Information and the intermediary: Are market intermediaries informed traders in electronic markets?

Amber Anand
Whitman School of Management
Syracuse University
Syracuse, NY 13244-2130
Email: amanand@syr.edu
Phone: (315) 443-3397

November 15, 2004

Preliminary. Please do not quote.

I am grateful to Dan Weaver for sharing the data, as well as for his comments. Thanks are also due to Sugato Chakravarty, Terry Martell, Marios Panayides, Ravi Shukla and seminar participants at SUNY, Buffalo for their comments. I thank Joel Hasbrouck for making the programs available for calculating information shares. The standard disclaimer applies.

Abstract

Given the lack of any clear evidence on the informational contributions of market intermediaries vis-à-vis their clients in the extant literature, we analyze whether market intermediaries are informed traders, and whether they trade ahead of their clients to buttress their profits. Using confidential trades and orders data from the Toronto Stock Exchange, we find that intermediaries account for a majority of price discovery, in spite of initiating fewer trades and volume than their clients. Our estimates of price discovery attributable to market intermediaries range between 55% and 62%, although these trades are responsible for only 37% of trades, representing 40% of total volume. We also analyze whether this result is driven by inappropriate handling of customer orders. We explicitly test for frontrunning and stepping ahead by intermediaries, and find no evidence of such behavior.

While the role of market intermediaries in equity markets has received significant attention in the literature, little has been done to explicitly compare the informational differences between market intermediaries and their clients. The issue of relative informational advantages of market intermediaries and outside customers is significant from the standpoint of price discovery. Price discovery primarily occurs through the trading of informed traders who trade based on their private signals about the asset's unobservable true value. Hence, an understanding of which traders are informed and contribute to price discovery assumes considerable significance. The literature on the source of information in equities has focused largely on differences between institutions and individuals. Chakravarty (2001) applies the Barclay and Warner (1993) methodology to the question and finds that institutions are responsible for most of the cumulative price impact in his sample. Other studies use the level of institutional ownership to assess the informational advantages of institutions vis-à-vis individuals. While these studies indicate that institutions are more informed than individuals, they do not shed any light on the informational differences between intermediaries and their clients. The questions that we seek to answer in this study then are: first, are intermediaries more, or less, informed vis-à-vis their clients? And second, do intermediaries inappropriately use their knowledge of client orders in placing their own orders? We answer these questions in context of a completely electronic, and highly transparent market - the Toronto Stock Exchange.

The question of whether intermediaries are informed is prominent in the current debate on the future of the NYSE floor. Views on both sides of the electronic versus floor trading assume that electronic trading would render intermediaries irrelevant. Critics of the NYSE floor have focused on the information available to traders and specialists on the floor of the exchange.[1] Hence, it is argued, that creating an electronic system would even out the playing field. Traders on the floor also seem to accept this in their reluctance to change. In an article on LaBranche Specialists, the WSJ reports, "… investors' ability to buy and sell without intervention could cut into specialists' ability to trade profitably for their own accounts, a

[1] *The Wall Street Journal* (August 2, 2004) quotes Scott DeSano, head of Trading at Fidelity, "…the thing that's so striking is how much information the specialists had versus the rest of the world." ("Big Board Chief's Tough Job: Selling Technology on the Floor")

practice that last year accounted for 66% of LaBranche's revenue."[2] With our study of intermediaries in an electronic market, we test whether the presumption of the irrelevance of intermediaries in a screen based environment is indeed true. We should point out here that our analysis focuses on trading of intermediaries as a group. Thus, specialists and broker-dealers are combined together in our analysis. In the electronic setup of the Toronto Stock Exchange, specialists do not enjoy any informational advantages over other traders and are employed by the same firms that operate as broker-dealers. Further, proprietary trading by member firms is important in different market structures. Broker-dealers provide competition to exchanges by frequently internalizing trades. This source of competition is significant even for a floor driven market like the NYSE.[3] It is also likely that such competition will intensify in an electronic environment. In the futures markets, dual traders are similar to broker-dealers in our setup in that they trade for their own account as well as on behalf of their clients. Hence, our definition of intermediaries encompasses this broader group of entities involved in facilitating trading in markets.

