The University of Georgia

Marc L. Lipson
Associate Professor

Terry College of Business
Department of Banking and Finance

451 Brooks Hall
Athens, GA 30602-6253
Ph: (706) 542-3644
Fax: (706) 542-9434
mlipson@terry.uga.edu

January 4, 2005

U.S. Securities and Exchange Commission
450 5th Street, N. W.
Washington, DC 20549-0609

Attention: Mr. Jonathan G. Katz, Secretary

Re: File No. S7-10-04, Regulation NMS, release No. 34-49325 (February 26, 2004; the "NMS Release") and File No. SR-NYSE-2004-5, Release No. 34-50173 (August 10, 2004)

Ladies and Gentlemen:

I am writing to draw your attention to two research studies I have conducted which I believe are relevant as you consider the above mentioned changes in market regulation. In particular, I believe the work provides a very different view on trading costs for NYSE-listed stocks than a study provided to you, and conducted by, Fidelity Investments entitled "Comparison of effective spreads for the NYSE trades versus Electronic market trades in the NYSE listed stocks, as published in reports filed pursuant to Exchange Act Rule 11Ac1-5". The papers accompany this summary letter.

In the first of the two studies, "Competition Among Market Centers", I use "Dash-5" data which are filed with the Commission under Rule 11Ac-5. This is the same data used for the Fidelity analysis. In one respect, the two studies note the same fact: that spreads for some trades executed through the NYSE are larger than those at some competing market centers. The principal difference is that my study draws attention to results which provide a compelling and sensible explanation for the difference – an explanation which suggests that Fidelities conclusion that "the NYSE is a substantially more costly trading environment that that of the fully automatic trading environment of the Electronic Market" cannot be justified by this data (cover letter submitted by Eric D. Roiter, date December 8, 2004).

The basic explanation is that the type of order flow routed to the NYSE differs from that routed to other market centers. The NYSE order flow is much more informed. This means, for example, due to characteristics of the order such as the timing of the order, that a buy order at the NYSE will move prices upwards more than a buy order at other exchanges. It should be clear that a liquidity provider (such as the specialist at the NYSE) will have to sell at a higher price on the NYSE since they will have to pay more to replace the shares in their inventory. This is what academics refer to as "adverse selection costs". I show in my paper that the difference in adverse selection costs more than offsets the difference in effective spreads between the NYSE and most other exchanges.

It should be noted that the Dash-5 data include a measure called the realized spread. This realized spread has a simple economic intuition – it is what the liquidity provider earns (realizes) once one accounts for the replacement value of the shares bought or sold. It is equal to the effective spread less the change in prices resulting from the trade. The amount earned by the liquidity provider is a very good economic measure of the true cost of trading – what the liquidity provider actually earns is want the trader is really paying. This is, undoubtedly, the reason the measure is included in the data. By this measure, the NYSE is among the least costly venues for executing shares. In other words, once we account for the difference in order flow difficulty, the NYSE is no more costly than other exchanges and much less costly than many.

The second paper I am including, entitled "Are Retail Orders Different?" and co-authored with Charles Jones (Columbia University), specifically looks at the execution costs for retail order flow at the NYSE. When people make comparisons across market centers, it is often to argue that retail order flow would be better off at an exchange with lower average spreads. There is, however, a flaw in this argument – it presumes that the average execution cost for all order flow types is equal to the average for the exchange. We look at retail order flow and document that the average execution cost for retail order flow is much less than for other order flow at the NYSE. The paper also provides an explanation for this difference. Similar to the point of the paper I discussed above, the explanation lies in order difficulty – non-retail order flow tends to arrive at times when execution is more difficult.

I hope these studies provide you with additional insight into the complexities of market architecture and help you to view skeptically studies which assert the superiority of one market system over another.

Sincerely,

Marc L. Lipson

Competition Among Market Centers

*Marc L. Lipson**
University of Georgia

November, 2004

* Contact information: Terry College of Business, University of Georgia, Athens, GA, 30602; 706-542-3644; mlipson@terry.uga.edu.

This paper has benefited from the comments of Paul Bennett, Hendrik Bessembinder, Ekkehart Boehmer, Michael Goldstein, Mark Madoff, Gideon Saar, Ingrid Werner (who discussed the paper at the American Finance Association Meetings), and seminar participants at Babson College, Emory University, and the University of Georgia. The New York Stock Exchange provided support for this study. The comments, opinions, and errors are those of the author only. In particular, the views expressed here do not necessarily reflect those of the directors, members, or officers of the New York Stock Exchange, Inc.

Competition Among Market Centers

We examine competition among six market centers for NYSE-listed stocks using SEC Rule 11ac-5 data. We find that market centers competing with the NYSE execute orders in only a subset of stocks, order types and order sizes; that differences in effective spreads are generally much smaller than suggested by comparisons of overall averages, though reliable differences remain; that order flow routed to the NYSE is substantially more informed than order flow routed to broker-dealers and other exchanges (though less informed than marketable limit orders routed to ECNs); that prevailing quoted spreads at time of order arrival differ among market centers; and that results are less variable for orders routed to the NYSE. More importantly, we find that differences in effective spreads between the NYSE and some market centers are related to characteristics of the stocks traded.

Competition Among Market Centers

1. Introduction

Stocks listed on the New York Stock Exchange (NYSE) may be traded through a number of competing market centers, each with unique technologies and trading protocols. These include exchanges with trading floors, broker-dealers that provide execution services, and electronic communication networks (ECNs) that offer order driven markets. While the business model of each market center may rely upon innovations in market design, it may also rely upon selectively competing in a subset of order flow and/or a subset of stocks. Thus, differences in execution costs between the NYSE and these competing market centers may reflect advantages and disadvantages inherent in market center design, or may simply reflect variation in the segments market centers choose to serve. Furthermore, the ability of market centers to compete may vary depending on characteristics of the order flow and stocks that are selected.

We use data mandated by U.S. Securities Exchange Commission Rule 11Ac 1-5 (Dash5) to examine three questions related to differences in execution results between the NYSE and competing market centers.[1] First, are there differences in trading activity

---

[1] This rule requires market centers to make public a specified set of execution quality measures. Among other benefits, these data allow us to evaluate execution results for broker-dealers and electronic communication networks (ECNs), provide better measurements of execution results than can be constructed from trade and quote data (see Bessembinder, 2002b), allow a comparison of time to execution, and allow a breakdown and control based on order type. The purpose of the data was to promote optimal order routing for retail order flow. As a result, the rule covers only orders less than 10,000 shares and detailed data are only required for market orders and marketable limit orders. Fortunately, these orders account for approximately 36% of NYSE executed share volume and comprise the vast majority of activity at most market centers.

related to order type, order size, or the stocks that are traded at the market center.[2] Second, controlling for differences in order flow, are there persistent and reliable differences in execution results between the NYSE and competing exchanges? Third, if there are reliable differences, are the magnitudes of these differences related to characteristics of order flow or the stocks being traded? If there are differences in execution costs after controlling for trading activity, and if these are related to order characteristics or vary across stocks, then differences in execution costs are not simply an outcome of market design.[3] In this case, it is more likely that competition is imperfect and market centers vary in their ability to serve various market segments. For example, if market centers attempt to attract and retain less informed traders, this strategy may be more effective for small orders in lower volume stocks.

We find substantial variation in trading activity among market centers. As expected, the NYSE is the most active center.[4] More importantly, while the NYSE has order flow for just about every order category (a combination of stock, order size range, and order type), some market centers execute orders in only a subset of stocks, typically execute smaller orders, and/or predominantly execute a specific order type. Given the differences in trading activity, we use category-by-category comparisons to compare execution statistics for the NYSE with those of competing market centers. For marketable limit orders, NYSE effective spreads are reliably lower than those at all the

[2] Selectivity may result from a conscious strategy, such as the purchasing of less informed order flow, or an outgrowth of decisions regarding the design of execution procedures, such as the anonymity of ECNs.
[3] See Domowitz and Steil (1999), Conrad, Johnson and Wahal (2002), and Hasbrouck and Saar (2002), among others, for discussions of alternative trading mechanisms and Stoll (2001) for a discussion of the benefits of encouraging innovation.
[4] We compare the NYSE to the broker-dealers Knight and Madoff, the ECN Archipelago, and the Boston, Cincinnati and Chicago stock exchanges. These market centers are the six most active centers for NYSE-listed securities during our study period.

market centers we examine. For market orders, NYSE effective spreads are reliably lower than spreads at four of six market centers and reliably higher than those at the remaining two. For example, effective spreads on the NYSE for very small market orders are about a half penny lower than for Knight Securities and about a half penny higher than for Madoff Securities. We note that these results are remarkably different than results based on market center averages, in which NYSE effective spreads are three to seven pennies higher than competing market centers. This illustrates the extent to which selectivity characterizes competition among market centers.

We find that order flow routed to the NYSE is substantially more informed than order flow routed to broker-dealers and other exchanges. For ECN Archipelago, on the other hand, there is no difference in information for market orders, though marketable limit orders routed to Archipelago are typically more informed than those routed to the NYSE. For example, after the arrival of very small market or marketable limit orders at the NYSE, prices move about twice as much as they do following similar orders at broker-dealers or other exchanges. Furthermore, NYSE order flow is sufficiently more informed than order flow routed to competing market centers, that the NYSE has substantially lower realized spreads than all other market centers except Archipelago.[5] We also examine quoted spreads at the time of order submission. In general, prevailing quoted spreads are larger at the time small orders arrive at the NYSE than they are when small orders arrive at competing market centers. The reverse is true for larger orders.

---

[5] While the effective spread is the difference between execution prices and value of a security at the time of order placement, the realized spread is the difference between the execution price and a subsequent value of the security. The difference in the spread measures is due to price movements subsequent to order placement and reflects order difficulty due either to characteristics of the order or market conditions around the time of order execution. For this reason, realized spreads can be considered a measure of the gross

More importantly, those market centers with effective spreads smaller than the NYSE are those that execute orders when prevailing spreads are narrower. These results suggest that order difficulty and order timing are important dimensions along which competing market centers may chose to specialize.

The potential benefits of catering to more profitable market segments have been well documented.[6] In particular, Easley, Keifer and O'Hara (1996) note that market centers capturing less-informed trading volume (cream skimming) can earn substantial profits.[7] Given the potential benefit from selective competition in equity markets and given our earlier results suggesting that specialization occurs, we provide evidence on the following question: when is selective competition actually more effective? We find that differences in effective spreads between the NYSE and some competing market centers vary systematically in the cross-section of stocks. For example, the quintile of firms for which the difference in effective spreads between Madoff Securities and the NYSE is greatest in July 2001, remains the quintile with the greatest difference each month over the remaining year. We find that factors related leading to successful competition vary across market centers. However, competition generally appears to be more effective for lower volume stocks, stocks with higher volatility, stocks where the NYSE provides less

---

revenue to providers of liquidity and are indicative of the expected cost of execution when order characteristics are controlled for.

[6] Much of the literature extends Rothschild and Stiglitz (1976), who note that insurance companies can benefit from offering contracts that attract lower risk individuals. Similar arguments have subsequently been made for capital market institutions, including banks making loans, venture funds selecting investments, and brokerage firms purchasing order flow.

[7] Chordia and Subrahmanyam (1995) and Battalio, Greene and Jennings (1997), describe the arrangements and agreements that route order flow to various market centers. Related evidence and discussions can be found in Battalio, Greene and Jennings (1997), Bloomfield and O'Hara (1998), Battalio and Holden (2001). Consistent with the effects of order routing, Blume and Goldstein (1997) find that orders do not always flow to the market center posting the best prices and Bessembinder (2002a) links higher trading costs on the regional exchanges to orders being routed to regional exchanges when their quotes are not competitive.

price improvement, and for smaller orders.[8] These results are consistent with predictions based on models of strategic competition in which successful entrants focus on smaller markets.[9] Thus, while our earlier results suggest that market centers may compete by selectively executing orders based on order and stock characteristics, these results suggest that whatever strategies are chosen are not equally effective for all order flow and for all stocks.

Finally, we note that there are significant differences in the time to execution among market centers and that execution results vary less on the NYSE than on most other market centers. Speed and predictability are factors that might attract specific trading clienteles and provide a basis for market center differentiation.[10]

Taken together, our results suggest that competition for execution services in NYSE-listed stocks is often characterized by specialization and selective competition. In fact, selective execution may be necessary for effectively competing against the NYSE: the two market centers that have reliably lower effective spreads than the NYSE are the two that execute orders in the fewest stocks while the two market centers with reliably higher effective spreads are those that execute orders in almost all the stocks in our sample.[11] In addition, the low spreads on broker-dealer Madoff seem to reflect their

---

Easley, Keifer and O'Hara (1996) and Bessembinder and Kaufman (1997) provide empirical evidence of cream skimming while Battalio (1997) does not.

[8] The phrase "competing more effectively" is used here as a comment on the comparative ability of one market center over another, not the absolute ability. For example, effective spreads on the Chicago Stock Exchange are typically higher than effective spreads on the NYSE. However, for those stocks with lower volume, the difference between the market centers is lower. In this case we would say the Chicago Stock Exchange competes more effectively in lower volume stocks.

[9] Porter (1985) outlines a number of strategies which were subsequently refined by Wright (1987).

[10] Boehmer (2002) and Hasbrouck and Saar (2002) discuss issues related to time to execution.

[11] We do not draw any conclusions in this study about the desirability of selective execution. Less informed traders will certainly prefer to pool only with other uniformed traders (Admati and Pfleiderer (1988)). On the other hand, fragmentation of order flow may degrade price discovery (see Mendelson (1987), Madhavan (1995), Stoll (2001), and Huang (2002), among others).

ability to select less informed order flow. If we view competition among market centers as a search for market segments in which a market can compete more effectively, our results suggest that competition focuses on subsets of order flow and stocks and that competition varies predictably in effectiveness across stocks.

Our results are related to a number of studies comparing market centers. In contrast to Lee (1993) and Blume and Goldstein (1997), we find that the regional exchanges are not uniformly more costly than the NYSE. Our results also contrast those of Conrad, Johnson and Wahal (2002), who find that institutional orders routed to ECNs are less costly to execute than orders routed to the listing market. On the other hand, our results are consistent with Barclay, Hendershott and McCormick (2002) and Huang (2002) who find that ECN spreads are higher than those at the listing exchange and that ECNs appear to attract more informed order flow. We augment this line of inquiry by updating earlier results and adding results on broker-dealers, which are organized quite differently than exchanges or ECNs. The importance of examining broker-dealers is apparent in the case of Madoff, whose performance differs most remarkably from that of the NYSE.

Studies of cream skimming have focused on whether reliable differences in the informativeness of order flow exist and whether these are related to agreements for the purchase of order flow. We extend this literature by adding results on when these agreements augment the competitiveness of market centers. For example, Easley, Keifer and O'Hara (1996) point out that cream skimming can be especially profitable when market centers match the spreads on markets where trading volume is more informed. Consistent with this argument, we find that competition is more effective when there is less price improvement by the NYSE. Furthermore, given the limited ability of the NYSE trading floor to treat small orders differently than large orders, it is not surprising that competition is more effective for smaller orders.

The paper is organized as follows. Section 2 discusses the market centers we examine and describes the Dash5 statistics. Section 3 discusses our data. Section 4 examines issues related to the calculation and appropriate use of Dash5 statistics and uses the statistics to evaluate execution quality. Section 5 provides a brief conclusion.

2. Market Centers and Dash5 Data Details

The landscape for execution services has been evolving at a rapid pace. A decade ago the central question was whether the regional exchanges were sufficient competition for the NYSE. Now there is a proliferation of trading venues and these venues employ vastly differing technologies and execution procedures. Even these details are changing at every moment. Below we discuss the market centers examined in this study. Following that we discuss the Dash5 data that are required by the SEC and used in this study.

2. 1 Market Centers

The NYSE is a centralized continuous auction market. All orders are routed to a single specialist who is responsible for maintaining an orderly market in the traded stock. The NYSE has both a physical trading floor and an electronic order routing system

(SuperDot).[12] The orders routed through SuperDot are the only NYSE orders subject to Dash5 reporting. The NYSE is the listing exchange for the stocks we examine. It is also, as we shall see, the dominant venue for order routing. For this reason, and to simplify the analysis, we will evaluate results at other market centers relative to the NYSE.

We examine six market centers in addition to the NYSE. Below is a description of each over the time period studied, classified by type of market center:

Broker-Dealers	
Knight	Knight Trading Group - A broker-dealer that provides execution services.
Madoff	Madoff Investment Securities - A discount brokerage that provides automated execution services.
ECN	
Archipelago	Archipelago ECN - An electronic communications network that has become the equity trading group for the Pacific Stock Exchange.
Exchanges	
Boston	Boston Stock Exchange – A floor based continuous auction market, originally specializing in New England based companies
Cincinnati	Cincinnati Stock Exchange - An electronic auction market
Chicago	Chicago Stock Exchange – An electronic auction market.

Each of the above market centers is unique in their organization and design.[13] Knight, and Madoff are broker-dealers who provide execution services specifically for their clients. We note that one contribution of this study to examine broker-dealers and that these market centers have the greatest degree of flexibility in limiting the services they provide. Archipelago is an ECN and is essentially an order driven electronic market that also accepts all orders in the stocks for it makes markets. The NYSE, Boston,

[12] There can be many providers of liquidity for NYSE listed stocks other than the specialist. In fact, the floor of the exchange encourages competition for liquidity provision. When we refer to the specialist in their general role as a liquidity provider, such statement applies all providers of liquidity.
[13] Since we do not attempt to link any specific features to differences in performance among market centers, we do not attempt to fully characterize the protocols at each market center.

Chicago and Cincinnati are stock exchanges and accept all orders in the stocks for which they make markets. Of these four, the NYSE and Boston retain traditional trading floors.

2. 2 The Details of SEC Rule 11Ac 1-5

This section discusses the critical features of the Dash5 data. From the many measures of execution quality that are available, Dash5 rules require the following:

- **Effective Spread.** The effective spread is equal to twice the difference between the price at which an order is executed and the midpoint of a benchmark quote, multiplied by -1 for sell orders.[14] The benchmark mid-quote should represent the price that would be obtained in the absence of transaction costs. Since the Dash5 statistics are based on orders (not transactions), the benchmark quote for Dash5 data is the prevailing quote at the time the order arrives at the market center.[15]
- **Realized Spread.** The realized spread is twice the difference between the execution price and the mid-quote five minutes after execution, multiplied by –1 for sell orders. The mid-quote represents the subsequent value of the security and the realized spread, therefore, reflects the gross trading profit to a liquidity provider from taking the other side of an order. The difference between effective and realized spreads reflects is the permanent price impact of the order. This impact reflects the difficulty of the order – essentially the expected information content of the order.[16]
- **Time to Execution.** The time from order arrival until it is executed.[17]
- **Price Improvement.** Price improvement measures the execution price relative to the quoted bid price for sell orders or the quoted ask for buy orders. In other words, it describes the execution relative to the execution price that would have occurred against the quoted spread at the time of order arrival.[18]

---

[14] The combination of the best bid and the best offer across all quotes is referred to as the National Best Bid and Offer (NBBO). This is the quote that is used to benchmark Dash5 data. Unless otherwise indicated, any reference to quoted spreads means the NBBO.

[15] The order arrival benchmark is more appropriate than an execution time benchmark because the effective spread will reflect any price movements that occur while an order is being executed. These price movements reflect real costs to traders and may differ across market centers. As mentioned previously, one advantage of Dash5 data is the accuracy of the results.

[16] Glosten and Milgrom (1985) and Easley and O'Hara (1987) describe how the information implicit in order flow is reflected in the spread. For a comprehensive discussion of these issues, the reader is referred to O'Hara (1997) and Harris (2002).

[17] The arrival time is technically the time that an order is captured by the market center's automated handling system. Also, as with many statistics, the execution time is the share-weighted time of execution if the order is executed in parts.

[18] This assumes, implicitly, that there is enough depth at quote to execute the order in full. The depth available at quote may vary across market centers. Bacidore, Battalio and Jennings (2001) examine issues related to depth and the ability of markets to provide more depth than what is displayed in quotes.

From the price improvement and effective spread data, one can infer the prevailing quote at the time of order arrival. We examine quoted spreads since they reflect market and order flow conditions at the time of order arrival.[19] Note that effective spreads not only reflect prevailing market conditions, but also reflect characteristics of the order, such as order size, that cannot be reflected in the quoted spread (see Blume and Goldstein (1992)).

Both effective and quoted spreads will change depending on market conditions at the time of an order. It should be stressed that spreads are not a perfect measure of trading costs for many reasons. For example, many trading decisions are worked over time and spreads cannot capture the change in prices brought about by earlier executions as part of the same decision. Furthermore, spreads ignore commissions and any other market center fees or costs. However, spreads are simple to measure, readily available, and are usually reasonable indicators of actual trading costs for very small orders.

Any comparison of spreads needs to condition, to whatever extent possible, on characteristics of the order and market conditions at the time of order placement. While the Dash5 data do not explicitly condition on prevailing market conditions, they do partition orders along two dimensions.

- **Order Size.** Orders are classified into four order size groups. These are indicated below along with the designation we use to describe the order size category. Note that Dash5 does not require statistics for order sizes below 100 shares (odd lots) or for 10,000 shares or more.

---

[19] The reason market participants use price improvement as a measure of market quality is precisely because the quoted spread may control for some characteristics of market conditions. For example, two market centers may provide the same price improvement, but one market center executes orders under more difficult conditions. The *realized* effective spreads would differ across the market centers, but price improvement would be the same. In this case, price improvement is a better reflection of the *expected* result of routing an order to a given market center.

Designation	*Order Size*
Very Small	100-499 shares
Small	500-1,999 shares
Medium	2,000-4,999 shares
Large	5,000-9,999 shares

- **Order Type.** Order conditions can dramatically affect how an order is executed and also reflect differing degrees of urgency on the part of customers. The more patient a customer, the lower the expected cost of execution (and the longer it will likely take to execute). The following categories require Dash5 statistics. The definitions below apply to buy orders; sell orders are defined analogously. The applicable quote is the quote prevailing at the time of order arrival.

Buy Order Type	*Description*
Market	No limiting price
Marketable Limit	Limit price equals or exceeds the ask
Inside the Quote Limit	Limit price is between the bid and ask
At the Quote Limit	Limit is equal to the bid
Near the Quote Limit	Limit is within 10 pennies below the bid

We refer to a combination of stock, order size, and order type as a category. Dash5 rules require statistics for any category for which an order was placed. We focus our analysis on market and marketable limit orders since the data required more other order types is limited (it does not included effective spreads, for example). Dash5 statistics are share-weighted within each category.[20] Dash5 statistics for each category also report the number of orders placed (the number of orders received by the market center), the number of shares placed (the number of total shares in the orders received by the market center), and the number of shares executed.

[20] Note that categories are determined by order size, while the share weighting is based on executed shares. For example, an order for 3,000 shares might execute in two equal parts, each part with a different execution result. The results would be weighted by the shares executed and reported in the medium size order category even though it executed in two parts each of which would be classified as a small order. The purpose is to represent the expected execution results for the *order* submitted. It is also the order characteristics, not execution sizes, which reflect the difficulty of executing an order.

