0007**	9.93	**-0.0075**	-16.48	**0.0005**	2.26	**-0.0008**	-9.92
Residual Ret (t-1)	**-0.6089**	-28.66	**-0.6032**	-36.74	-0.0012	-0.37	**-0.1179**	-16.51	**0.3406**	24.71	**-0.0060**	-4.20
Residual Ret (t-2)	**-0.1969**	-10.14	**-0.0750**	-6.10	0.0052	1.58	**-0.0896**	-12.10	**0.2014**	19.26	-0.0044	-1.21
Residual Ret (t-3)	0.0075	0.43	**0.0420**	3.92	-0.0022	-0.74	**-0.0747**	-9.32	**0.1038**	9.74	**-0.0071**	-2.88
Residual Ret (t-4)	0.0126	0.84	**0.0870**	7.02	-0.0018	-0.50	**-0.0628**	-8.95	**0.0755**	8.82	**-0.0085**	-6.16
Residual Ret (t-5)	**0.0251**	1.95	**0.1130**	10.50	-0.0057	-1.46	**-0.0550**	-8.21	**0.0530**	6.30	**-0.0056**	-2.60
Rm (t-1)	0.0569	1.65	**-0.5166**	-22.41	**-0.1360**	-11.93	0.0225	1.36	**0.2156**	10.35	0.0033	0.98
Rm (t-2)	**-0.0793**	-2.58	**-0.1585**	-7.40	-0.0067	-1.02	-0.0188	-1.21	**0.1550**	8.83	**0.0117**	3.27
Rm (t-3)	**-0.0786**	-2.47	**0.0477**	2.35	**0.0624**	10.45	**-0.0437**	-2.70	**0.0798**	4.28	**-0.0086**	-2.37
Rm (t-4)	0.0048	0.15	0.0215	1.06	**0.0308**	4.95	0.0092	0.60	0.0001	0.01	-0.0009	-0.27
Rm (t-5)	**0.0834**	2.94	**-0.0964**	-4.81	**-0.0963**	-15.27	-0.0008	-0.05	**0.0469**	2.78	0.0032	0.87

	Institutions Mean	t	Regular program trades Mean	t	Index arbitrage program trades Mean	t	Individuals Mean	t	Specialists Mean	t	Other market makers Mean	t
OIB (t-1)	**0.1820**	103.25	**0.2469**	114.15	**0.0547**	10.86	**0.1921**	84.28	**-0.1440**	-47.91	**0.1064**	32.72
OIB (t-2)	**0.0648**	41.55	**0.0836**	50.41	**0.0308**	9.83	**0.0932**	52.20	**-0.0559**	-24.15	**0.0490**	20.97
OIB (t-3)	**0.0415**	28.63	**0.0476**	29.04	**0.0419**	14.97	**0.0714**	41.53	**-0.0206**	-10.40	**0.0420**	21.68
OIB (t-4)	**0.0246**	16.70	**0.0372**	22.53	**0.0081**	3.02	**0.0573**	34.29	**-0.0079**	-4.25	**0.0359**	18.68
OIB (t-5)	**0.0209**	13.97	**0.0245**	16.08	**0.0113**	3.91	**0.0595**	38.38	**0.0035**	2.02	**0.0351**	19.91

Panel C: Order imbalances and previous-week cumulative returns by size quartiles (using Panel A regression, only average return coefficients shown)

	Institutions		Regular program trades		Index arbitrage program trades		Individuals		Specialists		Other market makers	
	Mean	t	Mean	t	Mean	t	Mean	t	Mean	t	Mean	t
Size quartile 1 (smallest)												
Residual Ret (t-5, t-1)	**-0.1691**	-8.66	**-0.1070**	-8.65	-0.0051	-1.46	**-0.1407**	-15.43	**0.2484**	12.09	**-0.0082**	-5.10
Rm (t-5,t-1)	-0.0112	-0.26	**-0.1274**	-4.66	0.0070	0.90	-0.0231	-0.88	**0.0703**	2.40	-0.0037	-0.75
Size quartile 2												
Residual Ret (t-5, t-1)	**-0.2324**	-6.44	**-0.1366**	-8.82	-0.0086	-1.72	**-0.0923**	-12.40	**0.2228**	14.46	-0.0010	-0.17
Rm (t-5,t-1)	**-0.0539**	-2.12	**-0.2006**	-10.17	-0.0168	-3.68	-0.0166	-1.78	**0.1928**	14.01	0.0010	0.50
Size quartile 3												
Residual Ret (t-5, t-1)	**-0.1465**	-10.62	**-0.1114**	-8.22	-0.0016	-0.70	**-0.0593**	-11.46	**0.1126**	15.61	**-0.0075**	-7.81
Rm (t-5,t-1)	-0.0001	-0.01	**-0.1479**	-9.76	-0.0289	-9.97	0.0000	-0.01	**0.1014**	15.55	0.0029	1.83
Size quartile 4 (largest)												
Residual Ret (t-5, t-1)	**-0.0856**	-10.35	**-0.0239**	-4.01	0.0105	5.04	**-0.0365**	-16.73	**0.0517**	17.21	**-0.0090**	-12.55
Rm (t-5,t-1)	**0.0612**	5.07	**-0.0723**	-8.67	-0.0736	-15.14	**0.0077**	2.41	**0.0356**	10.65	**0.0069**	6.91