Traditional market microstructure literature views market intermediaries as uninformed traders who cover their losses to informed traders by charging the adverse selection component of the bid-ask spread. More recent models (Saar (2001)) show that market intermediaries possess important order flow information which gives them an informational advantage. Madhavan and Smidt (1993) and Hasbrouck and Sofianos (1993) find that specialists have an investment motive to their trades. Given the conflicting views of the role of market intermediaries – as passive participants in the traditional microstructure literature, and as active traders taking positions in the market to trade profitably on their information in more recent studies - it becomes an empirical issue to examine whether market intermediaries are informed traders who contribute to price discovery. And if they are informed, how do they compare to outside customers? In the Holden and Subrahmanyam (1992) framework, the value of information is short

[2] "NYSE Specialist firms still face aftershocks from trading inquiry," *Wall Street Journal*, March 17, 2004

[3] *The Economist* ("Exchange places," May 3, 2001) comments on the role of investment banks as competition to the NYSE in, "It also says something else about the big investment banks: that they see themselves as the biggest potential rivals to the exchanges. They are, for instance, keen to internalize orders…."

In a letter to the editor of the Wall Street Journal, Thomas E. Haley, Vice President, Market Date at the NYSE claims that 13% of NYSE volume is internalized by NYSE member firms ("Big Board's 'Woes'? We don't see any," *Wall Street Journal*, October 3, 2003.)

lived. Hence, market intermediaries' ability to predict short term price movements might be as, or more, valuable than information held by outside customers.

The informational role of intermediaries has received more attention in futures markets, although evidence is mixed. Fishman and Longstaff (1992) find that dual traders (who trade for clients as well their own accounts) in futures markets are informed, while Chakravarty and Li (2003) find that dual traders act as liquidity suppliers. More recently, Kurov and Lasser (2004) find that trades initiated by locals in futures markets are responsible for a majority of the information share in the S&P 500 and Nasdaq 100 futures and E-mini market. This suggests that, in the futures markets, it is the market intermediaries who contribute most to price discovery. However, the combination of two factors – the advantage that market intermediaries draw from their presence on the floor in these markets, and the lower possibility of information asymmetry in futures contracts (as these tend to be on basket securities, or affected by macro-economic factors which are less prone to private information) – makes it difficult to extrapolate any results regarding informational contribution of intermediaries, from futures to equity markets. Kurov and Lasser (2004), in fact, find that the higher information share associated with locals is due to the locals' knowledge of large orders on the floor – information that they use to place orders in the simultaneous electronic market.

Our analysis uses detailed transactions data, with trader classifications, on the 100 most liquid stocks traded on the Toronto Stock Exchange (TSX), for September 2002, and the information shares methodology developed by Hasbrouck (1995).[4] The TSX is an electronic market, characterized by a high degree of pre-trade transparency, with member firms facilitating trading for their clients (institutional and individual), as well as trading for their own account, and specialists assigned to stocks to maintain a two-sided market. The electronic setup of the exchange precludes informational advantages that market intermediaries enjoy specifically in floor-based markets, from confounding our analysis.

[4] While a longer data period would have been desirable, data constraints limit our sample. However, an analysis of a different sample strongly supports our results.

Our results indicate that trades initiated by market intermediaries (for their proprietary accounts) are responsible for a majority of price discovery in our sample. Specifically, the information share associated with market intermediary initiated trades, ranges between 55% and 62% for our sample stocks. This result is even more striking when we consider that these trades form only 37% of all trades, representing 40% of the shares traded. Further, information shares are higher for intermediary initiated trades across different volume stocks (ranging from lower and upper bounds of 50% and 56% for the lowest volume quartile, to lower and upper bounds of 60% and 68% for the highest volume quartile). Intermediary initiated trades also have a statistically higher information share than client initiated trades, after controlling for the respective proportion of trades and proportion of volume. Analysis of a different sample period strongly supports these results.