Dash5 guidelines contain many provisions designed to prevent the statistics from being distorted by unusual orders. For example, orders that require special handling or have unusual restrictions are excluded. Stopped orders are excluded. Also, the portion of an order executed on a day different from when the order was placed is excluded. Orders that meet all the requirements for inclusion in the statistics are referred to as "eligible orders".

In general, requiring results by category attempts to create groups of orders that are essentially identical. These orders can then be compared to evaluate execution quality. However, these categories cannot capture all the characteristics that might affect order execution. For example, institutional orders are more likely to be motivated by price relevant information than orders originating from individuals. A market center whose order flow predictably originates from a given type of trader will have Dash5 statistics that disproportionately reflect the results that type of trader might expect. Furthermore, the categories do not acknowledge any differences in prevailing market conditions at the time an order is executed. For example, it is possible that orders are routed to a given market center when conditions for execution are most favorable and this market center's reported Dash5 statistics will be disproportionately determined by easy executions.[21]

3. Sample and Summary Statistics

This study examines 350 NYSE listed stocks over a one-year period spanning July 2001 through June 2002. We constructed the sample as follows. We began with all stocks available in the NYSE Trade and Quote (TAQ) data set for June 2001 that were matched to the Center for Research in Security Prices (CRSP) data set. We used the CRSP data to exclude from our study all securities that were not regular U.S. common equities (we excluded ADRs, funds, REITS, and unit rights offerings). Finally, we restricted our sample to those stocks with a June 2001 trade-weighted average price of at least $5.00.

From this initial sample, the 100 stocks with the largest share volume in June 2001 were selected (the most active stocks). The remaining firms were sorted into quintiles based, again, on June 2001 volume and 50 stocks were selected at random from each of the quintiles, providing an additional 250 stocks (less active stocks). In general, the results differed little between these two samples, so we present results for the combined sample of 350 firms except for sample summary statistics, which are presented in Table 1.[22]

The first section of Table 1 describes the characteristics of the companies whose stocks are studied. Not surprisingly, the most active stocks are substantially larger firms and firms with a slightly higher share prices than the less active firms. The second section of Table 1 describes the trading activity in the stocks, including the activity for orders subject to Dash5 statistics that are directed to the market centers we study. By construction, the most active stocks have substantially higher consolidated trading volume than the less active stocks – more than ten times more volume, on average.

---

[21] Additional categories would, of course, yield better comparisons. However, the realized spread does provide implicit controls on market conditions and order difficulty.
[22] We should note that for the less active stocks there are more cases where results are statistically insignificant. In many cases, this is simply a result of having few observations. For example, there are very few observations for marketable limit orders in the less active sample.

As for Dash 5 activity, for the most active stocks, there are 2,511 market orders and 1,392 marketable limit orders placed, on average, each day. For the less active stocks, trading activity is much lower: there are 227 market orders and 226 marketable limit orders placed each day.[23] Shares placed, as expected, follow the pattern in orders placed.

Table 1 also presents summary statistics for shares executed. As expected, virtually all the market orders are executed. Interestingly, about 20% of marketable limit orders are unexecuted. This reflects the fact that the determination of status of a limit order as "marketable" is made at the time of order arrival. Between that time and the time the order is made available for execution (e.g. the time at which it is displayed to the specialist), prices may have moved away from the limit order.

Table 2 presents summary statistics of trading activity for the market centers in this study. As might be expected, the NYSE is the most active market center with shares executed equal to over four times the total of all the other market centers combined and almost twenty times the next most active market center. More important, the NYSE has observations in more categories and stocks than other market centers. For example, the NYSE has observations in about 32 thousand categories while the next highest value is about 26 thousand for Knight. The lowest is Cincinatti with observations in just under 10 thousand categories. Similarly, while the NYSE has at least one observation in every

[23] For non-marketable limit orders, there are 4,363 and 1,078 orders placed each day, on average, for the two samples, respectively. To evaluate the activity subject to Dash 5 reporting to total trading activity, note that the order and share values are one-sided (buys and sells counted separately) whereas aggregate trading volume is two-sided (reflects both a buy and a sell).

stock in the sample, some market centers have observations in far fewer stocks. For example, there are results for Madoff in only 184 of the 350 stocks in our sample.[24]

These results show that there are clear differences in the samples for each of the market centers and we explore this further below. At this point, we note that the stocks in which competing market centers tend to have observations are the lower volume stocks, and that competing market centers tend to have smaller average order sizes.[25] For example, the total consolidated volume in the 184 stocks for which Madoff has observations is about 10% lower than for the NYSE, while the average order size for Madoff order flow is about 20% lower than order flow at the NYSE.

4. Differences Between the NYSE and Competing Market Centers

4. 1 Method of Comparison

Since market centers do not execute orders in every category and/or every firm, one has to be careful using Dash5 data to evaluate differences between any two market centers. In this section we highlight the magnitude of potential errors from inappropriate comparisons, illustrate the approach taken for the remainder of our analyses, and provide summary statistics for the magnitude of effective spreads (measured in various ways) at each market center. We consider three ways to make comparisons:

- *Simple Average*: Average the statistic across all categories and stocks and compare means.

---

[24] The maximum number of categories is 84,000 (350 stocks × 12 months × 4 order sizes × 5 order types). Thus, there are many cases of categories without observations. Even the NYSE has observations in only about 78% of the categories. These null observations, as might be expected, are predominantly in the larger order sizes and less active stocks.
[25] As expected, but not reported, the competing market centers therefore have more observations in the smaller order size categories.

- *Paired Differences by Stock*: Create a volume weighted average for each stock, and then calculate the difference between pairs of market centers for a given stock when there is an observation for both market centers. The distribution of paired differences can used to compare market centers.
- *Paired Differences by Category*: Calculate the difference between pairs of market centers for every category where there is an observation for both market centers. The distribution of the paired differences can be used to compare market centers.

The first approach is most simple, but will be distorted the most by variations in executed volume among market centers. For example, a market center that trades only high volume stocks will have a much lower average spread than a market center that also trades low volume stocks – regardless of actual execution quality. The next approach accounts for volume differences by volume weighting statistics and accounts for stock selection by comparing only stocks with observations. The final approach most closely adheres to the concept of holding all things constant – it looks at category-level differences (which includes the stock). Note that the last two approaches do not provide a simple single statistic that characterizes a given market center, though they do summarize the difference across any two market centers.

Table 3 presents our analysis of effective spread differences for market orders using the approaches outlined above. For each method, there is a column labeled "NYSE vs Other" which shows the inferred difference between the NYSE and the given market center using each approach. Consider the simple average across all category observations. The effective spread on the NYSE averages 11.50 pennies across all categories/stocks. The effective spread for other market centers varies from a low of 3.81 pennies (Madoff) to a high of 8.40 pennies (Knight). The NYSE effective spread is higher in every case and almost three times higher in some cases. All these differences are significant.[26]

Next we calculate volume weighted average spreads by stock and look at the distribution of stock-pair differences. This approach yields conclusions that are remarkably different from the simple average comparison. In particular, the only reliable differences are for three of the market centers and the magnitude of the differences is substantially attenuated relative to the simple average. Note (again) the variation across market centers in the number of observations – while many market centers have observations for virtually all stocks, others do not.

The final approach makes the most complete use of the data. This approach calculates the difference across market centers for every category/stock. Here we find that differences across some market centers just about vanish. For example, the difference between the NYSE and Cincinnati is about 2 hundredths of a penny. In this study, NYSE effective spreads are higher than two market centers and lower than three. Once again, it is important to note the variation across market center comparisons in the number of observations (categories) - from a low of 5,111 (Cincinnati) to a high of 13,752 (Knight). Given that we are examining market orders only, the maximum number of categories is 16,800 (350 stocks × 4 order sizes × 12 months).

In general, these results identify statistically reliable differences among market centers, though the magnitude is substantially smaller than what is suggested by looking

[26] One could also generate a share-weighted average. The results are not reported. However, we find that effective spreads are much lower: the NYSE effective spread drops precipitously to 4.57 pennies and the effective spread for other market centers ranges from a low of 1.99 pennies (Madoff again) to a high of 6.00 pennies (Instinet). More importantly, the inferences are changed. The NYSE has a lower trading cost

at simple averages. This illustrates the importance of looking at category level differences when drawing inferences.[27] In our subsequent analysis we will maintain this approach when documenting and exploring differences between market centers.

4. 2 Comparisons of Market Centers

We compare each market center with the NYSE in a series of tables that look at various Dash5 measures of execution quality. We begin by looking at spreads and spread components (realized spread and information component). We then examine price improvement, prevailing quoted spreads, and time to execution.

Spreads and Spread Components

The Dash5 statistics require reporting of effective spread and realized spread only for market and marketable limit orders. As discussed above, the effective spread is the cost to the trader, the realized spread is the gross revenue to liquidity providers (not necessarily the market maker or specialist). The difference is created by price movements subsequent to order arrival and reflects order difficulty, possibly due to the information implicit in the order (for this reason it is referred to as the information component of the spread).

Table 4 presents the analysis of spread components for very small orders. This table illustrates the analysis performed for each category. To conserve space, we will thereafter present summary tables covering all categories. We chose to present the details

---

than two of the other market centers and, in general, there is less difference across market centers. This shows that, on average, there is more aggregate volume in the smaller order sizes.

[27] Bessembinder (2003) makes a similar point by showing that there is a self-selection problem in comparing market centers. Bessembinder controls for the self-selection implicit in the information content differences across market centers and points out that remaining differences are much smaller than suggested by averages.

for the smallest orders since these are most likely to be retail orders and much of the discussion and concern about order routing relates to the treatment of retail orders. Furthermore, the results for very small orders are typical of many of the results we find.

In table 4, for each market center and order type, we provide the median of the indicated measure and, more importantly, we provide the median paired difference between the NYSE and the given market center. We use medians rather than means for our comparisons since outliers affect both the mean and variation in our results.[28] For simplicity, we present the median for the NYSE but not the median NYSE value for each market center comparison.[29] We include the number of pairs (in this case equal to the number of stocks since we are examining a single order size and type) used in the comparison so the reader can be aware of the extent to which the sample might differ in each comparison. For each paired test we not only list the median paired difference, but we indicate the proportion of pairs for which the NYSE value is greater (the proportion of strictly positive differences).

Associated with each paired difference is a test of statistical significance. These are based on non-parametric Wilcoxon signed rank tests. The first two columns indicate the number of pairs and the median number of shares executed in that category on the indicated market center. Consistent with our earlier results, the NYSE is far more active than any other market center and is active in at least as many of the stocks as any other

---

[28] The results for means, when significant, differ from the conclusions of medians in only one case. More often, but not frequently, the results for means are statistically insignificant while median tests are significant.

[29] The NYSE median for each market center comparison could be calculated only for those pairs of observations (stock and order category) for which the NYSE and the market center both have valid observations. In most cases, however, the NYSE median for each comparison is very close to the NYSE overall median.

market center. Note that the NYSE has far more marketable limit orders than all the other market centers.[30]

The next set of columns describes the comparison of effective spreads. For market orders, the NYSE effective spread is reliably larger than Madoff and Boston (though quite small in magnitude for Boston), and reliably smaller than two of the other market centers. For marketable limit orders, the NYSE has reliably lower effective spreads than *every* market center. The most notable differences in terms of magnitude are the results for Achipelago (over a penny worse than the NYSE) and Madoff (over a half-penny better than the NYSE). In terms of percentages of category pairs, Madoff has a lower effective spread 91% of the time, whereas Knight and Archipelago are lower only 22% and 24% of the time.

The third set of columns describes the realized spread. Here the evidence is fairly uniform. The NYSE is found to have reliably lower realized spreads than every other market center for market orders and lower than all but one market center (for which there is no reliable difference) for marketable limit orders. The final set of columns describes the information component of the effective spreads. These results shed light on important differences in the nature of order flow across market centers. In the case of market orders, the NYSE has more difficult order flow than every other market center except, once again, Archipelago. Marketable limit orders are similar.

In general, the information component results are the mirror image of the realized spread results. In other words, the NYSE has effective spreads generally in line with

---

[30] The raises an interesting question about order types. The choice between a regular market order and a marketable limit order may depend on the type of market center to which one is routing an order. It should

other market centers, but the difficulty of the execution for NYSE order flow dramatically reduces liquidity provider gross trading revenues. In this context, it is worth noting that the magnitude of the realized spread differences and information component differences (both generally over a penny), are often much greater than for the effective spread differences.

Table 5 summarizes the paired differences for market and marketable limit orders for all order sizes. For simplicity, we omit the other information from Table 4. There are few distinctive patterns in these results that we have not observed in the very small orders. In general, however, we observe fewer significant differences in the large order size category. While not reported, it should be noted that the number of observations decreases dramatically as we move to larger order sizes, consistent with the evidence in Table 2 that competing market centers provide services predominantly to smaller orders. Taken together, the Table 5 results suggest that Madoff and, to a much lesser extent, Boston, have consistently lower effective spreads while Archipelago has higher effective spreads.

These results identify a number of important characteristics of competition among markets. First of all, consistent with differences in size and type of orders, other characteristics of order flow are not the same across market centers. The substantial variation in the information component makes this point clear. Second, competition has minimized realized spreads on the NYSE. One interpretation of the realized spread is the equilibrium cost of executing an order conditioning on all order characteristics, including

---

be clear from the evidence we will present, that combining market and marketable order results will improve the relative performance of the NYSE.

informativeness and difficulty. Under this interpretation, then the NYSE is clearly the lowest cost market center.

Price Improvement, Quoted Spreads, and Time to Execution

There are two other dimensions of order execution that are emphasized in the Dash5 statistics. These are price improvement and time to execution. We begin with an analysis of price improvement. Recall that price improvement measures the price relative to the quoted spread. If the quoted spread represents a public benchmark price (like a suggested manufacturer's price), then price improvement is the discount from that benchmark. An interesting characteristic of the data, mentioned earlier, is that the combination of effective spread and price improvement (specifically the effective spread plus twice the improvement) must be equal to the prevailing quoted spread at the time of order arrival.

Examining quoted spreads provides information on the market conditions associated with the time of order arrival. In particular, variations in quoted spreads will reflect variations in liquidity due to market conditions or recent order flow. Thus, we can compare the average quoted spread at order arrival time across market centers to obtain another view (in addition to the information component) of the difficulty of executing order flow that is directed to various market centers.

The analysis of price improvement and quoted spreads is shown in Tables 6 in a format similar to Table 5. Specifically, the table lists the median difference between the NYSE and the indicated market center. For market orders, the NYSE provides more price improvement than most market centers (no worse than any market center) for very small orders. As order size increases, competing market centers tend to provide better improvement to the point where, for large orders, all market centers provide better price improvement. As with most of the measures, for marketable limit orders, there is evidence the NYSE does better than every market center.

The quoted spread analysis provides interesting results, particularly in light of the price improvement results. For very smaller orders, prevailing spreads are typically narrower, on average, at the time of NYSE order arrivals than at the time of competing market center order arrivals. As with price improvement, this reverses as we mover to larger order sizes. For the large orders, spreads are substantially smaller on the NYSE at the time of NYSE arrivals than at the competing market centers at the time of competing market center arrivals. In fact, the magnitude of the improvement typically parallels the magnitude of the quote difference. Thus, though there is clearly selection in the timing of orders being routed to various market centers, price improvements offset most of the differences.

Most notably, the two market centers that typically provide lower effective spreads than the NYSE (Madoff and Boston) are the two market centers that capture order flow at times when prevailing quoted spreads are smaller. This suggests these market centers are executing orders when market conditions are more favorable. Put another way, the results suggest market centers are filling small niches in the provision of execution services and these niches may not be just customer niches, but market condition niches.

Our analysis of the time to execution is presented in Table 7. The time to execution results vary by market center comparison, as might be expected given the differences in market center design (see Boehmer (2003)). We find that only the broker-dealers Madoff and Knight provide execution times that are reliably shorter than the NYSE for market orders. This is surprising in the case of Archipelago, Chicago and Cincinati since these are all electronic order driven markets. For marketable limit orders, the variation in execution time differences is much greater. In this case, Archipelago and Madoff are both faster. In general, the NYSE is faster than all the other exchanges, Archipelago is slower on market orders but faster for marketable limit orders, and the broker-dealer Madoff is reliably fast in every case.

Dispersion in Execution Quality Within Market Centers

In addition to the level of differences, it is important to assess the predictability of the difference. In other words, one might save a half a penny, on average, at a given market center, but that saving might be highly variable with extreme costs and extreme savings. We address this question by looking at the dispersion of reported results within each category. We concentrate on effective spreads, but results for other measures of execution quality provide similar inferences.

Table 9 presents measures of dispersion of the effective spread observations for each market center in a format similar to Tables 5 and 6. This analysis excludes all categories with less than 10 orders so that outliers do not drive the results. The table presents the standard deviation, minimum, and maximum effective spreads for four order sizes and for market and marketable limit orders. These statistics are measures of the dispersion across stocks and months. Consider, for example, very small market orders at the NYSE. The standard deviation in effective spreads is about 0.7 pennies. In every case but market orders at Madoff, the variation (whether in standard deviation or range) is lower at the NYSE. In some cases the variation is much greater – notably at

Archipelago. This basic result holds true in for all the order sizes, except that Madoff variation becomes larger than the NYSE for larger order sizes. In general, the results suggest that order execution at the NYSE is more predictable than at other market centers with the exception of Madoff. The order driven ECN Archipelago provides executions with extremely high variability.

These results suggests that predictability of execution costs may be a factor that should be considered when choosing execution venues. It is not a statistic that is generated for each stock. However, the variation across stocks, when market centers trade similar stocks, gives some idea of the predictability of execution. The caveat is that the variation may be generated by variation across stocks, rather than within stocks. However, since the NYSE trades every stock traded on other market centers, and the range will reflect the range across all stocks, we have evidence of greater consistency in NYSE results.

4. 3 Execution Quality in the Cross-Section

In the prior sections, we examined differences in average Dash5 statistics between the NYSE and competing market centers. Our results suggest that in many cases market centers are competing with the NYSE by targeting segments of the market. This process may involve innovation, but it would appear that success is related more to selectivity than to any specific market design innovation. If this is the case, then one might expect the ability to target segments of the market to vary depending on the stock being traded. For example, if Madoff is able to cream skim less informed order flow by purchasing orders, they may be better able to screen orders in this fashion for less active stocks. We explore this possibility in this section by examining the cross-section of differences in effective spreads.

Persistence of Market Center Differences

We begin by establishing that differences in effective spreads between market centers for a given stock is persistent over time. Specifically, we establish that the cross-sectional variation in differences persists over time.[31] This justifies our later tests that look for association between stock characteristics and the differences in spreads. Since we will be looking at variation across stocks, we normalize all spread differences by the average price of the stock during the month the data were generated. Thus, we perform our analysis on basis point differences.

To assess whether some market centers are consistently more or less costly venues for some stocks than others, we proceed as follows.[32] We take the effective spreads calculated for the first month of our sample (July 2001) and divide the stocks into quintiles based on the *difference* in effective spreads for each market center and the NYSE.[33] The first quintile is where the NYSE provides the highest effective spread relative to a given market center, and the fifth is where the NYSE provides the lowest effective spread relative to the given market center. We then calculate the difference for each of the subsequent 11 months based on the quintiles formed from the last month. We test whether the difference in the two extreme quintiles is significant using the time series over the 11 months.

---

[31] Incidentally, we also validate that our earlier conclusions about spread differences are not an artifact of the sample by showing that the results hold monthly.

[32] Note that we are looking for patterns across stocks in this analysis. Another possibility is that at certain *times* one market center might provide better execution services than another. However, we do not have the information needed to consider this possibility.

Before we consider the statistical tests, Figure 1 presents the graph of all five quintiles for Madoff. For the stocks in quintile 5, the difference between NYSE effective spreads and Madoff were the most positive in July – and the difference for these stocks remains the most positive in every other month. For the stocks in quintile 1, the difference was least positive in July (actually negative). These stocks remained the least positive in every month. In fact, the ordering remains remarkably consistent over time. This illustrates the concept we are examining in this section – whether there are patterns across stocks in how much better or worse execution is for one market center relative to another. After documenting the differences, we explore factors that might give rise to these differences.

Table 9 presents the tests of persistence for effective spreads for the whole sample, by market center and order size. There were no reliable cross-sectional differences for Archipelago, possibly a function of the large variation in Archipelago results noted in Table 8, and it has been omitted from this table. This may also reflect the possibility that this ECN does not screen order flow and, therefore, no variation would be expected. More likely, ECNs compete by attracting orders under certain market conditions, rather than from certain stocks.

The differences in Table 9, like our earlier table, are all relative to the NYSE. Consider the results for the smallest orders in Knight. The NYSE effective spreads are, on average, lower in each quintile since the difference is negative. More importantly, the difference for the quintile of stocks with the most positive difference is reliably more positive than for the quintile with the least positive difference. Put another way, for one

---

[33] One could assign quintiles based on any month and results would be similar.

set of stocks, the differences between Knight and the NYSE are reliably larger in magnitude than for another set of stocks. A few results are notable. First, there are there a number of cases where patterns are persistent. For example, regardless of order size, there are persistent differences between NYSE and Madoff. For Knight, Boston, and Chicago, at least half the order sizes exhibit persistent differences.

These results suggest that there are patterns in the cross-section of execution quality differences. It would seem that not all market centers provide comparable execution quality for all stocks.

Regression Analysis

In this section, we examine execution differences to see if they are related to characteristics of the stocks being traded. Specifically, we explore the relation between differences in effective spreads and firm characteristics in the cross section of firms. This analysis focuses on differences across stocks, so we use averages over our study period to prevent distortions related to pooling cross-sections and time-series.[34]

The dependent variable is the difference in relative effective spread calculated from the Dash5 data. The independent variables we examine are the following for each firm:

Quoted Spread	The trade-weighted average quoted relative spread, expressed in percent.
Price Improvement	The difference between relative quoted spreads and relative effective spreads, expressed as a percentage.

[34] Results are similar if we use monthly observations.

Turnover	The volume in the stock divided by the number of shares outstanding.
Log(Market Value)	The average of the daily log of market value.
Volatility	The standard deviation of the daily trade-weighted prices.
PIN	The probability of informed trading calculated using the methodology of Easley, Kiefer, O'Hara and Paperman (1996).
Size Indicators	Indicator variables for small, medium, and large orders (we omit an indicator for the very small orders).

These variables capture a number of possible determinants of effective competition. The quoted spread captures the opportunity for market centers to improve on spreads. Price improvement reflects the degree to which quotes are indicative of executions results – greater price improvement suggests that competition may be less driven by quote behavior and more related to preferencing and other arrangements (see Easley, Keifer, and O'Hara (1996) and Battalio (1997), Chordia and Subrahmanyam (1995) and Battalio, Greene and Jennings (1997), and Bloomfield and O'Hara (1998), among others). This may provide greater opportunities for competition. Size and volume (turnover) reflect trading information and trading interest, respectively. Larger firms have more analysts and more investors who follow or are aware of a firm. Larger firms also have greater trading activity (see Arbel and Strebel (1982), Merton (1987) Admati and Pfleiderer (1988), and Irvine (2003), among others). All else equal, market centers may choose to compete in the most active securities or, instead, target their efforts to the less visible and less active securities. Volatility imposes a cost on market making by increasing the risks associated with holding inventory (Demsetz (1968), Benston and

Hagerman (1974), and Stoll (1978)). This can affect competition market centers differed on their ability to absorb inventory risk.