Table 5. The price Impact of Order Imbalances.

For each security, we regress daily close-to-close quote-midpoint returns, R (t), on contemporaneous market returns, Rm (t), and current and lagged order imbalances, OIB (t-k). Order imbalances are measured in dollars and scaled by total dollar volume. Market returns are computed as the equally-weighted average of close-to-close midpoint returns across all sample stocks. The reported coefficients are cross-sectional averages of the time-series regression coefficients. The sample includes 1,300 NYSE common stocks from January 2000 to April 2004. Panel B uses the same model as Panel A, but provides separate average coefficients for each size quartile (using the time-series mean market value of equity for each firm). Numbers in boldface indicate that the mean coefficient is significant at the 5% level, using a t-test with an adjustment for cross-sectional correlations.

	Institutions		Regular program trades		Index arbitrage program trades		Individuals		Specialists		Other market makers	
	Mean	t	Mean	t	Mean	t	Mean	t	Mean	t	Mean	t
Panel A: Price impact regression												
Intercept	**-0.0001**	-2.55	0.0001	1.60	0.0000	-0.64	**-0.0001**	-1.96	0.0000	-0.23	**-0.0001**	-2.71
Rm (t)	**0.9783**	82.01	**0.9927**	82.45	**0.9806**	81.09	**0.9730**	82.46	**0.9152**	80.54	**0.9656**	83.54
OIB (t)	**0.0028**	4.47	**-0.0189**	-19.65	0.0193	0.84	**-0.0608**	-20.46	**-0.1271**	-34.93	**-0.2666**	-3.02
OIB (t-1)	**0.0017**	6.19	-0.0002	-0.30	-0.0254	-1.00	**0.0212**	16.50	**-0.0264**	-21.05	0.4199	1.13
OIB (t-2)	-0.0003	-1.30	**0.0012**	2.52	0.0013	0.06	**0.0110**	10.97	**-0.0078**	-6.74	-0.1109	-1.12
OIB (t-3)	0.0001	0.34	0.0007	1.60	0.0121	0.55	**0.0078**	9.36	**-0.0038**	-3.74	-0.1999	-1.02
OIB (t-4)	0.0000	0.16	0.0006	1.18	0.0050	0.46	**0.0079**	8.40	-0.0002	-0.23	-0.1846	-1.37
Panel B: Price impact by size quartile (using Panel A regression, only average coefficients of contemporaneous OIB shown)												
Size quartile 1 (smallest)												
OIB (t)	**0.0061**	6.19	**0.0088**	5.43	**0.1597**	2.20	**-0.0101**	-7.00	**-0.0769**	-34.60	-0.4055	-1.15
Size quartile 2												
OIB (t)	-0.0011	-1.51	**-0.0118**	-10.98	**-0.0243**	-4.22	**-0.0227**	-12.79	**-0.0875**	-28.45	**-0.0618**	-3.61
Size quartile 3												
OIB (t)	-0.0009	-0.84	**-0.0287**	-17.00	-0.0234	-0.47	**-0.0508**	-14.31	**-0.1241**	-20.46	**-0.1659**	-10.03
Size quartile 4 (largest)												
OIB (t)	**0.0071**	3.79	**-0.0439**	-23.69	-0.0347	-1.43	**-0.1596**	-17.84	**-0.2199**	-20.02	**-0.4334**	-19.81

Table 6. Explaining the Price Impact of Institutional Order Imbalances.