This study then presents compelling evidence that market intermediaries tend to be informed in the conventional sense of contributing to price discovery. Further, by establishing this result in a highly transparent electronic market, we show that this result is not driven by a privileged view of the market on a trading floor. Our results indicate that there is a significant informational role for intermediaries, even in a screen-based environment.

Recent allegations of improper trading by specialists on the NYSE, and the earlier investigations of intermediaries in futures markets, raise the possibility that the higher information share of market intermediaries on the TSX is a result of frontrunning, or stepping ahead by brokers. We test for such activities using both trades, and detailed orders data, and find no evidence of such trading by intermediaries on the TSX.

The plan for the rest of the paper is as follows. Section 2 presents the background and motivation of the research. Section 3 describes the institutional features of the Toronto Stock Exchange. Section 4 presents the details of the data used in the paper. Section 5 presents a brief overview of the implementation of the Hasbrouck (1995) information shares methodology in our context. Section 6 discusses the results and provides robustness tests. Section 7 concludes.

2. Background and Motivation:

The role of information, its sources and how it gets incorporated into prices has received considerable attention in the literature. In our summary of related literature we focus on three streams - price discovery, "informed" traders, and market maker behavior.

Price discovery has recently been studied on many dimensions. Hasbrouck (1995) provides an information shares methodology to measure the contribution of NYSE versus other exchanges to price discovery in the Dow 30 stocks. Huang (2002) and Barclay, Hendershott and McCormick (2003) focus on the ECNs on the Nasdaq market and document that ECNs are the preferred destination for informed traders, accounting for a majority of the information share for Nasdaq stocks. Hasbrouck (2003) studies the index futures markets and finds an overwhelmingly large information share associated with the electronically traded E-Mini contracts. Kurov and Lasser (2004) follow up on Hasbrouck (2003) and attribute this informational dominance of E-minis to trades initiated by locals. Chakravarty, Gulen and Mayhew (2004) find that options play a significant role in price discovery for the underlying stocks. Anand and Chakravarty (2004) analyze price discovery across trade sizes in options markets and find that small and medium size trades are responsible for a majority of price discovery. While these studies significantly further our understanding of how and where information gets incorporated into prices, the only study focusing on type of traders responsible for discovering prices is Kurov and Lasser (2004). However, their context is very specific to floor-based futures markets. Kurov and Lasser (2004) find that locals use information about large orders on the floor to trade in the simultaneously trading electronic E-Mini market. The analysis of basket securities (index futures) also mitigates private information effects (Subrahmanyam (1991)), and might increase the advantage that market intermediaries enjoy. Thus, these studies do not answer the general question of which traders' trades result in the discovery of prices. We contribute to this literature by examining the informational contribution of different trader types, specifically market intermediaries versus their clients, in a completely electronic (and highly transparent) environment, and in the context of equities where the potential for informational effects is known to be stronger than in basket securities.

Existing literature, on which trader types are informed in equities, centers on differences between institutional and individual investors. One approach to the issue has been to relate institutional ownership to the level of informed trading in the market (proxied by spreads or various measures of adverse selection). The evidence here appears to indicate a positive relationship between institutional ownership and information asymmetry (Kothare and Laux (1995), Heflin and Shaw (2000) Sarin, Shastri and Shastri (2000) and Dennis and Weston (2001)). Chiang and Venkatesh (1988) find no relationship between institutional ownership and information asymmetry. Chakravarty (2001) finds that institutional trades, in the TORQ data, account for a majority of the cumulative price impact in his sample stocks. Since member firms' proprietary trades are combined with trades of their institutional clients, and specialist trades are not included in the analysis (these trades, along with certain others, are not identified in the TORQ data[5],) his analysis focuses on informativeness of individual and institutional trades, rather than on the role of market intermediaries.[6] Thus, while this stream of literature concludes that institutions are more informed than individuals, it does not analyze the role of market intermediaries in price discovery.