Letting *c* represent a market center other than the NYSE, *n* represent the NYSE, *s* represent a stock, and *o* represent an order size, we first calculate the difference between the Dash5 NYSE effective spread and the Dash5 effective spread for a given market center *c* (in basis points relative to the monthly average price), order size category *o*, and stock *s*, after having calculated shares weighted averages over the twelve months.[35] This can be written as follows:

$$Difference_{c,o,s} = Effective\ Spread_{n,o,s} - Effective\ Spread_{c,o,s}$$

To prevent or results from being driven by a few outliers, we include only those observations with at least 10 orders placed during a given month.

We then run the following regression for selected market centers:

$$\begin{aligned} Difference_{c,o,s} &= \alpha + \beta_1\, Quoted\ Spread_s + \beta_2\, Price\, Improvement_s + \beta_3\, Turnover_s + \\ &\beta_4\, Log(Market\, Value)_s + \beta_5\, Volatility_s + \beta_6 PIN_s + IND(Small)_o + IND(Medium)_o + IND(L\arg e)_o + \varepsilon \end{aligned}$$

The results are given in Table 10. We provide the number of observations for each regression as well as adjusted *R*-squares. Given that we are looking at averages across the year, the most observations we would expect to see would be 350 stocks × 4 order sizes = 1,400 observations. Adjusted R squares vary from 0.036 (Knight) to 0.345 (Madoff). Perhaps not surprisingly, Madoff is the market center that provides the greatest improvement upon NYSE quoted spreads and statistically significant in the most number of cases. Furthermore, it is interesting to note that Madoff and Boston are the two market centers that improve upon the NYSE and have the two highest adjusted R squares – suggesting it is easier to explain better performance than worse.

Looking at the coefficients of the regressions we note the following. Market centers compete more effectively for lower turnover stocks (true in every case), for the smallest orders (true in every case), and where volatility is greater (true for two market centers). Results also show that Madoff does better for smaller stocks and Boston does better when quotes are wide. Results for price improvement are mixed, for Knight, Boston and Chicago, market centers compete less effectively when price improvement is high whereas Madoff does better when price improvement is high. In the case of Madoff this is consistent with greater opportunities to purchase order flow, while the result for the other three centers is hard to explain. It may be that the two exchanges do better when quotes are set more competitively and, one might expect, that aggressive quoting draws order flow. Note that the price improvement conditions on the quote size, so it is not inconsistent that Boston does better when quotes are higher and also when quotes are set competitively to draw order flow.

Of course, one should be cautious when interpreting these results. For example, our results would not imply that each market center should execute the smallest orders they can attract. What we are observing is that, given the characteristics of a market center and the orders that market center has chosen to attract, the market centers provide more favorable executions for smaller orders. Also, we emphasize again that this does not imply that the executions are favorable in absolute terms relative to the NYSE, but

[35] We conducted similar analyses for the components of the spread. In general, the results for realized spreads were similar to those for effective spreads and there were few reliable relations for the information component of the spread.

that they are *more* favorable in certain circumstances. Put another way, effective spreads for Knight are typically larger than for the NYSE, but not so much larger for the smaller orders. The central point of this analysis is that the effectiveness of the strategies employed by various market centers (which are not limited to structural innovation) differs across stocks, suggesting that business plans of competing market centers often include selective competition.

5. Conclusions

We examine competition among market centers using order level execution statistics. Our evidence suggests that competition may be best described as market centers developing strategies that allow them to profitably provide execution services to select segments of the market. Our evidence includes the fact that order flows are not similar across market centers, that effective spread differences exist but are related to characteristics of order flow, and that the ability to compete with the NYSE depends on characteristics not related to market center design. Thus, the market for NYSE-listed execution services is one where competition is not perfect and effective spreads are not driven to identical levels.

Interestingly, the differences in effective spreads for market orders are much smaller in magnitude than the differences in realized spread. This result can be interpreted in two ways. First, since the realized spread conditions on the price changes induced by an order, it implicitly controls for order difficulty. Under this interpretation, differences in effective spreads are not an appropriate basis for comparing execution results. Another interpretation (the two are not mutually exclusive) is that competition may limit potential differences in effective spreads, and that this competition, combined with more difficult order flow being routed to the NYSE, attenuates the profitability of liquidity provision on the NYSE.

Because of differences in order flow characteristics, and the lack of consistency in results (they depend on the order type, size, and stocks examined), drawing broad conclusions about the success of various market architectures is difficult. In fact, we find no evidence that any particular market architecture is uniformly more successful at providing execution services. For example, one broker-dealer, Madoff, does quite well, whereas another, Knight, does poorly. More importantly, competition among market centers appears to lead to specialization and fragmentation, not a convergence to a single, dominant, lowest cost architecture.

References

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

Arbel, Avner, and Paul Strebel, 1982, The neglected and small firm effects, *The Financial Review*, 201-218

Badidore, J., R. Battalio, and R. Jennings, 2001, Depth improvement and adjusted price improvement on the New York Stock Exchange, *Journal of Financial Markets* forthcoming

Bae,K. H., H. Jang, K. S. Park, Traders' choice between limit and market orders: Evidence from NYSE stocks, *Journal of Financial Markets* 6, 517-538.

Barclay, M. J., T. Hedershott, and D. T. McCormick, 2002, Competition among trading venues: Information and trading on electronic communications networks, forthcoming *Journal of Finance*.

Battalio, Robert, 1997, Third market broker-dealers: cost competitions or cream skimmers?, *Journal of Finance* 52, 341-352.

Battalio, Robert, Jason Greene, and Robert Jennings, 1997, Do competing specialists and preferencing dealers affect market quality?, *Review of Financial Studies*.

Battalio, Robert, and Craig Holden, 2001, A simple model of payment for order flow, *Journal of Financial Markets*.

Benston, George J., and Robert Hagerman, 1974, Determinants of bid-asked spreads in the over-the-counter market, *Journal of Financial Economics* 1(4), 353-64

Bessembinder, Hendrik, 2002, Quote-based competition and trade execution costs in NYSE-listed stocks, *Journal of Financial Economics*, forthcoming.

Bessembinder, Hendrik, 2002, Issues in assessing trade execution costs, *Journal of Financial Markets*, forthcoming.

Bessembinder, Hendrik, 2003, Selection biases and cross-market trading cost comparisons, Working Paper.

Bessembinder, Hedrik, and Herbert M. Kaufman, 1997, A cross-exchange comparison of execution costs and information flow for NYSE-listed stocks, *Journal of Financial Economics* 46, 293-319.

Bloomfield, R. and M. O'Hara, 1998, Does order preferencing matter?, *Journal of Financial Economics* 50, 3-37.

Blume, Marshall, and Michael Goldstein, 1992, Displayed and effective spreads by market, Rodney L. White Center for Financial Research Working Paper

Blume, Marshall, and Michael Goldstein, 1997, Quotes, order flow, and price discovery, *Journal of Finance* 52, 221-244.

Boehmer, E., 2003, Dimensions of execution quality: recent evidence for U.S. equity markets, Working Paper

Chordia, Tarun, and Avinidhar Subramanyam, 1995, Market making, tick size, and payment-for-order flow: Theory and evidence, *Journal of Business* 68, 543-575.

Conrad, Jenniger, Kevin M. Johnson, and Sunil Wahal, 2002, Institutional trading and alternative trading systems, *Journal of Financial Economics* forthcoming.

Demsetz, Harold, 1968, The cost of transacting, *Quarterly Journal of Economics* 82(1), 33-53

Domowitx, I. And B. Steil, 1999, Automation, trading costs, and the structure of the securities trading industry, *Brookings-Wharton Papers on Financial Services* 2, 33-92.

Easley, David, N. Kiefer, and Maureen O'Hara, 1996, Cream skimming or profit sharing? The curious role of purchased order flow, *Journal of Finance* 51, 811-834.

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

Easley, D., N. Kiefer, M. O'Hara, and J. Paperman, 1996, Liquidity, information and infrequently traded stocks, *Journal of Finance* 51, 1405-1436

Foucault, T., 1999, Order flow composition and trading costs in a dynamic limit order market, *Journal of Financial Markets* 2, 193-226.

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

Harris, Lawrence, 2002, "Trading and Exchanges: Market Microstructure for Practitioners", Oxford University Press.

Harris, Lawrence and Joel Hasbrouck, 1996, Market vs. limit orders: The SuperDOT evidence of order submission strategy, *Journal of Financial and Quantitative Analysis* 31, 213-231.

Hasbrouck, Joel, 1995, One security, many markets: Determining the contributions to price discovery, *Journal of Finance* 50, 1175-1199.

Hasbrouck, Joel, and Gideon Saar, 2002, Limit orders and volatility in a hybrid market: the Island ECN, Working Paper.

Huang, Roger, 2002, The quality of ECN and Nasdaq market maker quotes, *Journal of Finance* 57, 1285-1319.

Huang, Roger, and Hans Stoll, 1996, Dealer versus auction markets: A paired comparison of execution costs on NASDAQ and the NYSE, *Journal of Financial Economics* 41, 313-358.

Irvine, Paul J. A., 2003, The incremental impact of analyst initiation coverage, *Journal of Corporate Finance 9*, 431-451.

Lee, Charles,1993, Market integration and price execution for NYSE-listed securities, *Journal of Finance* 48,1009-1038.

Madhavan, A., 1995, Consolidation, fragmentation, and the disclosure of trading information, *Review of Financial Studies v8(3),* 579-603.

Mendelson, H., 1987, Consolidation, fragmentation and market performance, *Journal of Financial and Quantitative Analysis* 22, 189-207.

Merton, 1987, A Simple Model of Capital Market Equilibrium with Incomplete Information, *Journal of Finance* 42(3), 483-510

O'Hara, Maureen, 1997, "Market Microstructure Theory", Blackwell Publishers, Cambridge, Mass.

Petterson, Mark, and Erik Sirri, 2002, order submission strategy and the curious case of marketable limit orders, *Journal of Financial and Quantitative Analysis* 37, 221-241.

Porter, M, 1985, *Competitive Advantage: Creating and Sustaining Superior Performance,* Free Press, New York.

Stoll, Hans R., 1978, The pricing of security dealer services: an empirical study of Nasdaq stocks, *Journal of Finance* 33(4), 1153-1172

Stoll, Hans R, 2001, Market fragmentation, *Financial Analysts Journal* 57, 16-20.

Wright, P., 1987, A refinement of Porter's strategies, *Strategic Management Journal* 8, 93-101.

Table 1: Sample Summary Statistics

This table contains summary statistics for the sample firms. The sample is composed of the 100 NYSE-listed stocks with the highest June 2002 executed share volume (most active stocks) and a random sample of 250 additional stocks (less active stocks). The less active sample is composed of five sets of 50 stocks chosen at random from June 2002 executed share volume quintiles. Stocks are all NYSE listed common equity (REITs, ADRs and funds excluded) with a share-weighted average June, 2001 price of at least \$5.00. The Dash-5 activity levels are average of the daily totals across the 9 market centers studied.

Firm Characteristics	**Most Active Stocks**	**Less Active Stocks**
Price (Trade-Weighted June 2001)	41.82	30.69
Shares Outstanding (thousands, End-of-June 2001)	1,438,407	100,581
Market Value (thousands, End-of-June 2001)	64,447,948	3,457,944
Average Daily Trading Activity		
Daily Consolidated Volume (June 2001)	5,313,613	419,435
NYSE Market Share (June 2001)	85.67%	87.41%
Daily Dash 5 Orders Placed		
Market Orders	2,511	227
Marketable Limits	1,392	226
Daily Dash 5 Shares Placed		
Market Orders	1,710,188	109,909
Marketable Limits	1,352,277	151,617
Daily Shares Executed		
Market Orders	1,694,070	108,340
Marketable Limits	1,068,635	119,568

Table 2: Activity by Market Center

This table presents the Dash5 trading activity for each market center. Shares executed and consolidated volume in executed stocks are the totals over the 12-month sample period from July, 2001 through June, 2002. The number of categories is the number of categories with non-zero orders placed for execution. Stocks traded is the number of stocks with at least one observation over the 12-month sample period.

	Shares Executed (millions)	Number of Categories With Observations	Number of Stocks With Observations	Volume in Stocks Traded (millions)	Average Trade Size
NYSE	65,434	32,608	350	8,914	2,360
Broker-Dealers					
Knight	3,402	26,870	350	8,914	1,873
Madoff	2,143	14,427	184	8,182	1,944
ECN					
Archipelago	820	18,201	348	8,689	1,705
Stock Exchanges					
Boston	2,990	17,527	284	8,778	2,003
Chicago	2,566	25,646	343	8,907	1,896
Cincinnati	1,439	9,996	191	7,672	2,056

Table 3: Overview of Differences in Effective Spreads for Market Orders

This table presents and overview of differences in effective spreads (in pennies) for market orders using alternative methods. Each method includes a column "NYSE vs Other", which describes the inferred difference between the NYSE and the given market center based on the approach chosen. Statistical tests in the first case are *t*-tests or the difference in means between the NYSE and competing market center. In the second two cases, we use a *t*-test to determine whether the distribution of differences is different from zero.

	Simple Average		**Paired Stocks Differences**			**Paired Category Differences**		
	Spread	**NYSE vs Other**	**Num**	**NYSE Spread**	**NYSE vs Other**	**Num**	**NYSE Spread**	**NYSE vs Other**
NYSE	11.50							
Knight	8.40	+3.10***	350	8.66	-0.04	13,752	8.18	-0.19***
Madoff	3.81	+7.70***	184	5.19	+1.99***	7,497	5.38	+1.58***
Archipelago	6.17	+5.33***	343	8.53	-0.22	7,094	5.32	-0.86***
Boston	5.85	+5.65***	283	6.72	+0.47	9,401	5.99	+0.14*
Chicago	8.14	+3.35***	343	8.41	-0.21**	12,778	7.52	-0.62***
Cincinnati	4.81	+6.70***	149	4.68	+0.25**	5,111	4.79	-0.02

*, **, and *** indicate significance at the 10%, 5% and 1% levels, respectively.

Table 4: Spread Component Comparisons for Very Small Orders

This table presents our analysis of the difference in execution quality based on spread components. Values are medians across pairs and tests of significance are Wilcoxon signed rank tests on the distribution of paired differences. Also presented are the percentage of NYSE stocks whose observations greater than the given market center.

	Activity		Effective Spread			Realized Spread			Information Component		
	Num.	1,000 Shares Per Stock	Avg.	Paired Difference	NYSE Greater	Avg.	Paired Difference	NYSE Greater	Avg.	Paired Difference	NYSE Greater
Market Orders											
NYSE	350	6,241	3.62			0.41			3.00		
Knight	350	276	4.16	-0.45***	22%	2.81	-2.17***	4%	1.15	+1.60***	94%
Madoff	185	1,658	2.11	+0.55***	91%	1.62	-1.31***	3%	0.41	+1.94***	98%
Archipelago	337	7	4.51	-1.07***	24%	1.47	-0.91***	31%	2.75	-0.18	47%
Boston	277	193	3.13	+0.01*	50%	2.29	-1.97***	5%	0.44	+2.02***	93%
Chicago	345	317	3.58	-0.11***	40%	2.33	-1.98***	4%	0.88	+1.81***	94%
Cincinnati	144	966	2.60	-0.01	49%	2.01	-1.87***	6%	0.39	+1.79***	93%
Marketable Limit Orders											
NYSE	350	5,694	2.37			-0.35			2.63		
Knight	350	29	3.87	-1.33***	2%	2.24	-2.42***	9%	1.21	1.06***	79%
Madoff	173	61	1.80	-0.17***	31%	1.36	-1.69***	9%	0.35	1.47***	90%
Archipelago	347	47	3.64	-1.23***	14%	-0.31	0.07	51%	4.06	-1.37***	26%
Boston	251	26	2.39	-0.54***	20%	1.93	-2.25***	15%	0.20	1.55***	83%
Chicago	335	49	4.01	-1.54***	6%	1.08	-1.32***	31%	2.70	-0.32***	44%
Cincinnati	185	33	1.82	-0.22***	31%	1.55	-2.03***	18%	0.00	1.82***	84%

*, **, and *** indicate significance at the 10%, 5% and 1% levels, respectively.

Table 5: Paired Differences for Spread Components

This table summarizes the differences in spread components for market orders between the NYSE and other market centers, for each size category, and for each sample. Values are medians across pairs and tests of significance are Wilcoxon signed rank tests on the distribution of paired differences.

	Very Small			Small			Medium			Large		
	Effective	Realized	Info	Effective	Realized	Info	Effective	Realized	Info	Effective	Realized	Info
Market Orders												
NYSE Median	3.62	0.41	3.00	5.89	0.63	4.97	9.99	1.84	7.47	16.07	3.51	11.22
Median Difference Between Market Center and NYSE												
Knight	-0.45***	-2.17***	+1.60***	+0.13*	-1.58***	+1.62***	-0.25	-1.31***	+1.25***	-0.07	-1.33***	+1.11**
Madoff	+0.55***	-1.31***	+1.94***	+1.32***	-0.73***	+1.97***	+1.88***	+0.38***	+1.41***	+1.80***	+0.59***	+1.28***
Archipelago	-1.07***	-0.91***	-0.18	-0.43*	-0.03	-0.39	-0.14	-0.06	-0.14	+0.66**	-0.28	+0.55
Boston	+0.01*	-1.97***	+2.02***	+0.21***	-1.92***	+2.09***	+0.16***	-1.13***	+1.47***	+0.92***	-0.70**	+1.38***
Chicago	-0.11***	-1.98***	+1.81***	-0.36***	-2.44***	+1.83***	-0.79***	-2.17***	+1.36***	-0.32	-1.38***	+1.13***
Cincinnati	-0.01	-1.87***	+1.79***	+0.04	-1.58***	+1.66***	-0.34***	-1.20***	+1.10***	-0.02	-1.60***	+1.63***
Marketable Limit Orders												
NYSE Median	2.37	-0.35	2.63	2.80	-0.60	3.68	3.63	0.12	3.33	4.71	0.23	4.45
Median Difference Between Market Center and NYSE												
Knight	-1.33***	-2.42***	+1.06***	-0.73***	-1.38***	+0.61***	-0.50***	-0.89***	+0.35***	-0.39***	-1.29***	+0.85***
Madoff	-0.17***	-1.69***	+1.47***	-0.20***	-0.92***	+0.83***	-0.22***	-0.41	+0.08	-0.32***	-0.86***	+0.36
Archipelago	-1.23***	+0.07	-1.37***	-1.06***	+0.17	-1.05***	-0.47***	+0.72***	-1.07***	+0.32***	+0.98***	-0.75***
Boston	-0.54***	-2.25***	+1.55***	-0.59***	-2.26*	+1.57***	-0.73***	-1.80***	+1.08***	-0.60***	-2.14***	+1.17***
Chicago	-1.54***	-1.32***	-0.32***	-1.47***	-1.04***	-0.48***	-1.55***	-0.53***	-0.92***	-0.95***	-0.14	-0.73***
Cincinnati	-0.22***	-2.03***	+1.82***	-0.35***	-2.05***	+1.54***	-0.61***	-1.46***	+1.27***	-0.66***	-0.99	+0.46

*, **, and *** indicate significance at the 10%, 5% and 1% levels, respectively.

Table 6: Improvement and Quote

This table summarizes the differences in price improvement and prevailing quoted spreads between the NYSE and other market centers, for each size category, and for each sample. Values are medians across pairs and tests of significance are Wilcoxon signed rank tests on the distribution of paired differences.

	Very Small		Small		Medium		Large	
	Improvement	Quote	Improvement	Quote	Improvement	Quote	Improvement	Quote
Market Orders								
NYSE Median	0.74	5.24	-0.20	5.52	-2.06	5.45	-3.18	5.87
Median Difference Between Market Center and NYSE								
Knight	+0.28***	+0.12***	-0.01	+0.07***	-0.37***	-0.37***	-1.20***	-1.09***
Madoff	-0.09	+0.22***	-0.46***	+0.22***	-0.71***	+0.33***	-1.35***	-0.28***
Archipelago	+0.80***	+0.06**	+0.06	-0.08	-0.76***	-0.45***	-1.63***	-1.25***
Boston	+0.28***	+0.18***	-0.10***	+0.08***	-0.45***	-0.07***	-0.96***	-0.54***
Cincinnati	+0.01	-0.04	-0.01*	-0.03	-0.02	-0.20***	-0.72***	-0.99***
Chicago	+0.20***	+0.12***	+0.17***	+0.02	-0.04***	-0.53***	-1.68***	-2.29***
Marketable Limit Orders								
NYSE Median	0.50	3.41	0.28	3.50	0.04	3.95	-0.20	4.08
Median Difference Between Market Center and NYSE								
Knight	+0.42***	-0.34***	+0.19***	-0.25***	+0.12***	-0.03*	-0.01	-0.06**
Madoff	+0.30***	+0.33***	+0.18***	+0.08**	+0.18***	+0.10**	+0.10***	+0.02
Archipelago	+0.46***	-0.42***	+0.22***	-0.53***	+0.06***	+0.00	-0.15***	+0.25***
Chicago	+0.58***	-0.44***	+0.44***	-0.41***	+0.21***	-0.53***	-0.00	-0.52***
Boston	+0.38***	+0.01	+0.23	-0.19***	+0.19***	-0.09***	+0.07	-0.11***
Cincinnati	+0.35***	+0.18***	+0.28***	-0.10	+0.30***	-0.08	+0.18***	-0.19***

*, **, and *** indicate significance at the 10%, 5% and 1% levels, respectively.

Table 7: Time to Execution

This table presents our analysis of the time to execution partitioned by sample and order type. The column labeled "paired vs. NYSE" presents the difference for each pair of stocks for which reports were available. The number of stock pairs is given. Values are medians across pairs and tests of significance are Wilcoxon signed rank tests on the distribution of paired differences.

	Execution Time			Distribution of Execution Time (%)			
	Median Time	Num Pairs	Median Paired Difference	By 10 Seconds	11 to 30 Seconds	31 to 60 Seconds	Over 60 Seconds
Market Orders							
NYSE	23.45			32	44	16	9
Knight	21.07	350	+01.83***	54	18	15	12
Madoff	4.13	184	+16.09***	89	4	4	2
Archipelago	29.07	343	-06.06***	19	48	21	12
Boston	30.20	283	-08.26***	44	19	18	19
Chicago	39.46	343	-16.67***	41	17	14	29
Cincinnati	21.90	148	-01.03***	55	19	17	9
Marketable Limit Orders							
NYSE	48.47			41	36	12	11
Knight	74.16	349	-20.44***	44	18	17	21
Madoff	19.84	173	+13.16***	77	7	9	7
Archipelago	24.82	348	+16.37***	54	31	8	7
Boston	68.35	274	-27.55***	30	18	23	28
Cincinnati	42.36	191	-14.01***	50	19	17	14
Chicago	66.86	336	-12.16***	28	40	12	20

*, **, and *** indicate significance at the 10%, 5% and 1% levels, respectively.