The sample includes 1,300 NYSE common stocks from January 2000 to April 2004. We estimate cross-sectional regressions to explain the security-specific coefficient on institutional order imbalances (in dollar) in a regression of returns on contemporaneous market returns, contemporaneous institutional dollar imbalances, and lagged institutional dollar imbalances (see Table 5). In Panel A, the independent variables include relative effective spreads (RES), defined as twice the absolute difference between the execution price and the quote midpoint prevailing when the trade was reported; the temporary component of RES, measured as the price change from the trade price to the quote midpoint five minutes after the trade (multiplied by -1 for sell-signed trades); the permanent component of RES, defined as the change in quote midpoints from just before a trade to five minutes afterwards; and Size, the time-series mean of a firm's market value of equity. For the spread variables, we compute time-series means of the daily equally-weighted averages. In Panel B, R&D expenditures are scaled by sales and intangible assets are scaled by total assets. Both are measured at the year-end immediately preceding our sample period.

	Coefficient	t	Coefficient	t
Panel A: Intraday proxies for information asymmetry				
Intercept	-0.0031	-3.80	-0.0051	5.05
RES	1.2780	9.05		
Temporary component of RES			0.3120	0.96
Permanent component of RES			5.7530	5.74
Size * 10^{12} ($)	0.1997	7.70	0.2140	8.17
adjusted R^2	0.084		0.091	
Panel B: Financial statement proxies for information asymmetry				
Intercept	0.0030	3.32	0.0008	0.07
R&D/Sales in Dec 1999	0.1151	5.14	0.0898	3.88
R&D nonmissing dummy	-0.0026	-1.72	-0.0030	-2.16
Intangible/TA in Dec 1999			-0.0047	-1.33
Intangible nonmissing dummy			0.0028	0.26
adjusted R^2	0.021		0.013	

Table 7. The Predictive Power of Order Imbalances for Returns.
For each security, we regress daily close-to-close quote-midpoint returns, R (t), on contemporaneous market returns, Rm (t), and five lagged daily order imbalances, OIB (t-k). Market returns are computed as the equally-weighted average of close-to-close midpoint returns across all sample stocks. Order imbalances are measured in dollars and scaled by total dollar volume. The reported coefficients are cross-sectional averages of the time-series regression coefficients. The sample includes 1,300 NYSE common stocks from January 2000 to April 2004. Panel B uses the same model as Panel A, but provides separate average coefficients for each size quartile (using the time-series mean market value of equity for each firm). Numbers in boldface indicate that the mean coefficient is significant at the 5% level, using a t-test with an adjustment for cross-sectional correlations.

	Institutions		Regular program trades		Index arbitrage program trades		Individuals		Specialists		Other market makers	
	Mean	t	Mean	t	Mean	t	Mean	t	Mean	t	Mean	t
Panel A: Predictive regressions												
Intercept	-0.0001	-1.54	0.0000	0.20	0.0000	-0.66	0.0001	1.76	-0.0001	-1.07	0.0000	-0.21
Rm (t)	**0.9792**	81.68	**0.9807**	82.67	**0.9785**	82.32	**0.9791**	81.81	**0.9789**	82.45	**0.9785**	82.27
OIB (t-1)	**0.0018**	7.07	**-0.0059**	-10.90	-0.0265	-1.04	**0.0075**	8.00	**-0.0059**	-5.53	0.3859	1.05
OIB (t-2)	-0.0002	-0.67	-0.0002	-0.38	-0.0014	-0.06	**0.0044**	4.74	-0.0003	-0.29	-0.1318	-1.32
OIB (t-3)	0.0002	0.96	0.0000	0.03	0.0114	0.54	**0.0028**	3.41	-0.0016	-1.51	-0.1987	-1.07
OIB (t-4)	0.0001	0.54	0.0002	0.30	-0.0012	-0.10	**0.0033**	3.53	0.0005	0.42	-0.1985	-1.43
OIB (t-5)	0.0000	0.03	-0.0008	-1.67	-0.0413	-1.24	0.0008	1.06	-0.0005	-0.47	-0.3565	-1.00
Panel B: Predictive regressions by size quartile (using Panel A regression, only average coefficients of OIB (t-1) shown)												
Size quartile 1 (smallest)												
OIB (t-1)	**0.0010**	2.06	**-0.0080**	-4.83	**-0.1559**	-2.06	**0.0029**	3.23	**-0.0038**	-3.60	1.4067	0.96
Size quartile 2												
OIB (t-1)	0.0007	1.85	**-0.0054**	-9.82	-0.0075	-0.82	**0.0043**	2.61	**-0.0023**	-1.93	0.0955	1.16
Size quartile 3												
OIB (t-1)	**0.0013**	2.49	**-0.0048**	-4.64	0.0453	0.68	**0.0057**	3.02	**-0.0045**	-2.53	0.0122	1.10
Size quartile 4 (largest)												
OIB (t-1)	**0.0042**	7.12	**-0.0052**	-7.27	0.0121	1.07	**0.0173**	6.64	**-0.0131**	-3.71	**0.0291**	3.85