Table 8: Variation in Effective Spreads

This table presents measures of dispersion for effective spreads. The standard deviation and range of values is given for the NYSE. We then present the difference between the given market center and the NYSE. These results exclude any category with less than 10 orders and for which there were observations for less than 10 months. The statistics are calculated across stocks and months. Values are medians and tests of significance are Wilcoxon signed rank tests on the distribution of paired differences

	Very Small		Small		Medium		Large	
	Std. Deviation	Range	Std. Deviation	Range	Std. Deviation	Range	Std. Deviation	Range
Market Orders								
NYSE Median	0.69	2.17	1.11	3.65	1.90	6.13	2.05	6.61
Median Difference Between Market Center and NYSE								
Knight	-0.40***	-1.31***	-0.35***	-1.04***	-0.69***	-2.57***	-0.49***	-1.74***
Madoff	0.06***	0.16***	0.12***	0.38***	-0.02***	-0.13***	-0.26	-1.08
Archipelago	-1.19***	-4.26***	-1.06***	-3.57***	-1.34***	-4.10***	-2.62***	-11.02***
Boston	-0.21***	-0.82***	-0.18***	-0.68***	-0.50***	-1.51***	-0.40***	-1.15***
Chicago	-0.27***	-0.96***	-0.53***	-1.84***	-0.75***	-2.88***	-0.75***	-2.36***
Cincinnati	-0.08***	-0.23***	-0.23***	-0.85***	-0.73***	-2.39***	-0.92***	-2.44***
Marketable Limit Orders								
NYSE Median	0.37	1.19	0.49	1.57	0.73	2.35	0.90	3.10
Median Difference Between Market Center and NYSE								
Knight	-0.42***	-1.41***	-0.43***	-1.54***	-0.40***	-1.36***	-0.42***	-1.22***
Madoff	-0.26***	-0.94***	-0.28***	-0.98***	-0.24	-0.58	-0.42*	-1.17
Archipelago	-0.91***	-3.17***	-0.88***	-3.01***	-0.73***	-2.42***	-	-
Boston	-0.39***	-1.40***	-0.46***	-1.60***	-0.54***	-1.78***	-0.30***	-0.87***
Cincinnati	-0.50***	-1.82***	-0.96***	-3.18***	-0.43***	-1.50***	-0.49***	-1.61***
Chicago	-0.82***	-2.74***	-1.10***	-3.50***	-1.27***	-4.04***	-1.20**	-4.25***

*, **, and *** indicate significance at the 10%, 5% and 1% levels, respectively.

Table 9: Persistence of Spread Performance Relative to the NYSE

This Table presents tests of whether differences in effective spreads between the NYSE and other market centers for subsets of stocks are persistent. Stocks were placed in quintiles based on their July, 2001 differences in effective spreads between the NYSE and the indicated market center (NYSE – market center). This table shows the median difference for the remaining 11 months for the fifth and first quintiles and a Wilcoxon test comparing the difference between the fifth and first quintiles.

	Very Small	**Small**	**Medium**	**Large**
Knight				
Quintile 5 (largest)	-7.95	-1.25	-2.17	-6.94
Quintile 1 (smallest)	-0.68	1.42	0.64	-6.07
Difference	-7.27***	-2.67***	-2.80**	-0.87
Madoff				
Quintile 5 (largest)	0.72	2.27	3.89	4.92
Quintile 1 (smallest)	3.94	9.04	8.37	8.99
Difference	-3.22***	-6.77***	-4.48***	-4.07
Boston				
Quintile 5 (largest)	-2.44	-1.45	-2.45	0.14
Quintile 1 (smallest)	0.89	1.84	-1.05	2.52
Difference	-3.33***	-3.28***	-1.40	-2.38
Chicago				
Quintile 5 (largest)	-4.48	-4.27	-12.29	-11.15
Quintile 1 (smallest)	0.01	-0.80	-2.90	-5.33
Difference	-4.49***	-3.47**	-9.39***	-5.82*
Cincinnati				
Quintile 5 (largest)	-0.37	0.17	-1.31	2.02
Quintile 1 (smallest)	1.48	0.48	-1.47	-1.73
Difference	-1.85**	-0.31	0.17	3.76

*, **, and *** indicate significance at the 10%, 5% and 1% levels, respectively.

Table 10: Regression Analysis

Regression of differences in effective spread on stock characteristics. The dependent variable is the difference between the NYSE effective spread and the effective spread for the indicated market center/order size. The sample in each regression is all monthly observations with both NYSE and market center effective spreads. The explanatory variables are characteristics of stock trading activity during June, 2001 (out of sample). They include the NYSE quoted spread, total executed volume, turnover (volume divided by shares outstanding, and volatility (the standard deviation of June, 2001 daily trade-weighted transaction prices). Standard errors are reported in parentheses and the sample size is given for each regression. We report the adjusted *R*-squared from the OLS regression.

	Knight	**Madoff**	**Boston**	**Chicago**	**Cincinnati**
Intercept	-7.845***	33.605***	-11.670***	-8.480***	10.255***
	(6.834)	(4.575)	(6.053)	(10.418)	(6.535)
Quoted Spread	0.046	0.065	0.207***	0.131	0.046
	(0.057)	(0.080)	(0.064)	(0.092)	(0.086)
Price Improvement	-0.075***	0.058**	-0.104***	-0.144***	0.008
	(0.025)	(0.024)	(0.022)	(0.036)	(0.021)
Turnover	-0.613***	-1.172***	-2.415***	-3.085***	-2.647***
	(0.705)	(0.518)	(0.718)	(1.087)	(0.750)
Log (Market Value)	0.395	-1.927***	0.358	0.551	-0.505
	(0.303)	(0.228)	(0.271)	(0.460)	(0.322)
Volatility (%)	-1.098	1.057**	1.696***	-0.235	0.864
	(0.722)	(0.425)	(0.649)	(1.060)	(0.542)
PIN	0.117	-0.090	0.198*	0.151	-0.019
	(0.129)	(0.069)	(0.112)	(0.203)	(0.084)
Indicator Small	3.429***	2.943***	1.890***	-0.596	
	(1.000)	(0.559)	(0.714)	(1.547)	
Indicator Large	1.108	6.510***		-2.581	
	(1.162)	(0.571)		(1.764)	
Indicator Very Large				-9.577***	
				(2.371)	
Observations	885	502	453	913	141
Adjusted R-Square	0.036	0.345	0.077	0.055	0.050

*, **, and *** indicate significance at the 10%, 5% and 1% levels, respectively.

.

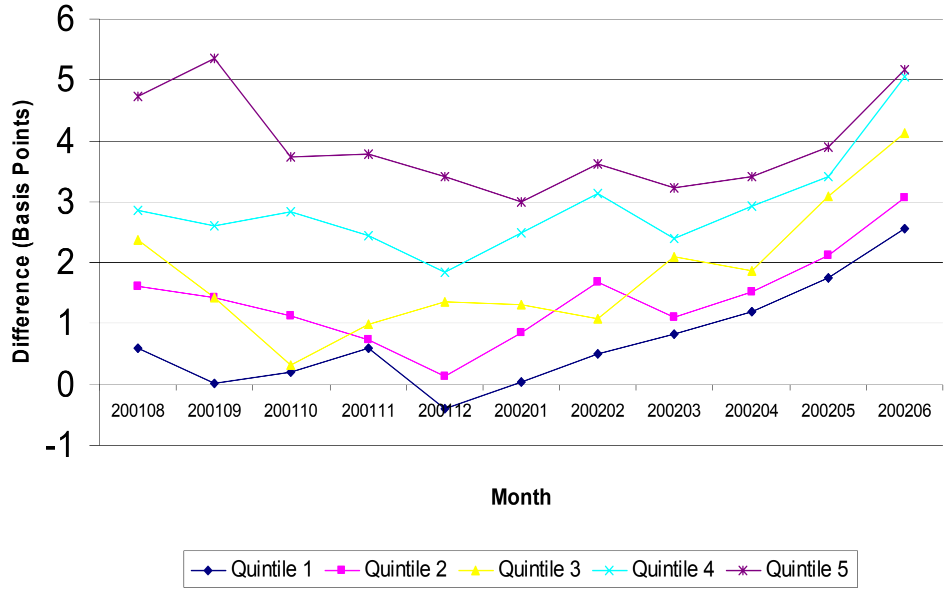

Figure 1: NYSE Effective Spreads Differences

These figures graph the difference in effective spreads (in basis points) between the NYSE and tMadoff for quintiles based on that difference in June 1001, over twelve months (July 2001 to June 2002). Stocks were assigned to quintiles based on their June 2002 differences.

ARE RETAIL ORDERS DIFFERENT?

Charles M. Jones
Graduate School of Business
Columbia University

and

Marc L. Lipson
Department of Banking and Finance
Terry College of Business
University of Georgia

First draft: April 16, 2003
This draft: August, 2004

We thank Ekkehart Boehmer, Rich Lyons, Gideon Saar, Patrik Sandas, Kumar Venkataraman, and participants in the 2003 NBER Summer Institute for helpful comments. An earlier version of this paper was titled "Why NYSE retail orders get better (yes, better) executions." Part of this research was undertaken while Jones was the visiting economist at the New York Stock Exchange, and the NYSE provided the data as well as financial support for this study. The comments, opinions, and errors are those of the authors only. In particular, the views expressed here do not necessarily reflect those of the directors, members, or officers of the New York Stock Exchange, Inc. Contact Charles Jones at Columbia Business School, Columbia University, New York, NY, 10027, or at cj88@columbia.edu, or contact Marc Lipson at the Terry College of Business, University of Georgia, Athens, GA, 30602, or at mlipson@terry.uga.edu.

ARE RETAIL ORDERS DIFFERENT?

Abstract

We use proprietary order-level data to examine the execution quality for retail order flow on the New York Stock Exchange (NYSE). We find that retail and non-retail orders do not have the same average execution costs. Effective spreads for retail orders are smaller than effective spreads for similar orders originating from institutions, program trades, or other sources. We show that this difference is not due to differential treatment of retail orders, to prevailing quote conditions, or to a different distribution in order flow across the a day. Neither institutional nor retail orders appear to chase price trends. The main difference is that non-retail order flow is strongly positively autocorrelated, while retail order flow is close to random over time. As a result, retail order flow appears to be less correlated with information flows; the lower trading cost is driven by lower price impacts. Interestingly, some of the initial price response to retail order flow is reversed in the first ten minutes after execution. VAR evidence indicates that retail orders tend to be followed by institutional orders in the opposite direction. Also, trading volume is lower before and after the execution of retail orders. Finally, non-retail order flow appears to take advantage of liquidity changes, jumping in when spreads narrow, while retail order flow does not. Among other things, these results imply that retail orders should not necessarily be routed to the market center with the lowest average spreads.

ARE RETAIL ORDERS DIFFERENT?

1. Introduction

In the United States, many venues execute individual investors' equity orders. How should an individual (retail) investor gauge whether he or she is getting best execution? In an effort to provide information helpful to investors, the U.S. Securities and Exchange Commission has mandated that each market center disseminate average execution quality measures.[1] These average data are only useful, however, if retail and non-retail order flow obtain similar executions, or if the mix of retail and non-retail order flow is the same at every venue. In this paper, we examine the nature of, and execution measures for, retail and non-retail order flow routed to the New York Stock Exchange. In the process, we shed light on the usefulness of these aggregate execution quality data.

While NYSE retail orders are not explicitly identified as such, and presumably would be treated no differently than other orders, it is still possible for average execution results to differ between retail and non-retail orders. For example, if retail order flow is less likely to arrive at difficult periods than other order flow, execution results for retail order flow will be relatively better. In fact, studies of execution quality data show that lower average execution costs at a given market center are often associated with smaller post-execution price movements.[2] Furthermore, if retail order flow contains less price-relevant information, the differences in execution quality across market centers may reflect the mix of order flow rather than just the procedures and structure of the market center.[3]

---

[1]Public dissemination of these data are required under SEC Rule 11Ac1-5. As stated in U.S. Securities and Exchange Commission Staff Legal Bulletin No. 12, "One of the primary objectives of the Rule is to generate statistical measures of execution quality that provide a fair and useful basis for comparisons among different market centers."

[2] See Lipson (2003), Huang (2002), and Barclay, Hendershott and McCormick (2002). A relation between execution costs and the information content of order flow has been suggested by Demsetz (1968), Glosten and Milgrom (1985), Easley and O'Hara (1987), among others. See O'Hara (1997) and our discussion below for additional details.

[3] Glosten and Milgrom (1985) and Easley and O'Hara (1987) show that trading costs reflect information, suggesting that less informed order flow is less costly to execute. Easley, Keifer, and O'Hara (1996) and Battalio (1997) point out that order routing agreements can be used by market centers to draw more profitable uninformed order flow (cream skimming). Chordia and Subrahmanyam (1995) and Battalio,

We find that the two types of order flow are very different. On average, we find that retail order flow obtains substantially more favorable executions than other order flow in our sample.[4] For example, effective spreads for retail orders in our sample are about 2.60 pennies and are a half a penny lower, on average, than effective spreads for comparable institutional orders. Retail orders also obtain better executions than orders associated with program trading and all other orders. These results are more pronounced for market orders than marketable limit orders and for smaller order sizes. Retail orders have a higher realized spread (a measure of gross trading profits to liquidity providers), which makes it clear why market centers prefer to execute these orders (see Bessembinder and Kaufman (1997) and Huang and Stoll (1996)).

Given the difference in execution results for retail orders, we explore the underlying causes of these differences. First, we verify that the results are not due to retail orders being treated differently. Second, we verify that the differences are not driven by variation in order flow during the day. We also find that retail orders are only modestly correlated with institutional, program and other order flows while these other order flows are much more highly correlated with each other. The explanation for generally lower effective spreads must relate to the timing of order flows.

We examine quoted spreads and price movements immediately around order execution. Non-retail order flow seems better able to time changes in liquidity. Spreads narrow markedly before a non-retail order arrives, while spreads narrow less before a retail order arrives. Clearly, liquidity timing would seem to indicate more favorable execution for non-retail orders, so it cannot explain the narrow effective spreads for retail orders.

We find no evidence that retail or institutional orders are chasing price trends. Interestingly, we find that program trades do tend to follow recent trends (with buys

---

Greene and Jennings (1997), describe the arrangements and agreements that route order flow to various market centers. Related evidence and discussions can be found in Battalio, Greene and Jennings (1997), Bloomfield and O'Hara (1998), Dutta and Madhavan (1997), Bessembinder (1999), Bessembinder (2002a).

[4] We examine a random sample of 60 stocks chosen from the most active 1,000 symbols in November of 2002. The order-level data we obtain provide particularly accurate measures of execution results since the quality measures can acknowledge the time of order submission and can, therefore, incorporate price movements that affect execution results (Harris and Hasbrouck (1996) and Bessembinder (2002b) discuss the advantages of order-level data relative to transaction data). Most importantly, our data allow us to

following price rises and sells following price declines). Most importantly, prices move dramatically during and immediately following execution. As expected, prices move on average against the order (up for buys, down for sells) and, consistent with retail orders being less informed, prices move less for retail orders. For example, between order arrival and order execution, prices move about 0.13 pennies more for institutional than retail orders. Just after execution, the difference is even larger – about 0.88 pennies. These price movement differences more than offset the slightly larger spreads at the time of order arrival for retail orders, resulting in lower effective spreads.

Our results suggest that retail orders arrive relatively more often when prices respond less dramatically to order flow. We examine one possible factor that would contribute to differential price response. Since more active markets are an indicator of more information flows, we look at trading volumes around order arrival and execution. Both before and after order arrival, aggregate order flow is smaller around retail orders. For example, the average share volume for system orders (electronic orders) before a retail order arrives for execution is about 3,263, which is about 458 fewer shares than for institutional orders. Thus, differences in price response may be related to the intensity of trading around execution.

To address the evolution of prices and order flow in an integrated framework, we estimate Hasbrouck (1991) vector autoregressions of quote returns and net order flow by account type. We find that a unit of retail order flow has a small permanent price impact relative to non-retail order flow. Non-retail order flow is strongly persistent, and the steady stream of orders in one direction continues to move the price. There is little such persistence in retail order flow, so prices do not continue to move.

In fact, much of the initial price response to retail orders dissipates (on average) during the following ten minutes. The vector autoregression evidence indicates that institutions are likely to trade in the opposite direction for the first few minutes after a retail execution, bringing prices partially back to their initial levels.

Taken together, our results suggest a sensible explanation for the smaller spreads on NYSE retail orders compared to non-retail orders. On average, retail orders arrive at

---

identify the type of account associated with an order and we distinguish between retail, institutional, program and other orders.

calmer times, and they do not take advantage of short-term price momentum. Retail orders are not timed to take advantage of momentary changes in liquidity. Retail order flow does not persist through time, is largely uncorrelated with non-retail order flow, and in fact is followed by institutional order flow in the opposite direction. All these things make it profitable for market-makers to trade with retail orders, even at narrower spreads.

The remainder of the paper is organized as follows. Section 2 provides a discussion of background issues including the type of data used. Section 3 discusses our sample. Section 4 presents basic results, and Section 5 presents results in a vector autoregression framework. Section 6 concludes the paper.

2. Background

The purpose of this introduction is to provide a brief theoretical and empirical background for discussing statistical evaluation of execution quality. The first section discusses the typical spread measures employed when analyzing trade and quote data. The second section discusses the unique issues and measures associated with order level data.

2. 1 The Measurement and Determinants of Spreads

Spreads are a simple and intuitive measure of trading costs. They reflect the difference between the price at which one sells a security and the price at which one buys. From an investor's point of view, the spread quantifies the round-trip cost of acquiring and then liquidating an investment. Two spread measures are commonly used: the quoted spread and the effective spread.

The quoted spread is equal to the difference between quoted bid and ask prices, expressed either in dollars or as a percentage of the quote midpoint. Quoted spreads reflect a market center’s posted willingness to trade.

In contrast, effective spreads are based on actual transaction prices. The effective spread is defined as twice the distance between the price at which an order is executed and the midpoint of a benchmark quote. The benchmark mid-quote should represent the price that would be obtained in the absence of transaction costs. In most studies that look at transaction data, the benchmark quote is the quote prevailing at the time of execution. Here, we take advantage of our order level data and use as our benchmark the quote in effect at the time of order arrival. Effective spreads measure realized execution costs and differ from quoted spreads due to price or depth improvement. Effective spreads also vary with characteristics of the order, such as order size. This variation cannot be easily reflected in a single quoted spread number.

Both effective and quoted spreads vary over time and across securities and depend on market conditions and stock characteristics at the time an order arrives for execution. For example, the spread may reflect the inventory risk faced by liquidity providers from holding the security at that time.[5] As mentioned, the effective spread also reflects characteristics of the order. Liquidity providers incur less risk when trading with a small order, for example, and thus spreads should vary with order size.

It should be stressed that spreads are not a perfect measure of trading costs for many reasons. For example, many orders are worked over time, and spreads cannot capture the price impact of working an order. Furthermore, spreads ignore commissions and any other market center fees or costs.[6] However, spreads are simple to measure, readily available, and are usually reasonable indicators of actual trading costs for small orders.

Theoretical and empirical studies tend to divide the effective spread into two spread components: the information component and the realized spread. These components are important to drawing inferences about execution quality from spread numbers.

The realized spread is the gross trading revenue to liquidity providers. The realized spread is defined as twice the signed difference between an execution price and the mid-quote five minutes after execution. This mid-quote is designed to measure the post-trade value of the security, and therefore the realized spread reflects the gross trading profit to a liquidity provider from taking the other side of an order.

---

[5] For NYSE stocks, there are many providers of liquidity other than the specialist. In fact, the floor of the exchange encourages competition for liquidity provision. When we refer to the specialist as a liquidity provider, we mean to include all providers of liquidity.

[6] The conclusions drawn from examining spreads may actually differ from the conclusions reached with more extensive data. For example, almost all studies find that spreads decline with a reduction in tick size, but studies of order level data find little if any change (see Jones and Lipson (2002) and Goldstein and Kavajecz (2002)).

The difference between the effective spread and the realized spread reflects the five-minute price impact of the order. The price impact is often referred to as the information component or adverse selection cost, as it presumably reflects the information content of the order (see, for example, Huang and Stoll (1996)). To put it another way, the liquidity provider initially receives the effective spread, loses the information component as prices move against her, and thus earns only the realized spread as gross trading revenue.

These spread components are important to understanding the characteristics of particular order flows. If an order is perceived to be more informed (whether through characteristics of the order or the time of order arrival), then the order will move prices relatively more than another order. Along the same lines, if a trading venue is earning economic rents by successfully cream-skimming uninformed order flow, realized spreads should be relatively large.

Effective spreads and realized spreads are some of the quantities mandated by SEC Rule 11Ac1-5 (Dash5). Dash5 has become a standard for evaluating execution costs at various market centers. Thus, the Dash5 approach seems particularly suited to an investigation of retail order flow, and we follow many of the conventions established by the Dash5 regulations. For example, as mentioned above, we use order arrival times to benchmark effective spreads. We also examine the set of orders for which Dash5 statistics are required. Most importantly, our data allow us to identify the type of account associated with an order, and this allows us to compare retail, institutional, program and other orders.

2. 2 Order Level Data

In this study, the order level data are data captured by the NYSE SuperDOT system for orders submitted electronically. Order level data have two main advantages. First, it is possible to identify many of the characteristics of executed orders, such as the account type and order type. Second, order level data allow a more accurate measure of the full cost of execution since the data reflect order arrival times, not just execution times.

Execution costs should be evaluated as much as possible conditioning on characteristics of an order. We follow the Dash5 rules and partition orders across two dimensions:

- **Order Size.** Orders are classified into four order size groups. These are indicated below along with the designation we use to describe the order size category. As with Dash5 statistics, this study does not examine orders of 10,000 shares or more.

Designation	*Order Size*
Very Small	100-499 shares
Small	500-1,999 shares
Medium	2,000-4,999 shares
Large	5,000-9,999 shares

- **Order Type.** Among other things, the order type reflects a customer's degree of urgency. In general, the more patient a customer, the lower the expected cost of execution (and the longer the expected time to execution). Dash5 distinguishes between the following order types. The definitions below apply to buy orders; sell orders are defined analogously. The applicable quote is the quote prevailing at the time of order arrival.

Order Type	*Description*
Market	No limiting price
Marketable Limit	Limit price equals or exceeds the ask
Non-Marketable Limit	Limit price is below the ask

Throughout the paper, we refer to combinations of order size and order type as a "category". In general, we report average share-weighted execution results within each category. We do not examine non-marketable limit orders. Spread measures are problematic for these orders, and Dash5 regulations do not require their publication.

Dash5 guidelines contain many provisions designed to prevent the statistics from being distorted by unusual orders. For example, orders that require special handling or have unusual restrictions are excluded. Also excluded is any portion of an order executed on a day different from when the order was placed. Orders that meet all the requirements for inclusion in the statistics are referred to as "eligible orders". We follow the NYSE implementation of Dash5 rules to identify eligible orders, and we limit our analysis to these orders.

The system data include an indicator of the account type originating the order. We partition the indicators into four groups: retail, institution, program, and other. The orders in the "other" category are generally of less interest but are included for completeness. The account type partitions are:

Account Type Designation	*Description*
Retail	Agency orders that originate from individuals
Institution	Agency orders that do not originate with individuals
Program	Orders associated with program trades.
Other	Mostly orders where NYSE members are trading as principal.

Account types are coded by the submitting broker-dealer based on a set of regulations issued by the NYSE. While they are generally unaudited, these classifications are important to the NYSE and to broker-dealers because they are required for a number of compliance issues. For example, NYSE Rule 80A suspends certain types of index arbitrage program trading on volatile trading days, and account type classifications are important for enforcing this ban. The specialist and traders on the floor do not, however, observe this account type indicator for an incoming system order. In general, these market participants observe only the type, size, and limit price (if applicable) of an order. It is possible for the specialist to research a particular order in real-time and obtain the account type as well as information about the submitting broker. However, this takes a number of keystrokes and requires a certain amount of time, and given the pace of trading on the exchange and our conversations with specialists, we conclude that the account type indicator is seldom if ever observed before execution.

We believe we are the first academic researchers to study execution quality and order timing for these different groups. Using proprietary Nasdaq data, Griffin, Harris, and Topaloglu (2003) classify trades as either individual or institutional, but they focus instead on momentum trading at the daily horizon for each of these groups. Battalio, Hatch and Jennings (2003) examine compare retail order flow sent to a third-market dealer with similar order flow sent to the New York Stock Exchange.

3. Sample and Summary Statistics

This study examines a sample of 60 symbols for which NYSE system order data were gathered. The sample was chosen as follows. First, NYSE executed share volume for all NYSE listed common equity symbols trading above $5.00 a share was gathered for November of 2002. From this sample, the 1000 most active symbols were identified and were divided into trading volume quintiles. From the most active quintile, we chose 20 symbols at random. From each of the remaining four quintiles, we choose 10 symbols at random. Appendix A lists the symbols studied along with their November consolidated trading volume. Order level data for this sample were collected for every order in the month of November 2002 (twenty trading days).

Table 1 presents summary statistics for the sample. The statistics are given for the full sample and then separately for the 20 symbols from the most active quintile and the remaining symbols. The first part of the table describes firm and share characteristics. Note that the active symbols have a higher share price, greater market capitalization (over $34 billion on average), and by construction a much higher trading volume – over ten times more active than symbols in the less-active subsample. Note that daily trading volume is based on the consolidated tape and includes all trades at all market centers.

The second part of Table 1 describes all NYSE system orders in our sample stocks. It gives the executed share volume for all orders and for relevant partitions.[7] Note that these executed order data count buy and sell orders separately. Hence, overall volume figures should be compared to twice the consolidated volume from the first part of the table. Overall, about 36% of (twice) consolidated volume involves NYSE system orders.

The last part of Table 1 describes the Dash5 eligible orders that make up our sample. Compared to twice the consolidated volume from the first part of the table, our sample covers about 17% of total volume. These numbers are much lower because we follow the Dash5 selection criteria and limit the analysis to system market and

[7] We could also have provided results on orders rather than executions. For market orders, order volume and executed volume will be almost identical. However, for marketable limit orders, order volume will

marketable limit orders below 10,000 shares. The sample excludes large institutional orders and orders sent to floor brokers. Since the focus of the paper is retail orders, and our methodology seeks similar institutional orders as a basis for comparison, excluding these large or difficult orders should not affect the results.

About 55% of the executed shares in the sample are market orders. The remaining 45% are marketable limit orders. In addition, retail order flow represents only 4% of the executed shares in the sample. There are several reasons this percentage is so low. First, retail orders tend to be relatively small. Second, while most institutional orders and program trades are routed to the NYSE, a substantial amount of retail order flow is either internalized or channeled to alternative venues. Unfortunately, we do not have order level data on retail orders executed elsewhere. Thus, we do not know whether NYSE retail orders are similar to retail order flow that is internalized or sent to other venues. Finally, the account type codes are imperfect. Based on conversations with exchange officials, we are confident that nearly all orders marked as retail are in fact submitted by individual investors. However, some orders submitted by individual investors are not recorded as retail orders, particularly if they are executed by an NYSE member firm on behalf of another broker-dealer.

It is typically argued that retail order flow is less informed than other order flow. To take this to the extreme, if retail order flow arrives randomly over time and is uncorrelated with contemporaneous informed order flow, then it must be uninformed. Table 2 assesses this null hypothesis by calculating the autocorrelation of and the correlation between the net order flow of different account types. For the 60 stocks in our sample during November 2002, we aggregate all orders of a given account type that execute in the same minute and measure net order flow as the excess of buys over sells during that minute. Net order flow is measured in shares as well as orders executed. The resulting time series has 7,800 observations for each account type (390 minutes per trading day × 20 trading days).

Table 2 contains the relevant correlations and autocorrelations, and the evidence rejects the extreme null. Like other account types, retail order flow is positively

---

exceed executed volume since the market may move away from a marketable limit order before it is executed. Lipson (2003) provides more detailed results on system order disposition.

autocorrelated, with a one-minute autocorrelation of 0.10. Retail order flow is also positively correlated with order flow from other account types. If measured in shares, retail order flow has a contemporaneous correlation of 0.05 with institutional order flow, and 0.06 with program trades. However, all of these correlations are extremely small, and they are only marginally statistically different from zero. Economically, retail order flow is quite close to being random over time.

Though the absolute correlation levels are different from zero, we might expect relative differences if retail order flow is less informed than other types of order flow. More precisely, we would expect non-retail order flow to be more highly correlated if the different classes of non-retail order flow are motivated by the same information flows. Table 2 shows that, indeed, retail order flow is much less correlated with other order flow. This is particularly true if we consider correlation in the number of orders rather than the number of shares. For example, different types of non-retail orders have correlations that range between 0.30 and 0.55, while the correlation of retail order flow with other account types is between 0.03 and 0.06. In addition, we find that retail orders are the least autocorrelated, and institutional orders the most, with a one-minute autocorrelation coefficient of 0.34.

Similar evidence emerges from the cross-autocorrelation of retail and non-retail order flow. Institutional, program, and other non-retail order flows have similar characteristics, while retail order flow is very different. Retail order flow has almost no predictive power for non-retail order flow in the next minute, with cross-autocorrelations between 0.027 and 0.041. Retail orders seem to lag other orders slightly, as the cross-autocorrelations between non-retail order flows and lagged retail order flow are a bit higher, ranging from 0.062 to 0.079. Of course, the correlation evidence is only suggestive and needs to be confirmed by a closer look at the execution of retail orders.

4. A Detailed Look at Retail Order Execution

4. 1 Execution Quality Measures

Table 3 presents a summary of standard execution quality statistics for our sample by account type. These are simple share-weighted averages across the whole sample. Results are presented for the whole sample, by order type, and by order size. We also indicate the total shares executed in each category.[8] Finally, we include tests of the hypothesis that the given value differs from the corresponding value for retail order flow. Throughout the paper, we conduct statistical inference by aggregating all observations on a single day and base statistical tests on the variation in the weighted time series of daily observations, thus assuming independence across days but not across orders.

For the whole sample, the average effective spread for retail orders is 2.60 cents. This compares to 3.07, 3.05 and 2.46 for institution, program, and other order types. The retail orders have reliably lower spreads than institutional orders and program trades. The differences are substantial – almost half a penny separates institutional and retail spreads. Generally, the results for realized spreads and information component are similar to those in Lipson (2003) – realized spreads are small and the information component is large. The notable difference here is that realized spreads are substantial for retail order flow. The realized spread is over a penny whereas, for example, it is negative (on average) for institution orders. This illustrates the trading revenue that might be available to a market center that can attract retail order flow. From narrow effective spreads and high realized spreads, it follows directly that retail orders have little price impact. Average price impacts are 1.38 cents for retail orders, compared to 3.22 cents for institutional orders and 2.66 cents for program trades. We often refer to the price impact as the information component, because all else equal, a smaller price impact implies that retail orders are relatively more "uninformed". However, it is worth noting that these are simple averages and make no attempt to set all else equal. For example, perhaps retail orders pay smaller spreads because they are simply smaller than other orders on average.

The quoted spread at the time of order execution is reliably smaller for retail than institution orders, though reliably larger than for program and other orders. As we shall see later, these results change considerably once we apply appropriate control variables.

To begin to control for differences in order flow characteristics, we calculate execution quality measures for various partitions of the data. When we partition by order type, the results are weaker for marketable limit orders (see Peterson and Sirri (2002) for

---

[8] This differs from Table 1, which presents daily averages by symbol. To obtain the totals in Table 3, multiply Table 1 values by 20 (days) × 60 (symbols).

issues related to the execution costs of marketable limit orders). For example, the effective spread difference between retail and institutional order flow is about 1.20 cents for market orders, but only about 0.30 cents for marketable limit orders. It should be noted that individuals submit proportionally far fewer marketable limit orders than do the other account types – the market and marketable limit breakdown is more than 80/20 for retail orders and roughly 50/50 for other account types.

A more important control is order size. For smaller order sizes, retail effective spreads are statistically narrower. For the smallest orders of less than 500 shares, retail effective spreads average 1.69 cents, while institution orders' effective spreads average 2.57 cents. For the large orders in our sample (over 5,000 shares), there is no reliable difference in effective spreads between retail and either institution or program trades. As expected, effective spreads are increasing with order size (consistent with Easley and O'Hara (1997)).

These simple controls may not be enough. One possibility is that retail investors trade more in liquid stocks. For example, if retail orders are proportionally more likely in symbols with lower spreads, then effective spreads would be smaller. Table 4 contains the analysis with a full set of control variables. The reported numbers focus on retail orders relative to institutional orders; results for other account types are generally similar.

Table 4 presents a comparison of retail and institution orders using four control variables. Specifically, all orders are aggregated (using a share-weighted average) if they are on the same date in the same stock with the same order size category, same order type, and same account type. Pairs are formed when there are both retail and institutional orders that match along all four other dimensions, and the table reports equal-weighted averages across these pairs. Again, statistical inference is performed using the 20-day time series of these average pair-wise differences. It should be noted that we do not necessarily have observations for every category, so we also report the number of pairs in our analysis.[9]

[9] The maximum number of pairs would be equal to 20 (days) × 60 (symbols) × 2 (order types) × 4 (order sizes) = 9,600. Thus, for all orders, we only have pairs for about half the possible categories.

Across all such pairs, the average effective spread for retail orders is 2.81 pennies. This is 0.50 cents less than the average for institutional orders.[10] We find that effective spreads are reliably smaller than effective spreads for institutions in every case except for the largest order size, where the differences are not statistically reliable. Once again we see that realized spreads are much larger and the information component much smaller for retail orders.[11] Finally, after controlling for stock, trading day, order type, and order size category, it appears that retail orders are submitted when the spread is relatively wide, while institutional orders are submitted when the quoted spread is 0.23 cents narrower. This could indicate that institutions are closely monitoring liquidity as it varies through time, and they pounce when the market is relatively liquid. We return to this issue later in greater detail.

4. 2 Are Retail Orders Treated Differently?

Among other things, the previous section establishes that cheaper retail executions are not an artifact of individuals trading more liquid stocks or submitting smaller orders. In this section, we address another possibility – that retail orders sent to the NYSE are actually treated differently by the specialist or other intermediaries. For example, Benveniste, Marcus, and Wilhelm (1992) argue that the lack of anonymity in the NYSE's floor-based market structure allows the specialist to separate relatively informed and uninformed order flow, thereby reducing adverse selection risk. Their model implies that uninformed orders should have lower trading costs, which is consistent with the results found here.

However, in the case of retail order flow, differential treatment seems unlikely, since these orders arrive at the trading post electronically, and the specialist cannot easily observe the account type indicator, though he may be able to draw some inference from, say, the size and timing of the order. However, to rule out differential treatment, we construct matched pairs of retail vs. non-retail orders that occur within 5 seconds of each

---

[10] The magnitude of the spreads is much larger in Table 4 than Table 3 because we are equally weighting across symbols rather than share weighting. Thus, Table 4 reflects to a greater degree the conditions for smaller and less active symbols.

[11] Interpreting the magnitude of values in Tables 3 and 4 is somewhat complicated. In Table 3, the results are those that would be expected for a trader whose orders are distributed across symbols and days in line

other. These matched pairs are in the same symbol and are also the same order type (market or marketable limit), same direction (buy or sell), and also in the same order size category.

Results of the matched order analysis are given in Table 5. There are 3,306 order pairs that match retail and institution orders, and fewer retail orders that match the other account types. We report equal-weighted averages across all relevant pairs. The execution quality measures for retail orders are generally indistinguishable from the spreads for other account types. Retail orders have slightly lower effective spreads than matched program orders, but this difference is only marginally significant at the 10% level, and the result may be due to imperfect controls (e.g., matched orders need not be exactly the same size or arrive at exactly the same time). Overall, the evidence indicates that orders that arrive around the same time receive the same execution. Thus, it must be the case that retail orders execute at tighter spreads because they arrive at different times than other orders. Our goal in the rest of the paper is to explore market conditions before, during, and after retail order arrival.

4. 3 Time-of-day Differences

One simple possibility is that retail orders tend to trade at different times during the trading day. In general, spreads follow a U-shaped pattern during the trading day. They are higher at the start of trading, decline over the next few hours, and rise again near the close. If retail orders are predominantly executed in the middle of the day, then this might explain the results. Figure 1 presents the distribution of trading volume over the course of the day. Share volume is aggregated by 5-minute intervals, and the plot records the proportion of total volume in the sample that occurs during that 5-minute interval for that account type. All account types have very similar trading patterns. Retail order flow closely tracks the intraday regularities in other order flows. There are no discernible time-of-day differences in order flow.

---

with aggregate volume for that trader type. The results in Table 4 are what a trader might expect for a randomly chosen symbol and trading day.

4. 4 Quoted spreads before and after execution

Next we explore a number of possible determinants of execution quality differences. In this section we examine quoted spreads and in the next section we examine price changes.

We begin by examining conditions immediately surrounding the time of order arrival and execution. Figure 2 presents the quoted spread at 15 one-minute intervals prior to and at order arrival time, and at 15 one-minute intervals at and subsequent to order execution. The time between order arrival and execution (denoted in the graph by a gap) varies from order to order. All one-minute intervals are calculated relative to the order arrival time (for pre-arrival) and order execution time (for post-execution). The graph only includes orders that arrive later than 15 minutes after the start of trading and are executed at least 15 minutes before the close of trading.

Other than this filter, we apply control variables and aggregate orders following a procedure identical to that used for Table 4. That is, all orders are aggregated (using a share-weighted average) if they are on the same date in the same stock with the same order size category, same order type, and same account type. Pairs are formed when there are both retail and non-retail orders that match along all four other dimensions, and Figure 2 reports equal-weighted averages across these pairs. Statistical inference is performed using the daily time series of these average pair-wise differences.

Figure 2 shows that market conditions are similar 15 minutes before the order arrives. There is little difference in quoted spreads fifteen minutes before a retail vs. non-retail order. The notable feature of this graph is what happens just before retail order arrival. For the non-retail account types, the quoted spread declines markedly in the minutes just before order submission. In contrast, there is relatively little change in quoted spreads in the minutes before a retail order. Thus, it would appear that non-retail orders are timing their order arrivals to take advantage of changes in quoted spreads. For example, these orders may be picking off a limit order that has just arrived to narrow the spread. Retail orders, on the other hand, exhibit less liquidity timing.

At the time of order execution, quoted spreads are narrower for institutional orders than they are for similar retail orders. This matches the evidence in Table 4.

In all cases, quotes widen subsequent to order execution. For retail orders, the quotes narrow back down within a few minutes, whereas spreads do not narrow as much for non-retail orders. Once again, this is consistent with the timing of order flow to take advantage of temporary improvements in spreads. The slow decline may reflect the amount of time it takes for the book to fill back in.

Are non-retail orders simply quicker at pouncing on improved liquidity? To address this question, Table 6 looks at the time between the most recent liquidity improvement and the arrival of the market or marketable limit order for different account types. We look at the time between the last quote change and order arrival, the time since the last limit order arrival that improves the existing quote, and the time since the last quote narrowing. The general empirical strategy is the same as for Table 4. That is, all orders are aggregated (using a share-weighted average) if they are on the same date in the same stock with the same order size category, same order type, and same account type. Pairs are formed when there are both retail and non-retail orders that match along all four other dimensions, and Table 6 reports equal-weighted average times or price changes across these pairs. Statistical inference is performed using the daily time series of these average pair-wise differences.

Table 6 shows no evidence that institution or program trades are quicker at taking advantage of liquidity improvements. For example, the most recent improving limit order arrives an average of 94 seconds before a retail market order arrival, while the corresponding figure for institutional orders is almost identical at 93 seconds. There is some evidence that other (non-retail, non-institution, non-program) orders are quicker, at 83 seconds since the last improving limit order vs. 91 seconds for the matched sample of retail orders. These are mostly proprietary trades by member firms, so it makes sense that these entities would be the quickest on the trigger following an improvement in liquidity.

Overall, there is no evidence that institutions or program trades are faster at taking advantage of improved liquidity. Instead, the evidence suggests that institutions are waiting for substantial improvements in price before submitting a market order.

4. 5 Price changes before, during, and after execution

In addition to timing liquidity, perhaps some order submitters are responding to recent price changes in an effort to time the market. Also, market movements may affect the willingness of market participants to provide liquidity. For example, price movements might affect inventory holdings. We explore this possibility in Table 7 and Figure 3, where we examine price changes before order arrival, between order arrival and execution, and after execution. Table 6 breaks the order execution process into three parts that are analyzed separately:

- *Pre-Arrival* This is the five-minute period before an order arrives at the NYSE.

- *Execution* This period begins when the order arrives at the exchange and ends when the order is reported as executed. This takes an average of about 20 seconds. This interval matches the period used to calculate the effective spread.

- *Post-Execution* This is the five-minute period after an order is executed. This interval matches the period used to determine the realized spread.

We are most interested in the movement of prices around order arrival and execution. We measure this using momentum, which is defined as the average signed change in the midquote return (measured in cents) over the relevant time period.[12] Returns are signed by multiplying by 1 for a buy order and –1 for a sell order. That is, if prices are moving up during a buy order execution or down during a sell, momentum is positive. When positive momentum occurs before order execution, it reflects an adverse move in prices for the order submitter. However, when positive momentum occurs after order execution, the price move favors the order submitter. There are several possible sources of momentum during and after an order executes. The momentum could be the result of the executed order itself (reflecting prevailing market conditions), it could be due to other orders arriving at the same time, it could be due to price changes in other stocks, or it could be any other new information that causes the specialist to change the quotes.

---

[12] We also examined the volatility of returns around order arrival and execution. Results are not reported, because there were no discernible patterns in volatility before, during, or after order execution.

The basic idea is to see whether some classes of traders are responding to price trends, to see whether some traders are better able to anticipate short-term price moves, and to document the extent of price responses to orders. On average, program trades in our sample are short-term trend chasers, with prices moving a statistically significant 1.26 cents in the five minutes before order arrival.[13] Institutions also trade in the direction of previous price moves, while retail buy (sell) orders tend to arrive after modest and statistically insignificant price declines (increases) averaging 0.35 cents.

To compare momentum across account types, we again use the Table 4 approach to control for the symbol traded, trade date, order type, and order size category. In terms of five-minute pre-arrival momentum, program trades are statistically distinct from retail orders. However, pre-arrival momentum for retail is not significantly different from that of institutional or other order flow.

Table 7 also reveals that the most interesting quote changes happen during execution. Between order arrival and execution, quoted prices all move in the same direction as the order (up for buys, down for sells). But the price changes are the smallest for retail orders. After controlling for stock, trading day, and order characteristics, average momentum during retail order execution is always statistically lower than average momentum for other account types. Retail vs. institutional momentum is 0.34 vs. 0.61 cents, retail vs. program momentum is 0.31 vs. 0.70 cents, and retail vs. other momentum is 0.32 vs. 0.54 cents.

These differences in price moves during execution account could be the explanation for the difference in the effective spread paid by market order and marketable limit order submitters. To see this, consider again the retail vs. institutional comparison. The momentum numbers during execution (0.34 cents retail vs. 0.61 cents institutional) imply that this slippage contributes 0.68 cents to the (round-trip) cost of a retail trade and 1.21 cents to the cost of an institutional trade. The difference between the two is 0.53 cents, which is about the same as the 0.50 cent difference in effective spreads for these two account types from Table 4. This is also consistent with the large information

[13] Share-weighted average momentum is calculated for all orders in the same stock on the same day with the same order type, order size category, and account type. The table reports equal-weighted averages for all non-empty classifications of a given account type.

component we observe for non-retail orders; interestingly, some of this information is already being incorporated into price prior to execution.

One might worry that momentum during execution might depend on the time required to execute the order. But this does not seem to explain the differences between retail and non-retail momentum. The bigger price moves in non-retail orders are not the result of large systematic differences in the time to execution. The average time from order arrival to order execution is about 20 seconds for all account types.

In the first minute after execution, Table 7 shows that prices move more for non-retail orders. For example, retail vs. institutional price moves are 1.81 vs. 2.56 cents, a statistically reliable difference. Over the next four minutes, the contrast between retail and non-retail orders becomes especially stark. Following a retail order, prices revert by 0.49 cents during this interval. In contrast, comparable institutional orders show a continued average price move of 0.53 cents in the direction of the original order. The net result over the 5-minute post-execution period is not surprising; it is simply another manifestation of the greater information component for non-retail orders found in Table 4. But the pattern of adjustment is very different, with reversion in prices only after retail order executions.

Figure 3 tells the same general story graphically. It presents the cumulative price impact (cumulative momentum) around order arrival and execution. The graph begins fifteen minutes prior to order arrival, extends fifteen minutes subsequent to order execution, and documents the price change each minute. Orders are aggregated as in Table 4; to make comparisons across types, we control for symbol, trade date, order type, and order size category. Also included is a single point that captures quote changes between order arrival and execution, regardless of the elapsed time between arrival and execution.

Figure 3 shows that, in aggregate, neither retail nor institutional orders are chasing trends. The figure confirms that program trades chase recent trends, though it also indicates that these trends have been short-lived, beginning on average 10 minutes prior to the order. Figure 3 shows that institutional orders have a bigger price impact than retail orders. Figure 3 also confirms the Table 7 evidence of mean reversion in prices.

While the price impact for institutional orders is permanent at least 15 minutes out, after retail orders prices tend to partially revert to their earlier levels.

Overall, Table 7 and Figure 3 tell a very interesting story. Program trades tend to be short-term trend chasers, while retail and institutional orders do not exhibit any strong trend-chasing or trend-reversing behavior on average. However, during execution, prices start to move in the direction of trade, and they move much more for institutional orders. After an institutional order, the mini-trend continues, as prices continue to move in the same direction. After a retail order, however, prices move less initially, and they tend to revert significantly over the next 10 minutes. This price reversion is an important part of the high realized spreads on retail orders at a five-minute horizon, and the evidence indicates that realized spreads on retail orders are even higher at a horizon of ten minutes post-trade.