Table 8. The Predictive Power of Order Imbalances for Overnight Returns and Trading-day Returns.
This table presents predictive regressions of returns on past order imbalances. We divide returns into an overnight portion (measured from the closing midpoint on day t to the opening midpoint on day t+1) and a trading-day portion (measured from the opening midpoint on day t+1 to the closing midpoint on day t+1). For each security, we regress overnight returns (Panel A) or trading-day returns (Panel B), on contemporaneous market returns, Rm (t), and five lagged daily order imbalances, OIB (t-k). Market returns in Panel A (B) are computed as the equally-weighted average of overnight returns (trading-day returns) across all sample stocks. Order imbalances are measured in dollars and scaled by total dollar volume. The reported coefficients are cross-sectional averages of the time-series regression coefficients. The sample includes 1,300 NYSE common stocks from January 2000 to April 2004. Numbers in boldface indicate that the mean coefficient is significant at the 5% level, using a t-test with an adjustment for cross-sectional correlations.

	Institutions		Regular program trades		Index arbitrage program trades		Individuals		Specialists		Other market makers	
	Mean	t	Mean	t	Mean	t	Mean	t	Mean	t	Mean	t
Panel A: The predictive power of order imbalances for overnight returns												
Intercept	**0.0003**	5.86	**0.0003**	5.92	**0.0003**	5.74	**0.0005**	8.19	**0.0003**	5.65	**0.0004**	6.85
Rm (t)	**0.9853**	44.19	**0.9879**	44.33	**0.9899**	44.52	**0.9849**	44.16	**0.9888**	44.47	**0.9871**	44.46
OIB (t-1)	**-0.0005**	-2.57	**-0.0016**	-5.67	-0.0024	-0.29	**0.0076**	10.79	**0.0036**	5.24	0.3816	1.04
OIB (t-2)	-0.0002	-1.15	**-0.0008**	-2.66	0.0070	0.74	**0.0042**	6.99	**0.0039**	5.44	0.0087	0.35
OIB (t-3)	-0.0002	-1.48	0.0004	1.72	0.0137	1.51	**0.0036**	5.28	0.0005	0.69	-0.0048	-0.69
OIB (t-4)	-0.0002	-1.15	0.0002	0.72	**-0.0231**	-2.22	**0.0031**	4.90	**0.0018**	2.48	-0.1141	-1.24
OIB (t-5)	0.0000	0.03	0.0000	0.15	-0.0157	-1.38	**0.0030**	5.08	0.0009	1.32	-0.1332	-0.99
Panel B: The predictive power of order imbalances for trading-day returns												
Intercept	**-0.0004**	-4.24	**-0.0003**	-3.36	**-0.0003**	-3.67	**-0.0004**	-3.41	**-0.0004**	-3.76	**-0.0004**	-3.98
Rm (t)	**0.9653**	16.46	**0.9667**	16.39	**0.9638**	16.39	**0.9653**	16.36	**0.9648**	16.40	**0.9640**	16.29
OIB (t-1)	**0.0027**	4.43	**-0.0044**	-6.67	-0.0141	-0.61	-0.0043	-1.39	**-0.0118**	-7.28	0.0133	0.61
OIB (t-2)	-0.0008	-0.98	0.0013	1.14	-0.0050	-0.25	-0.0006	-0.37	**-0.0046**	-3.76	-0.1439	-1.36
OIB (t-3)	0.0009	1.91	-0.0015	-1.48	0.0028	0.14	0.0016	0.92	-0.0002	-0.15	-0.1876	-1.11
OIB (t-4)	0.0002	0.59	0.0004	0.54	0.0250	1.71	-0.0007	-0.62	-0.0024	-1.41	-0.0920	-1.09
OIB (t-5)	0.0007	0.95	-0.0007	-1.29	-0.0275	-1.10	-0.0011	-0.70	-0.0005	-0.43	-0.2170	-0.92