These results indicate that, for whatever reason, retail orders tend to arrive when prices respond less dramatically to order flow. What might contribute to a differential price response? It is possible that non-retail orders arrive in more active markets. These active markets might be associated with greater information flows. Active markets might also increase the amount of inventory risk borne by the specialist or other liquidity suppliers.

To investigate this, we look at trading volumes around order arrival and execution. As in the rest of the paper, we compare similar retail vs. non-retail orders, controlling for order type, order size category, symbol, and trade date. We look at net signed trading volume (buyer-initiated less seller-initiated volume) as well as unsigned trading volume.

The results are in Table 8. Non-retail orders tend to execute at relatively active times. Both before and after order arrival, aggregate system volume is smaller around retail orders. For example, average system order volume (electronic orders) is about 3,263 shares in the minute before a retail order arrives, which is about 458 fewer shares than for institutional orders. There is a similar differential during the minute after order execution. The difference in signed volume is even more dramatic. In the minute before a retail order arrives, net signed volume averages 169 shares in the same direction, compared to 702 shares in the minute before an institutional order. This differential persists during order execution and in the minute after order execution. This confirms the evidence in Table 2. Marketable retail orders are close to random over time and are largely uncorrelated with order flow from other account types, while institutional orders tend to cluster in the same direction over short intervals of time. In addition, the unsigned volume evidence indicates that retail orders tend to arrive in calmer times. Thus, it is not surprising that prices do not adjust as strongly in response to a given retail order.

5. Vector autoregressions

In most of the previous section, we take a typical market order or marketable limit order and examine the nearby behavior of prices, spreads, and volume. Table 2 gives some hints about how order flow is related to nearby order flow but does not consider order flow and prices at the same time. In order to model the evolution of order flow and prices over time in an integrated framework, we turn in this section to a vector autoregression of trades and quotes.

Based on Hasbrouck (1991, 1996), we construct a vector autoregression that distinguishes between different types of order flow (see also, for example, Hendershott and Jones (2003)). This involves separate equations for the order flow of each account type, yielding five equations in total: a quote midpoint equation, an equation that describes the evolution of retail signed order flow, and so on for institutional, program, and other order flow. Specifically, for a given stock define x_t^I to be the sum of the signed order flow in shares (positive for market and marketable limit orders to buy and negative for sells) during the one-minute interval t for retail account types. Similarly, define x_t^I for institutional account types, x_t^P for program trades, x_t^O for other order flow, and define r_t to be the percentage change (log return) in the quote midpoint during interval t. The following VAR with five lags is estimated for each stock for each trading day:[14]

[14] For actively-traded stocks, the results are insensitive to the length of the interval and the number of lags estimated in the VAR. VAR estimation is limited to the most active stocks, because the lack of order flow in other stocks makes it very difficult to pin down their transition matrices.

$$\begin{bmatrix} x_t^R \\ x_t^I \\ x_t^P \\ x_t^O \\ r_t \end{bmatrix} = \sum_{j=1}^{5} \Phi_j \begin{bmatrix} x_{t-j}^R \\ x_{t-j}^I \\ x_{t-j}^P \\ x_{t-j}^O \\ r_{t-j} \end{bmatrix} + \varepsilon_t \tag{1}$$

where Φ_j is a 5 x 5 autoregressive matrix and ε_t is a 5 x 1 vector of innovations with covariance matrix Ω.

The VAR is inverted to get the vector moving average representation in order to focus on the impulse response functions to shocks in various types of order flow. Among other things, this allows us to measure the permanent price impact from a shock to each trade equation, as well as the effect of an order flow shock on later order flow of the same or different account type. As discussed in Hasbrouck (1991), this method is robust to price discreteness, lagged adjustment to information, and lagged adjustment to trades. Note that in this case, the order flow variables include only system market orders and marketable limit orders of 10,000 shares or fewer. Executed floor orders are excluded because we lack account types and order types for these executions.

We calculate the response of each variable to a unit shock in net order flow of a certain account type, assuming that all other types of order flow are zero. The unit shock is normalized to 1,000 shares, and contemporaneous quote midpoint changes are included.[15] There is a separate VAR for each trading day, so we average the impulse response curves across the 20 trading days in our sample and report the average impulse response. Estimated impulse responses are assumed independent across trading days, and 95% confidence intervals are constructed using the variability in the impulse response across days. Impulse reponses are calculated for a total of twenty minutes following the initial shock.

Figure 4 reports results for a single large stock, ExxonMobil. This is the third-largest American company by market capitalization, a member of the Dow Jones Industrial Average, and the third most-active stock by share volume during November 2002. Its VAR results are also representative of the broader sample of active stocks.

[15] We accomplish this by working with orthogonalized residuals, where the order flow type being shocked is the penultimate variable, and the midpoint return is the last variable.

Figures 4a, 4b, 4c, and 4d give impulse response functions for shocks to retail order flow, institutional order flow, program order flow, and other order flow, respectively. Non-retail order flow is qualitatively similar. The strongest finding is that own order flow shocks persist over time. For example, a 1,000 share institutional buy tends to be followed by institutional purchases totaling an additional 642 shares over the next 20 minutes. Effects across order types tend to be much weaker. For example, a shock to institutional order flow alone does not tend to be followed by order flow in the same direction from other account types.

The same size trade has very different permanent price impacts for different account types. The permanent price impact is 0.13 basis points for a retail order flow shock of 1,000 shares, 0.64 basis points for an institutional order, and 0.44 basis points for a program order flow shock. The retail price impact is statistically distinct from the other two.

To help us understand why retail price impacts are so low, Figure 4a shows the response to a unit shock in retail order flow. Unlike institutional and program order flow, there is much less persistence in retail order flow. On average, a 1,000-share buy order is followed by only about 40 additional retail shares in the same direction over the next 20 minutes. The cumulative price response shows an initial price move of about one-half basis point in the direction of the trade. Only a little order flow follows in the same direction, so it is not surprising that prices do not continue to adjust in the same direction. In fact, the initial price move reverses quickly, with more than half of the initial move reversed over the next three minutes.

Why does this reversal take place? The answer lies in institutional order flow. In the first five minutes following a retail order execution, institutional order flow arrives in the opposite direction. This institutional order flow is fairly substantial: an unexpected retail order of 1,000 shares is followed by more than 400 institutional shares in the opposite direction. This countervailing order flow is significant and continues in the same direction for the entire twenty minute period studied. We cannot, of course, be sure why institutions are trading in the opposite direction, but this institutional order flow appears to explain the strong temporary component in the cumulative price response.

It is also worth noting that there is also a small reversal following program order flow (Figure 4c). This too appears to be driven by institutional order flow in the other direction, though the magnitudes are smaller. A shock of 1,000 shares in program order flow tends to be followed by 81 institutional shares in the opposite direction in the next two minutes, when the reversal occurs, and 227 institutional shares in the opposite direction over the next 20 minutes. However, it is important to note that program order flow is positively autocorrelated, with the unit shock of 1,000 shares followed by an average of 625 more program shares in the same direction over the next twenty minutes. This is likely to limit the effect of institutional trades in the opposite direction. In any case, program trades have substantial permanent price impacts, so they are qualitatively very different from retail orders.

Next, we report impulse response functions that are aggregated across stocks. For each of the twenty most active stocks in the sample, impulse response functions are calculated for each stock for each trading day, standardized to reflect the impact of an order flow innovation of 1,000 shares. An equal-weighted cross-sectional average impulse response function is calculated for each trading day, and these are then averaged across trading days. Time-series independence of the daily cross-sectional averages is used to conduct statistical inference.

The results are in Figures 5a, 5b, 5c, and 5d for retail, institutional, program, and other order shocks, respectively. The results are qualitatively similar to the single stock counterparts in Figure 4. For non-retail order flow, there is strong own order flow persistence, and modest positive cross-persistence in various types of non-retail order flow. Only retail order flow engenders order flow in the opposite direction. On average across these twenty stocks, an unexpected marketable order of 1,000 shares results in about 130 institutional shares in the opposite direction over the next five minutes, though there is only marginal statistical evidence that the institutional response is different from zero.

Permanent price impacts continue to differ across account type. The pooled average permanent price response to a unit shock of 1,000 shares is lowest for retail orders, at 1.33 basis points. Corresponding figures for institutional orders are 1.82 basis points and 2.37 basis points for program orders.

Figure 5a shows that the price reversal following retail orders is not unique to ExxonMobil. For the twenty active stocks, the response in quote midpoints maxes out at 1.91 basis points after one minute, and about one-third of this initial price response reverses in the next twenty minutes. Only retail order flow engenders such a price reversal.

Overall, the VAR evidence confirms that retail orders have smaller price impacts, and it confirms that the permanent price impact is much lower than the price impact one or two minutes after the order is executed. It also reveals at least part of the mechanism behind this quote reversion: market orders and marketable limit orders in the opposite direction are being sent by institutional accounts.

6. Conclusions

In this paper, we use proprietary system order data from the NYSE to examine the execution quality of NYSE retail order flow. It turns out that retail orders get better executions, on average, than similar non-retail orders. Effective spreads for retail orders are smaller than effective spreads for comparable orders originating from institutions, program trades, or other sources. Nevertheless, retail orders have larger realized spreads, which explains why other market centers are trying to siphon off these orders. This also implies that retail orders have a smaller price impact, which we confirm using impulse response evidence from vector autoregressions.

We rule out a number of explanations for these results. Retail orders are not treated any differently; comparable retail and non-retail orders that arrive at nearly the same time obtain similar executions. Retail and non-retail orders are distributed similarly throughout the day. The results are not driven by differences in quoted spreads at the time of execution, which are actually slightly larger, on average, when retail order flow arrives. In fact, we find that non-retail orders are able to time liquidity, jumping in when quoted spreads narrow substantially. But this effect goes the wrong way, so it cannot explain lower effective spreads for retail orders. Finally, neither institutional nor retail orders are chasing price trends, on average (though program trades do tend to chase them).

The explanation appears to be related to two important differences between retail and institutional orders. First, prices tend to rise (fall) immediately after any kind of buy (sell) order is executed, but the price reaction is smaller for retail orders. There is also a temporary component. For ten minutes after a retail execution, prices tend to partially revert toward their earlier levels. Vector autoregressions reveal that this reversion is at least partially due to institutional order flow in the opposite direction in the first few minutes following a retail order arrival. Second, retail orders seem to arrive at relatively calm times. There is more volume both before and after a non-retail order execution.

Most of this paper focuses on the search for what makes retail order flow different. But the stark differences in retail vs. non-retail order execution quality have important policy implications. Most importantly, Dash5 statistics may not provide sufficient information for routing retail order flow. For example, it is misleading to compare aggregate NYSE execution quality to that of market centers that execute predominantly retail order flow. Unfortunately, only aggregate statistics are required under Dash5 rules, and this promotes "apples-to-bicycles" comparisons. Among other things, our results suggest the New York Stock Exchange should voluntarily publish Dash5 statistics on its retail order flow so order-routers and others can draw meaningful comparisons between the NYSE and retail-oriented market centers.

References

Barclay, Michael, Terrence Hendershott, and Timothy McCormick, 2003, Electronic communications networks and market quality, working paper.

Battalio, Robert, 1997, Third market broker-dealers: cost competitors or cream skimmers?, *Journal of Finance* 52, 341-352.

Battalio, Robert, Jason Greene, and Robert Jennings, 1997, Do competing specialists and preferencing dealers affect market quality?, *Review of Financial Studies* 10, 969-993.

Battalio, Robert, Brian Hatch, and Robert Jennings, 2003, All else equal? A multi-dimensional analysis of retail, market order execution quality, Journal of Financial Markets 6, 337-362.

Benveniste, Lawrence M., Alan J. Marcus, and William J. Wilhelm, 1992, What's special about the specialist?, *Journal of Financial Economics* 32, 61-86.

Bessembinder, Hendrik, 1999, Trade execution costs on NASDAQ and the NYSE: A post-reform comparison, *Journal of Financial and Quantitative Analysis*, 34, 387-407.

Bessembinder, Hendrik, 2003a, Quote-based competition and trade execution costs in NYSE-listed stocks, *Journal of Financial Economics*, forthcoming.

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

Bessembinder, Hedrik, and Herbert M. Kaufman, 1997, A cross-exchange comparison of execution costs and information flow for NYSE-listed stocks, *Journal of Financial Economics* 46, 293-319.

Chordia, Tarun, and Avinidhar Subramanyam, 1995, Market making, tick size, and payment-for-order flow: Theory and evidence, *Journal of Business* 68, 543-575.

Demsetz, Harold, 1968, The cost of transacting, *Quarterly Journal of Economics* 82(1), 33-53

Dutta, Prajit, and Ananth Madhavan, 1997, Competition and collusion in dealer markets, *Journal of Finance* 52, 245-276.

Easley, David, N. Kiefer, and Maureen O'Hara, 1996, Cream skimming or profit sharing? The curious role of purchased order flow, *Journal of Finance* 51, 811-834.

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

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

Goldstein, Michael, and Kavajecz, 2000, Eighths, sixteenths and market depth: Changes in tick size and liquidity provision on the NYSE, *Journal of Financial Economics* 56, 125-149.

Griffin, John M., Jeffrey H. Harris, and Selim Topaloglu, 2003, The dynamics of institutional and individual trading, *Journal of Finance*, forthcoming.

Harris, Lawrence and Joel Hasbrouck, 1996, Market vs. limit orders: The SuperDOT evidence of order submission strategy, *Journal of Financial and Quantitative Analysis* 31, 213-231.

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

Hasbrouck, Joel, 1995, One security, many markets: Determining the contributions to price discovery, *Journal of Finance* 50, 1175-1199.

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

Hendershott, Terrence and Charles M. Jones, 2003, Island goes dark: transparency, fragmentation, and liquidity externalities, Columbia University working paper.

Huang, Roger, 2002, The quality of ECN and Nasdaq market maker quotes, *Journal of Finance*, 57, 1285-1319.

Huang, Roger, and Hans Stoll, 1996, Dealer versus auction markets: A paired comparison of execution costs on NASDAQ and the NYSE, *Journal of Financial Economics* 41, 313-358.

Jones, Charles M., and Marc L. Lipson, 2001, Sixteenths: Direct evidence on institutional trading costs, *Journal of Financial Economics* 59, 253-278.

Lipson, Marc L., 2003, Competition among market centers, NYSE working paper.

O'Hara, Maureen, 1997, "Market Microstructure Theory", Blackwell Publishers, Cambridge, Mass.

Peterson, Mark, and Erik Sirri, 2002, Order submission strategy and the curious case of marketable limit orders, *Journal of Financial and Quantitative Analysis* 37.

Appendix A
List of symbols studied.

Symbol	November 2002 Trading Volume	Name
AMD	291,517,400	ADVANCED MICRO DEVICES INC
HI	271,039,900	HOUSEHOLD INTERNATIONAL INC
XOM	224,264,100	EXXON MOBIL CORP
CD	102,219,600	CENDANT CORP
FNM	86,419,500	FEDERAL NATIONAL MORTGAGE ASSN
UNH	85,477,300	UNITEDHEALTH GROUP INC
SWY	80,895,100	SAFEWAY INC
ABT	77,328,900	ABBOTT LABS
WM	69,010,300	WASHINGTON MUTUAL INC
G	63,427,700	GILLETTE CO
ABC	51,175,400	AMERISOURCEBERGEN CORP
TJX	50,004,900	T J X COMPANIES INC NEW
DAL	45,435,800	DELTA AIR LINES INC
SLE	44,582,300	SARA LEE CORP
PRU	44,418,900	PRUDENTIAL FINANCIAL INC
ACS	40,492,300	AFFILIATED COMPUTER SERVICES INC
KFT	39,333,500	KRAFT FOODS INC
CAT	35,640,200	CATERPILLAR INC
OHP	33,217,100	OXFORD HEALTH PLANS INC
COX	32,347,000	COX COMMUNICATIONS INC NEW
Z	29,385,700	FOOT LOCKER INC
CMS	28,927,100	C M S ENERGY CORP
PFG	26,335,800	PRINCIPAL FINANCIAL GROUP INC
ETR	20,167,300	ENTERGY CORP NEW
BRO	15,833,800	BROWN & BROWN INC
CTL	15,831,000	CENTURYTEL INC
ROK	14,982,500	ROCKWELL INTERNATIONAL CORP NEW
SHW	14,758,300	SHERWIN WILLIAMS CO
PTV	14,656,800	PACTIV CORP
TXT	12,728,800	TEXTRON INC
GTK	12,108,600	GTECH HOLDINGS CORP
AW	11,699,400	ALLIED WASTE INDUSTRIES INC
TCB	11,618,500	T C F FINANCIAL CORP
PPD	10,734,000	PRE PAID LEGAL SERVICES INC
DST	9,431,600	D S T SYSTEMS INC DEL
NCF	7,972,800	NATIONAL COMMERCE FINANCIAL CORP
TEX	7,855,700	TEREX CORP NEW
ATI	6,688,700	ALLEGHENY TECHNOLOGIES
ION	6,155,900	IONICS INC
MW	6,136,200	MENS WAREHOUSE INC
PER	4,355,000	PEROT SYSTEMS CORP
HGR	4,280,400	HANGER ORTHOPEDIC GROUP INC
GVA	4,274,900	GRANITE CONSTRUCTION INC
EV	3,999,200	EATON VANCE CORP
GAS	3,899,900	NICOR INC
CXR	3,808,300	COX RADIO INC
NUI	3,496,600	N U I CORP NEW
BTU	3,266,300	PEABODY ENERGY CORP
PNM	3,188,800	P N M RESOURCES INC
GPN	3,017,500	GLOBAL PAYMENTS INC
BBX	2,808,700	BANKATLANTIC BANCORP INC
BKH	2,449,400	BLACK HILLS CORP
CBM	2,433,200	CAMBREX CORP
HAE	2,167,700	HAEMONETICS CORP MASS
BWS	2,155,800	BROWN SHOE INC NEW
KCP	2,095,800	COLE KENNETH PRODUCTIONS INC
KFY	1,636,300	KORN FERRY INTERNATIONAL
MHO	1,449,400	M I SCHOTTENSTEIN HOMES INC NEW
BKI	1,227,000	BUCKEYE TECHNOLOGIES INC
AIT	1,178,300	APPLIED INDUSTRIAL TECHS INC

Table 1
Summary statistics

The sample combines the 20 most active symbols for the month of November 2002 (measured by consolidated trading volume), plus a stratified random sample of 40 additional symbols. All symbols are common equity with a trade-weighted price of at least $5.00 during November. Dash-5 eligible trades represent SuperDot executions of market and marketable limit orders of 9,999 shares or fewer.

	Full Sample	Active 20	Remaining 40
Symbol Characteristics			
Price (dollars)	26.64	35.38	22.27
Shares Outstanding (thousands)	344,074	869,426	81,399
Market Value (thousands of dollars)	12,921,236	34,389,762	2,186,973
Consolidated Daily Volume (shares)	1,757,870	4,420,618	426,496
All NYSE System Trading Activity (daily average shares executed)			
All Orders	1,262,357	3,095,893	345,590
Market Orders	479,732	1,229,213	104,992
Marketable Limit Orders	404,086	959,080	126,590
Retail Orders	38,501	97,831	8,835
Orders from Institutions	691,991	1,720,103	177,935
Program Trades	356,980	842,435	114,252
Other Orders	174,886	435,523	44,568
Dash-5 Eligible NYSE System Trading Activity (daily average shares executed)			
All Orders	599,952	1,466,543	166,657
Market Orders	330,388	849,474	70,845
Marketable Limit Orders	269,565	617,069	95,812
Retail Orders	22,081	54,802	5,721
Orders from Institutions	328,285	812,025	86,414
Program Trades	190,549	451,533	60,056
Other Orders	59,038	148,183	14,466

Table 2
Correlation of signed order flow for one-minute intervals

Signed order flow is the net of buys minus sells (measured using the number of shares or the number of orders) for all market orders and marketable limit orders of less than 10,000 shares, aggregated across all stocks in the sample over one-minute intervals. Inference assumes time-series independence.

	Retail	Institution	Program	Other
Autocorrelation (Shares)	0.1026***	0.3443***	0.3638***	0.2810***
Contemporaneous Correlation (shares)				
Institution	0.0537**			
Program	0.0635*	0.5512***		
Other	0.0371**	0.3516***	0.2935***	
Contemporaneous Correlation (orders)				
Institution	0.0347**			
Program	0.0522*	0.5130***		
Other	0.0388**	0.3043***	0.2665***	
Cross-Autocorrelation (shares)				
Lagged Retail	0.1026***	0.0351**	0.0273***	0.0410***
Lagged Institution	0.0668***	0.3443***	0.2421***	0.2488***
Lagged Program	0.0787***	0.2559***	0.3638***	0.2187***
Lagged Other	0.0621***	0.1736***	0.1123***	0.2810***

*, **, and *** indicate significance at the 10%, 5% and 1% levels, respectively.

Table 3
Transaction cost measures by account type

Standard trading cost measures for the entire sample and selected partitions. Values are in pennies and are share-weighted across all observations. For each account type, we test whether the given value differs from the corresponding value for retail orders. Statistical tests are based on the daily time series of share-weighted averages.

			Spread Decomposition			
		Shares (1,000)	Effective Spread	Realized Spread	Information Component	Quoted Spread
All Orders						
	Retail	26,497	2.60	1.22	1.38	3.04
	Institution	393,941	3.07***	-0.15***	3.22***	3.19***
	Program	228,658	3.05**	0.39***	2.66***	2.78***
	Other	70,846	2.46	0.11***	2.34**	2.93***
By Order Type						
Market Orders	*Retail*	21,908	2.82	1.13	1.69	3.12
	Institution	217,028	4.09***	-0.06***	4.15***	3.66***
	Program	121,339	4.38***	0.95	3.44***	3.31**
	Other	36,190	3.38**	0.11***	3.27***	3.48***
Marketable Limit Orders	*Retail*	4,589	1.53	1.63	-0.10	2.66
	Institution	176,913	1.83**	-0.25***	2.07***	2.62
	Program	107,319	1.55	-0.24***	1.79***	2.18***
	Other	34,656	1.49	0.12**	1.37**	2.35***
By Order Size						
Very Small (100 – 499 shs)	*Retail*	5,927	1.69	1.10	0.59	3.24
	Institution	85,411	2.57***	-0.32***	2.89***	3.36**
	Program	77,997	2.93***	-0.26***	3.20***	3.06***
	Other	12,719	2.38***	0.10***	2.28***	3.36*
Small (500 – 1,999)	*Retail*	10,448	2.39	1.09	1.30	3.09
	Institution	165,176	3.11***	-0.58***	3.69***	3.28**
	Program	100,532	2.80**	0.15***	2.65***	2.61***
	Other	29,760	2.52	-0.11***	2.63***	3.11
Medium (2,000 – 4,999)	*Retail*	6,265	2.99	1.08	1.91	2.87
	Institution	86,251	3.17	0.31*	2.85**	3.08
	Program	38,207	3.71**	1.82	1.90	2.69**
	Other	16,104	2.47	-0.20***	2.68	2.60***
Large (5,000 – 9,999)	*Retail*	3,857	3.92	1.96	1.96	2.88
	Institution	57,104	3.56	0.67**	2.90	2.87
	Program	11,922	3.87	2.11	1.76	2.66
	Other	12,263	2.36***	1.08	1.28	2.48***

*, **, and *** indicate significance at the 10%, 5% and 1% levels, respectively.

Table 4
Differences Between Retail and Institutional Orders

All orders are aggregated (share-weighted average) if they are on the same date in the same stock with the same order size category, same order type, and same account type. Pairs are formed when there are both retail and institutional orders that match along all four other dimensions, and the table reports equal-weighted averages or average differences across these pairs. The reported difference is the retail value minus the institutional value. Statistical tests are based on the time series of daily averages.

			Spread Decomposition			
		Category Pairs	Effective Spread	Realized Spread	Information Component	Quoted Spread
All Orders						
	Retail		2.72	0.96	1.76	3.58
	Difference	4,388	-0.50***	1.57***	-2.06***	0.23***
By Order Type						
Market Orders	*Retail*		3.33	1.18	2.15	3.92
	Difference	2,819	-0.61***	1.66***	-2.27***	0.12*
Marketable Limit Orders	*Retail*		1.63	0.57	1.06	2.97
	Difference	1,569	-0.30**	1.40***	-1.70***	0.41***
By Order Size						
Very Small (100 – 499 shs)	*Retail*		1.95	0.92	1.03	3.71
	Difference	1,619	-0.65***	1.42***	-2.07***	0.29***
Small (500 – 1,999)	*Retail*		3.04	0.79	2.26	3.73
	Difference	1,548	-0.27***	1.72***	-2.00***	0.22**
Medium (2,000 – 4,999)	*Retail*		2.98	0.94	2.05	3.37
	Difference	801	-0.70**	1.11**	-1.81***	0.24**
Large (5,000 – 9,999)	*Retail*		4.02	1.84	2.18	2.95
	Difference	420	-0.37	2.40***	-2.77***	-0.03

*, **, and *** indicate significance at the 10%, 5% and 1% levels, respectively.

Table 5
Analysis of Matched Orders

Standard execution cost measures for matched pairs of orders arriving within five seconds of each other. Matches must be the same order type (market or marketable limit), and order direction (buy or sell). The table reports averages across all matched pairs. Inference is conducted using the time series of daily average paired differences.

		Spread Decomposition			
	Matched Pairs	Effective Spread	Realized Spread	Information Component	Quoted Spread
Retail		3.269	0.103	3.167	3.528
Institution		3.288	0.149	3.139	3.518
Difference	9,705	-0.018	-0.047	0.028	0.010
Retail		3.497	0.435	3.062	3.186
Program		3.681	0.564	3.117	3.167
Difference	4,935	-0.184***	-0.129**	-0.055	0.019
Retail		3.377	0.686	2.691	3.637
Other		3.356	0.697	2.659	3.626
Difference	2,070	0.021	-0.011	0.032	0.012

*, **, and *** indicate significance at the 10%, 5% and 1% levels, respectively.

Table 6
Analysis of Order Timing

This table describes the timing of order flow relative to recent changes in quotes and the magnitude of the quote change. For a marketable buy (sell) order, the change in the relevant side is the last change in the ask (bid) price, and a negative number indicates that the terms of trade are improving. Price changes are in cents.

	Time (in seconds)			Last Quote Change	
	Since Last Quote Change	Since Last Improving Limit Order	Since Last Spread Decrease	Change in Relevant Side	Change in Spread
Retail	66.52	94.46	78.95	-0.33	-0.62
Institution	62.97	93.47	78.84	-0.26	-0.74
Difference	3.55	0.99	0.11	-0.07*	0.12**
Retail	65.38	92.96	78.00	-0.34	-0.63
Program	65.24	93.71	78.68	-0.17	-0.71
Difference	0.14	-0.75	-0.68	-0.17***	0.08
Retail	64.39	91.33	77.88	-0.34	-0.62
Other	56.54	82.55	69.91	-0.18	-0.61
Difference	7.85***	8.78***	7.97*	-0.16***	-0.01

*, **, and *** indicate significance at the 10%, 5% and 1% levels, respectively.

Table 7
Momentum Analysis

Price momentum around execution and duration of order executions. Momentum is the price change in the same stock over the specified interval, signed by the direction of the order. Momentum is measured in cents using quote midpoints and is positive if price is moving up around a buy or down around a sell. Comparisons across account types use the approach described in Table 4, which controls for stock, trading day, order type, and order size category. Statistical tests are based on the time series of daily averages.

	Pre Arrival	**Execution**	**Post Execution**	
	5 Minutes Before Arrival	*Arrival to Execution*	*1 Minute After Execution*	*Next 4 Minutes After Execution*
ALL ORDERS				
Momentum (tests are against null of zero)				
Retail	0.104	0.348***	2.150***	-0.546***
Institution	0.409**	0.632***	3.459***	0.373***
Program	1.839***	0.574***	2.711***	0.079
Other	-0.257	0.508***	2.622***	0.010
Average time from arrival to execution (in seconds)				
Retail		22.49		
Institution		21.46		
Program		18.04		
Other		23.66		
COMPARABLE ORDERS ONLY				
Retail vs Institutional				
Retail Momentum	0.162	0.341	1.811	-0.491
Institution Momentum	0.519	0.605	2.556	0.527
Difference	-0.357	-0.264***	-0.745***	-1.018***
Retail vs Program				
Retail Momentum	0.103	0.311	1.413	-0.444
Program Momentum	1.854	0.699	2.294	0.065
Difference	-1.751***	-0.388***	-0.881***	-0.510***
Retail vs Other				
Retail Momentum	0.262	0.320	1.392	-0.520
Other Momentum	-0.215	0.538	2.027	-0.069
Difference	0.477	-0.218***	-0.635***	-0.451***

Table 8
Volume Analysis

NYSE system order volume around retail and institutional order execution, in shares. Comparisons across account types use the approach described in Table 4, which controls for stock, trading day, order type, and order size category. Statistical tests are based on the time series of daily averages.

	Number of Categories	**Pre Arrival** *1 Minute Before Arrival*	**Execution** *Arrival to Execution*	**Post Execution** *1 Minute After Execution*
Retail vs Institutional				
Volume Retail		3,263	1,314	3,152
Volume Institutional		3,721	1,187	3,614
Difference	4,099	-458***	127	-462***
Net Signed Volume Retail		169	20	94
Net Signed Institutional		702	231	597
Difference	4,099	-532***	-212***	-503***
Retail vs Program				
Volume Retail		3,422	1,387	3,301
Volume Program		3,898	1,271	3,748
Difference	3,670	-476***	117	-447***
Net Signed Volume Retail		160	15	65
Net Signed Volume Program		1,108	371	783
Difference	3,670	-948***	-356***	-718***
Retail vs Other				
Volume Retail		3,537	1,440	3,422
Volume Other		3,932	1,319	3,788
Difference	3,548	-395***	121	-366***
Net Signed Volume Retail		177	23	73
Net Signed Volume Other		699	219	494
Difference	3,548	-522***	-196***	-421***

Figure 1
Trading Volume by Time of Day

Distribution of trading volume, by account type, over the course of the trading day. Chart excludes first and last 15 minutes and each point represents a five minute block.

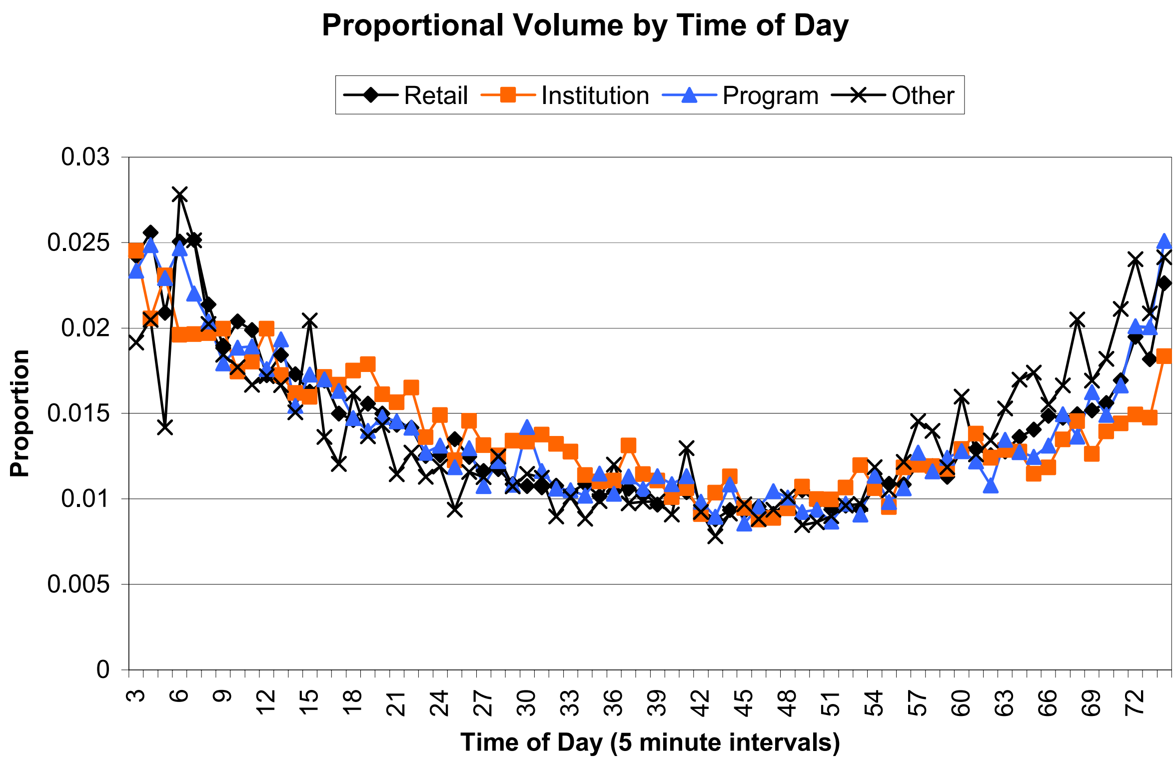

Figure 2
Quoted Spread Around Orders

Share-weighted average quoted spreads in pennies at various times before order arrival (negative numbers) and after order execution (positive numbers) Orders are aggregated and weighted using the approach in Table 4, which controls for stock, trading day, order type, and order size category.

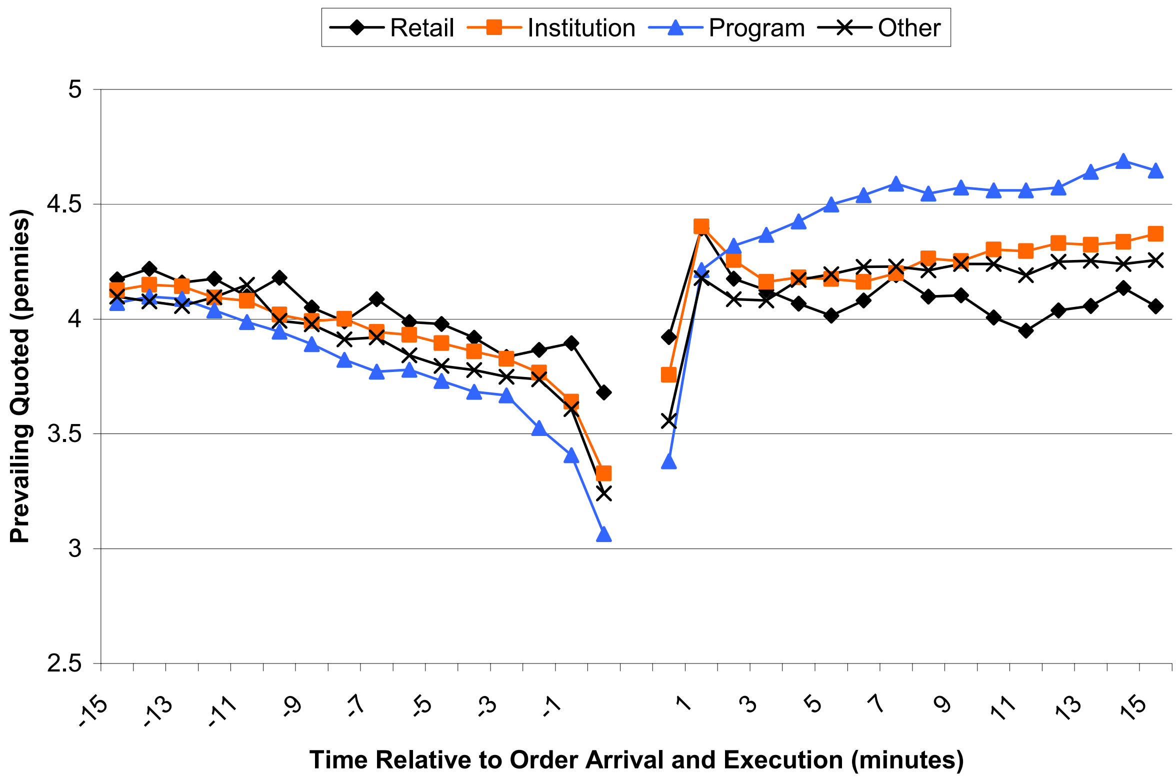

Figure 3
Cumulative Momentum

Cumulative price change over the specified interval, signed by the direction of the order. Price change or momentum is measured using quote midpoints and is positive if price is moving up around a buy or down around a sell. Single points at time zero include the earlier cumulative price changes plus the price change between order arrival and order execution. Orders are aggregated and weighted using the approach in Table 4, which controls for stock, trading day, order type, and order size category.

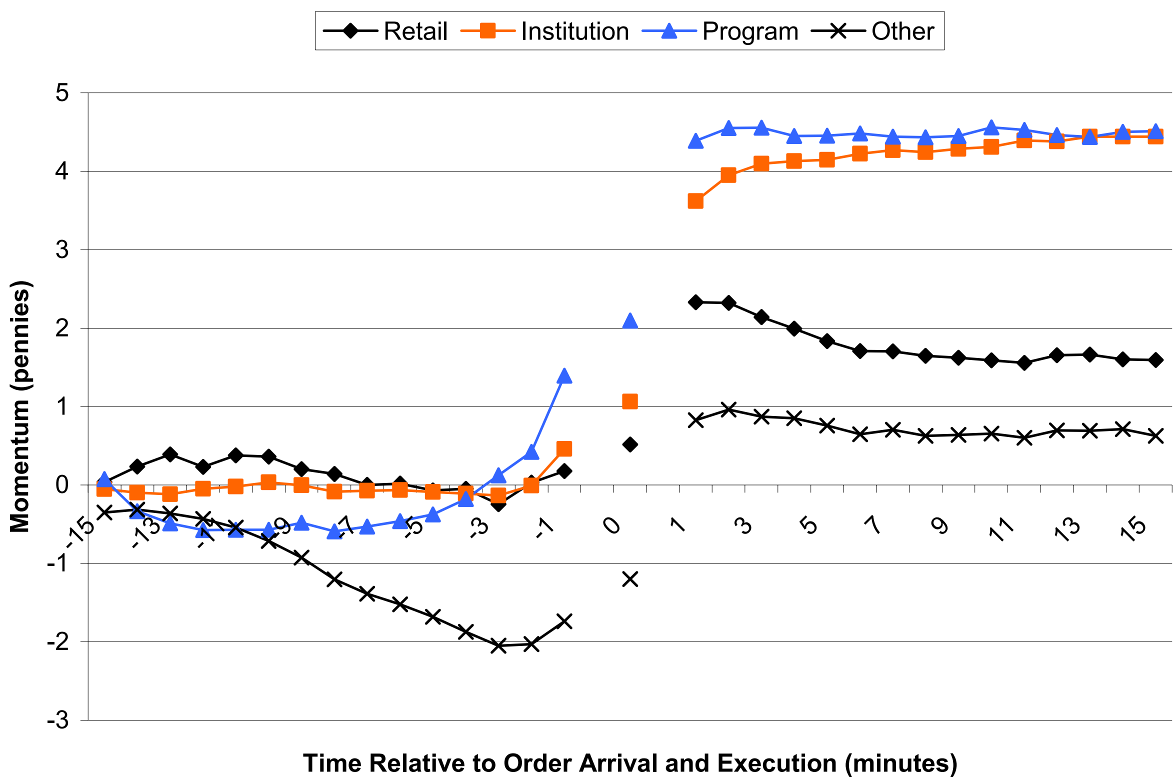

Figure 4a. Unit shock (1,000 shares) to retail net order flow in XOM

Impulse response functions for a vector autoregression in one-minute quote returns and net order flow of various account types. Confidence intervals are constructed by estimating a separate VAR and impulse response function for each trading day and then assuming independence over time. Dashed lines are two standard errors away from the average estimated impulse response and reflect approximate 95% confidence intervals.

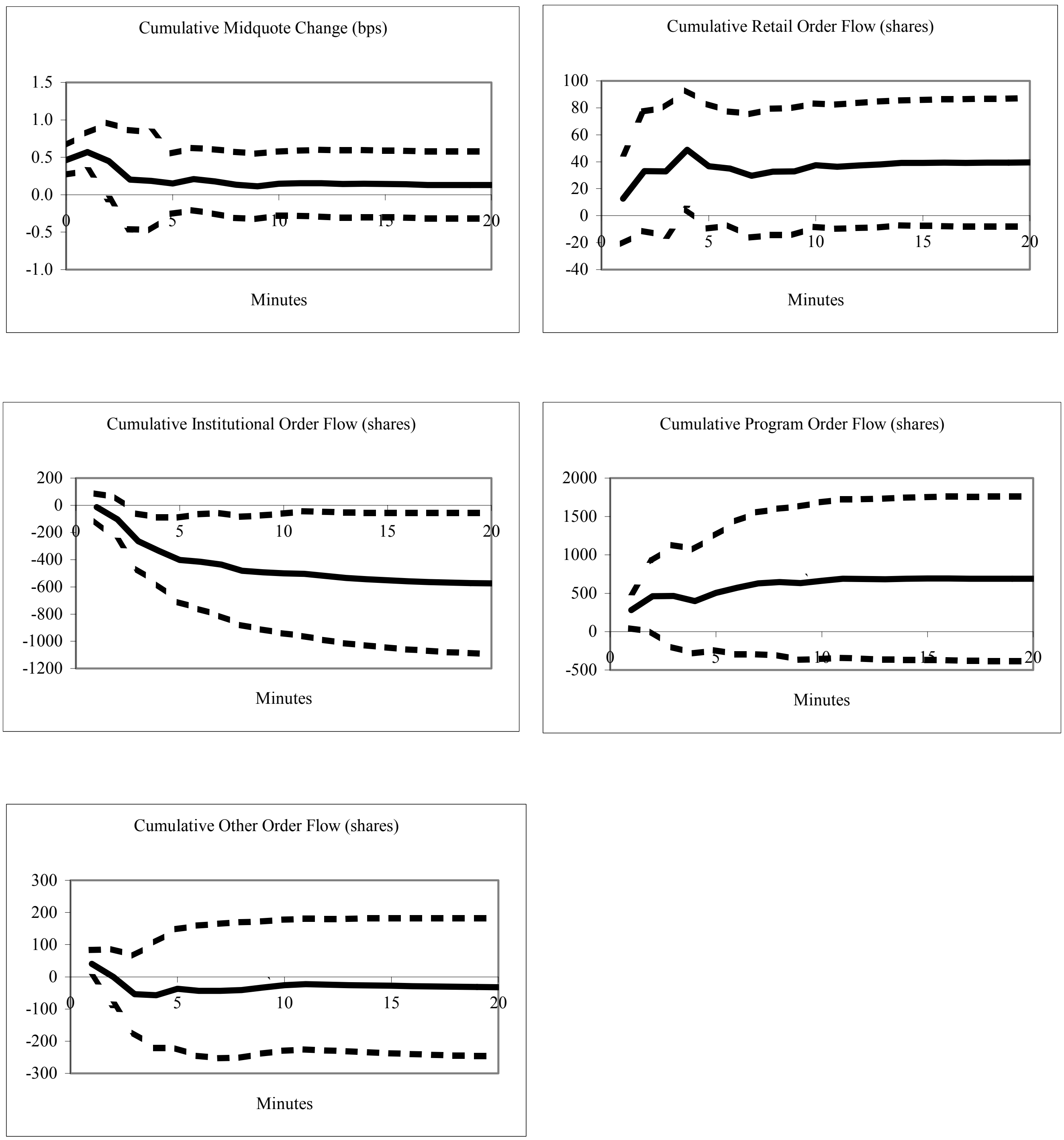

Figure 4b. Unit shock (1,000 shares) to institutional net order flow in XOM

Impulse response functions for a vector autoregression in one-minute quote returns and net order flow of various account types. Confidence intervals are constructed by estimating a separate VAR and impulse response function for each trading day and then assuming independence over time. Dashed lines are two standard errors away from the average estimated impulse response and reflect approximate 95% confidence intervals.

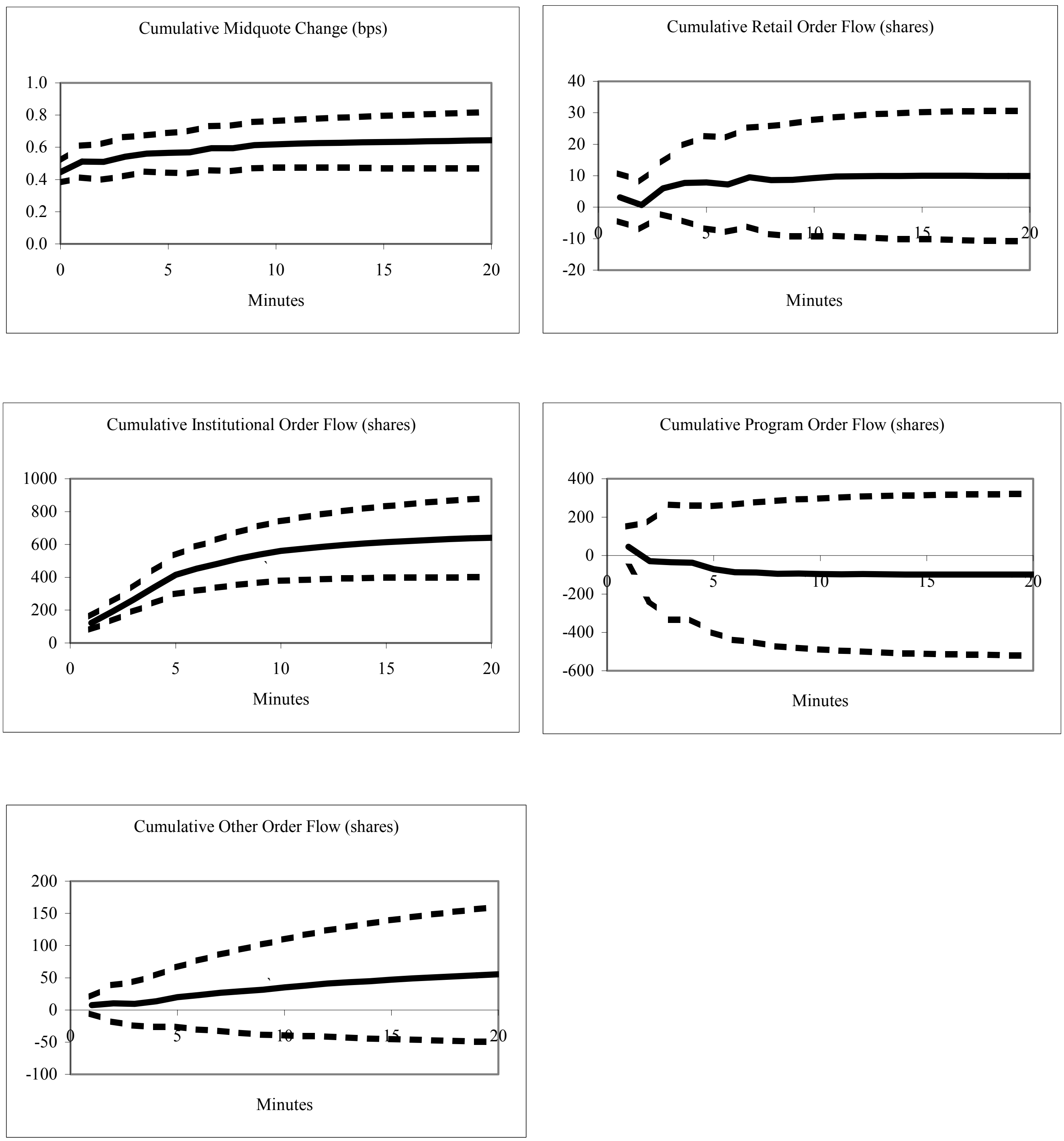

Figure 4c. Unit shock (1,000 shares) to program net order flow in XOM

Impulse response functions for a vector autoregression in one-minute quote returns and net order flow of various account types. Confidence intervals are constructed by estimating a separate VAR and impulse response function for each trading day and then assuming independence over time. Dashed lines are two standard errors away from the average estimated impulse response and reflect approximate 95% confidence intervals.

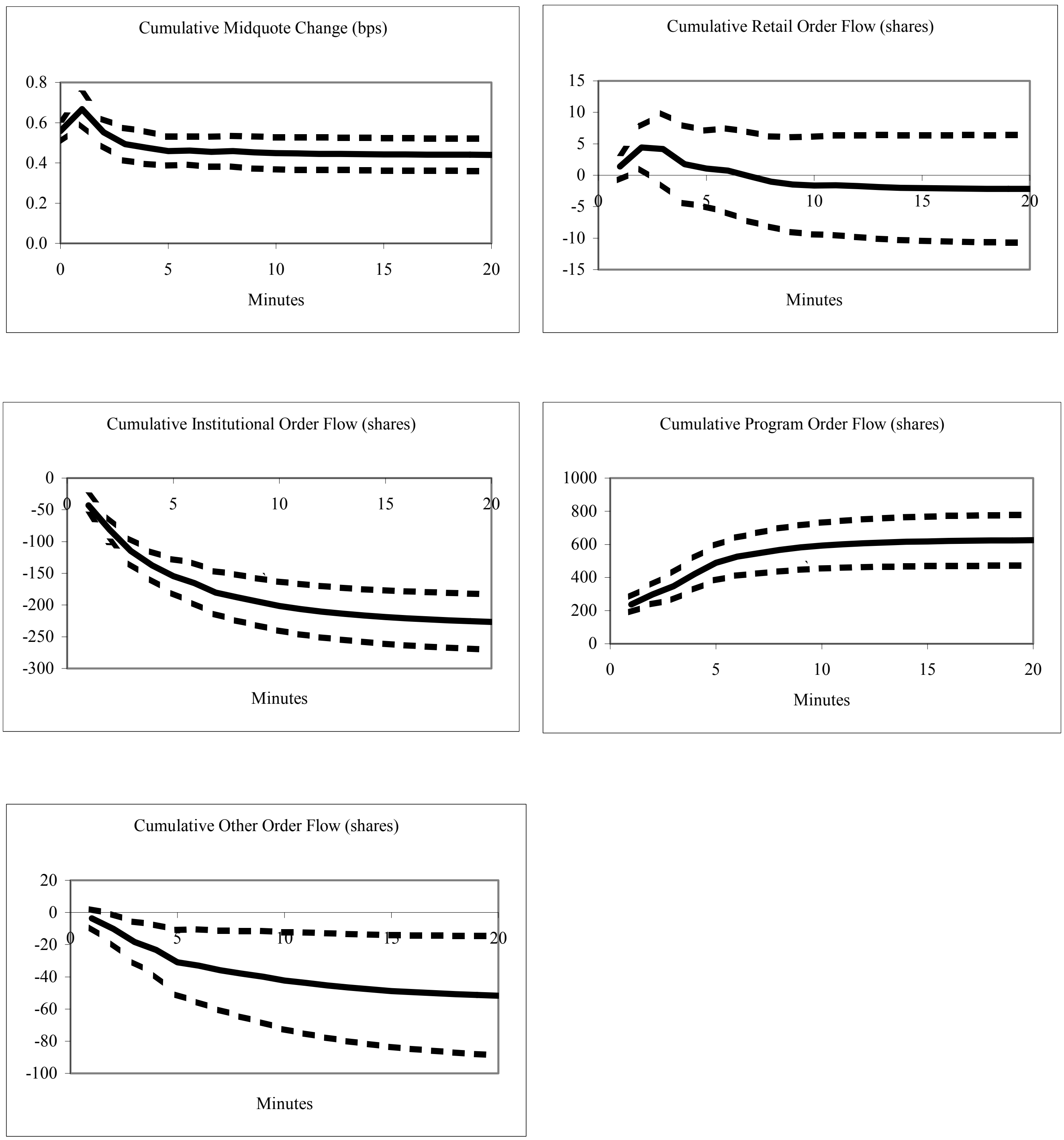

Figure 4d. Unit shock (1,000 shares) to other net order flow in XOM

Impulse response functions for a vector autoregression in one-minute quote returns and net order flow of various account types. Confidence intervals are constructed by estimating a separate VAR and impulse response function for each trading day and then assuming independence over time. Dashed lines are two standard errors away from the average estimated impulse response and reflect approximate 95% confidence intervals.

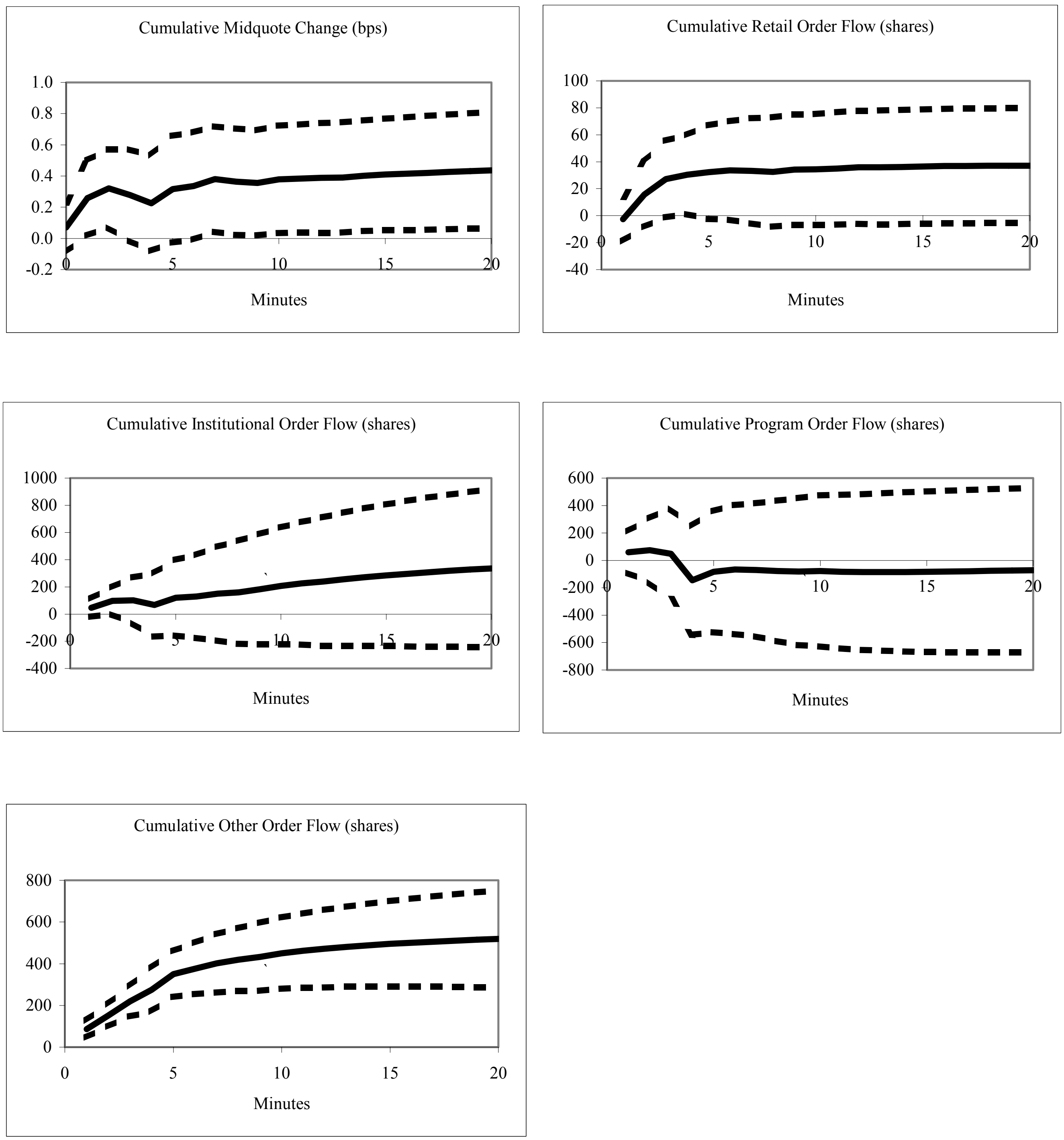

Figure 5a. Unit shock (1,000 shares) to retail net order flow, average of 20 most-active stocks

Results of a vector autoregression in one-minute quote returns and net order flow of various account types. A separate VAR is estimated for each trading day for each stock. The figures report impulse responses averaged across stocks and over time. Confidence intervals assume independence over time. Dashed lines are two standard errors away from the average estimated impulse response and reflect approximate 95% confidence intervals.

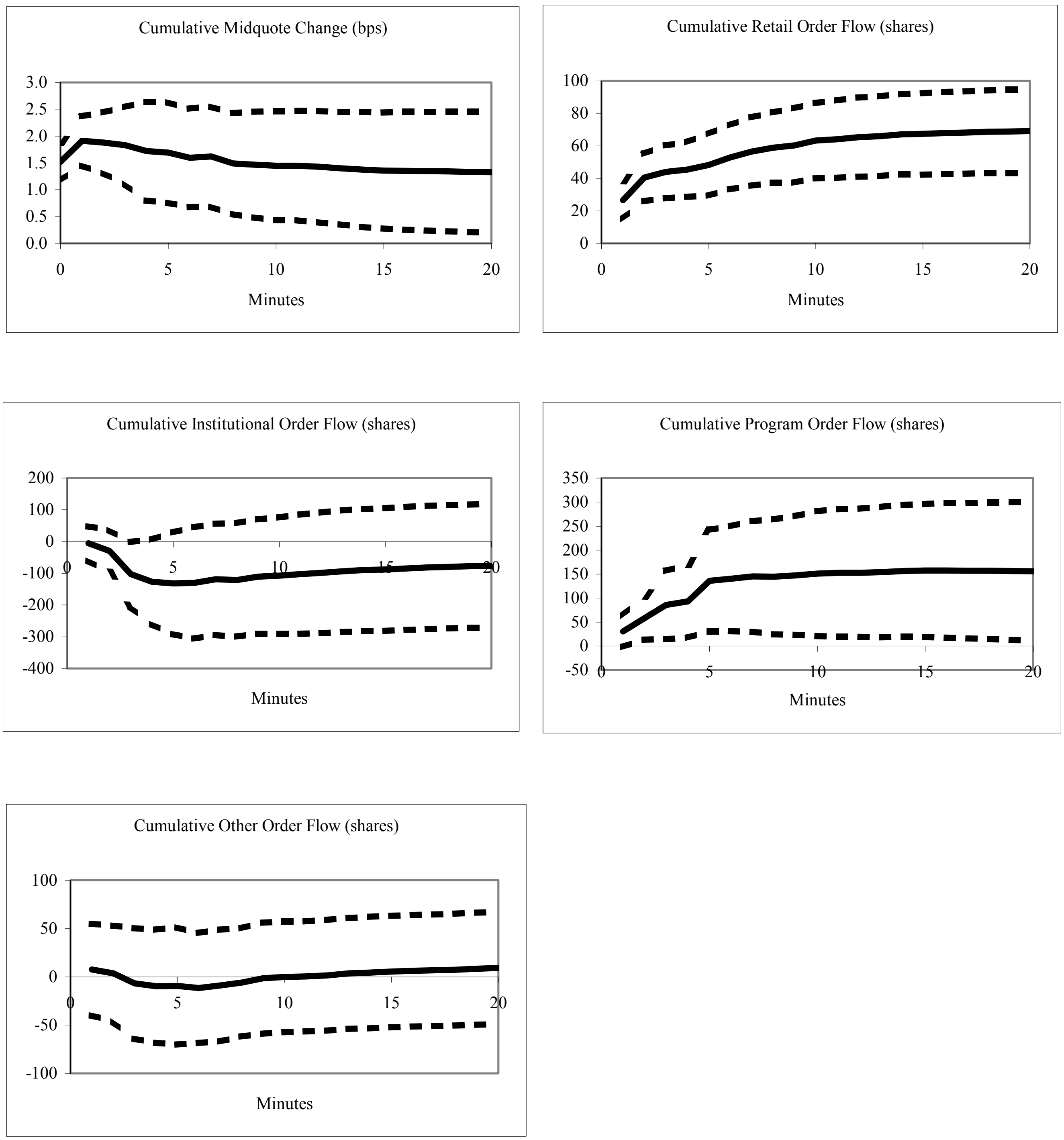

Figure 5b. Unit shock (1,000 shares) to institutional net order flow, avg of 20 most-active stocks

Results of a vector autoregression in one-minute quote returns and net order flow of various account types. A separate VAR is estimated for each trading day for each stock. The figures report impulse responses averaged across stocks and over time. Confidence intervals assume independence over time. Dashed lines are two standard errors away from the average estimated impulse response and reflect approximate 95% confidence intervals.

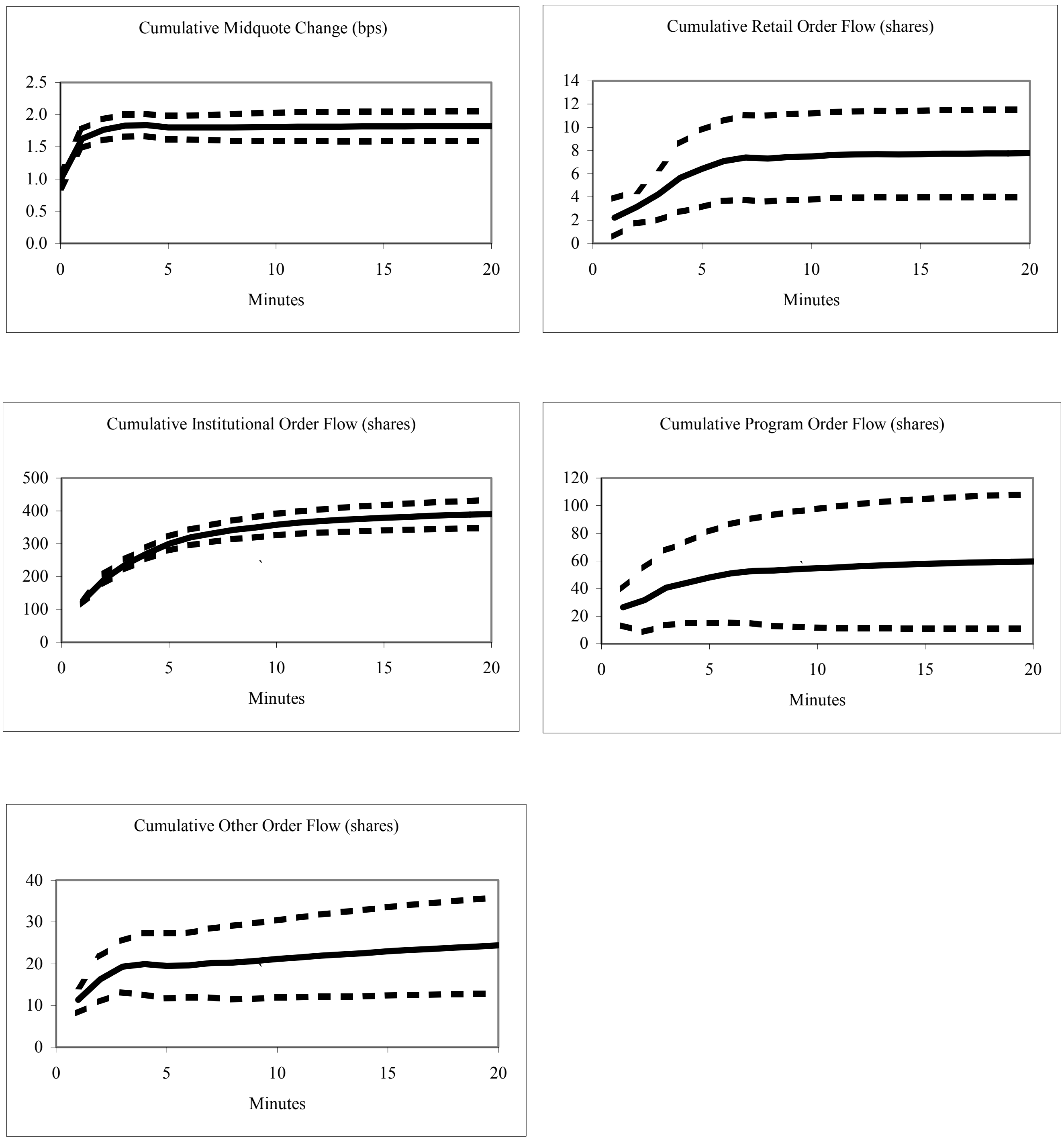

Figure 5c. Unit shock (1,000 shares) to program net order flow, avg. of 20 most-active stocks

Results of a vector autoregression in one-minute quote returns and net order flow of various account types. A separate VAR is estimated for each trading day for each stock. The figures report impulse responses averaged across stocks and over time. Confidence intervals assume independence over time. Dashed lines are two standard errors away from the average estimated impulse response and reflect approximate 95% confidence intervals.

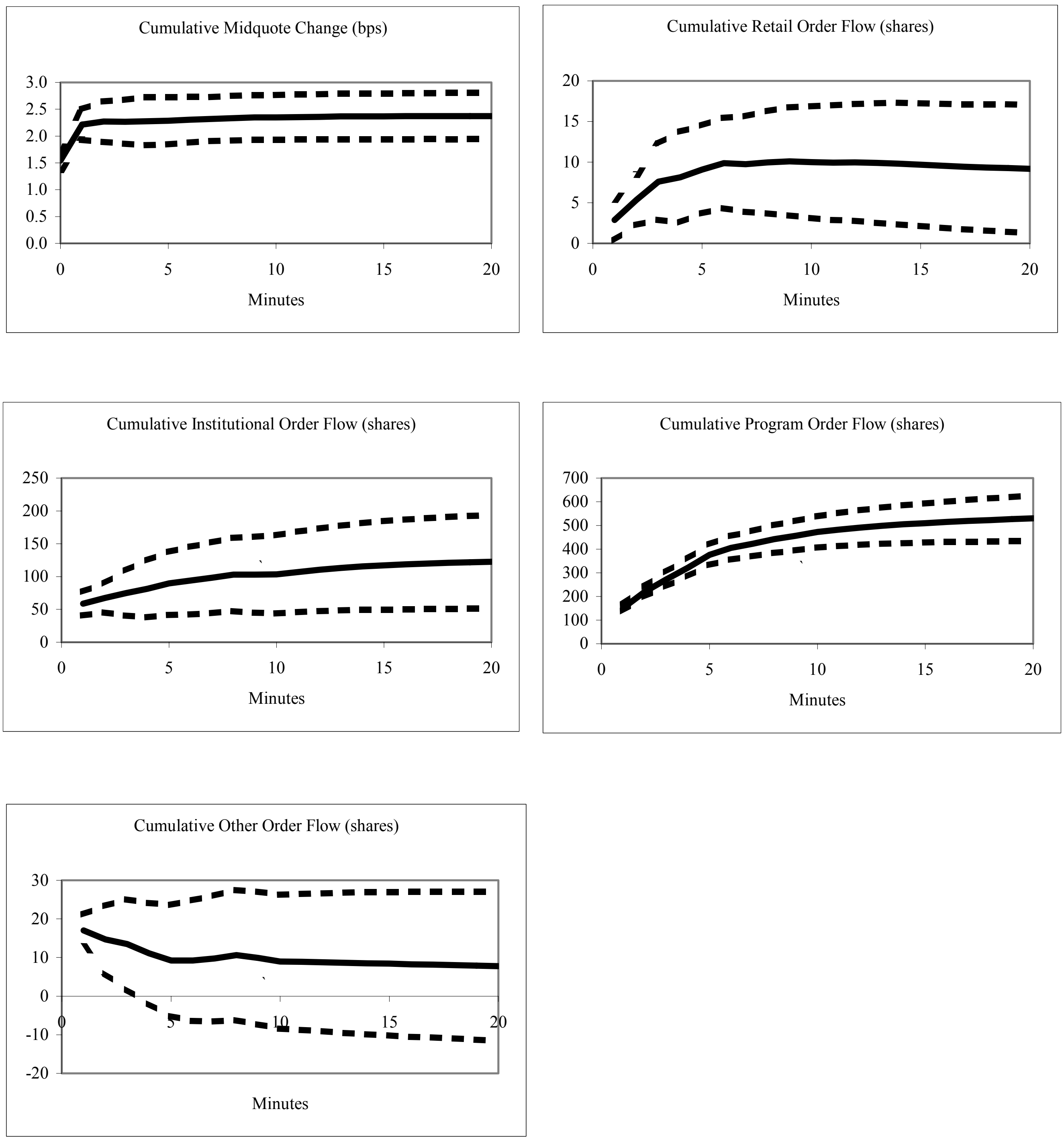

Figure 5d. Unit shock (1,000 shares) to other net order flow, average of 20 most-active stocks

Results of a vector autoregression in one-minute quote returns and net order flow of various account types. A separate VAR is estimated for each trading day for each stock. The figures report impulse responses averaged across stocks and over time. Confidence intervals assume independence over time. Dashed lines are two standard errors away from the average estimated impulse response and reflect approximate 95% confidence intervals.