The third stream of literature relevant to our study is that on market maker behavior. Our study focuses on a context where security dealers as well as specialists co-exist in an electronic market. Therefore, we realize that some of the models discussed here will not strictly apply in our context. However, we will limit ourselves to the most general principles that come out of these studies for market intermediaries. Traditional market microstructure models divide the trading population into the informed, the uninformed, and market makers. In this framework, market makers lose from trading to informed traders, and the adverse selection component of the spread compensates them for this risk. However, some studies have also emphasized market maker expertise. Madhavan and Panchapagesan (2000) find that the NYSE specialist uses his knowledge of future order flow to set opening prices that are more informative than would result from a simple matching of buy and sell orders. Saar (2001) presents a

[5] In the Koski and Scruggs (1998) TORQ sample, 29% of buy side and 28% of sell side volume doesn't have any attached identifiers (Table 2, page 62.)

[6] Koski and Scruggs (1998) use TORQ data to analyze trading by securities dealers around ex-dividend days. Their focus, however, is not on information or price discovery.

model where market makers gain useful information about investor endowments through their knowledge of the order flow. Madhavan and Smidt (1993) analyze trades of NYSE specialists on the NYSE and find that specialists possess information about impending order flow.[7] Thus, there is reason to believe that market intermediaries possess important information about the evolution of future order flow and are able to trade as active investors based on this information. Given the short term nature of informational advantages in the Holden and Subrahmanyam (1992) framework, it is then not clear who is more informed – intermediaries or their customers.

3. Institutional details:

The Toronto Stock Exchange (TSX) organizes its trading as a pure electronic limit order book, characterized by a high degree of transparency. Participants in the market include member firms, specialists,[8] and outside investors who trade through TSX member firms. Smith, Turnbull and White (2001) illustrate the routing of client orders on the TSX (Figure 1, page 1727). All retail and institutional order flow is routed through a member firm. A member firm can then execute the order within the firm (with the member firm's own account, or another client's account as the counterparty), or send it to the limit order book. The TSX disseminates, in real time, details on orders away from the best bid and ask prices, including broker identifications for the orders submitted. Subscribers also have the option of viewing data aggregated by price, order or submitting broker.

The Universal Market Integrity Rules (UMIR), established by Market Regulation Services Inc. (the Self Regulatory Organization for Canadian equity markets) govern equity trading in Canada. UMIR provide clear guidelines giving priority to client orders. Member firms trade for their clients as well as for

---

[7] Chae and Wang (2003) find that market makers on the Taiwan Stock Exchange trade as informed traders rather than liquidity suppliers. In the context of a floor based futures market, Manaster and Mann (1999) find that market makers have an informational advantage over off-exchange traders.

[8] The assignment of specialists in an electronic market, although unusual, is increasingly being adopted by other such markets (for example, the Paris Bourse, International Securities Exchange and the Stockholm Stock Exchange. Charitou and Panayides (2004) provide a comprehensive review of specialist systems used in various markets.) The specialist, on the TSX, is assigned by the exchange and is responsible for maintaining two sided markets, moderating price volatility, guaranteeing executions for odd lot orders, and for a specified number of shares, also known as Minimum Guaranteed Fill (MGF) orders. Each stock is assigned to a single specialist.

their own accounts (and for the accounts of their employees) with strict rules prohibiting frontrunning. Specifically, trading for the firm's account is prohibited if the participant has knowledge of a client order "that on entry could reasonably be expected to affect the market price of a security" (UMIR 4.1-1). Once an order is submitted, client limit orders get priority over principal orders at the same price. Similarly, client market orders gain priority over the member firm's own market orders.

4. Data:

We use confidential transactions data, on the 100 most liquid stocks, for a 19 day period - September 3 to 27, 2002 from the Toronto Stock Exchange. The data include, apart from time stamped prices and quantities of each trade, the trader type, and the broker associated with the trade on the buy and sell side. The data used in this study identify each counterparty to a trade as a member of one of four trader types – members' trades representing public customers (client orders), members trading for their firm's account (principal orders) or for the private accounts of the firm's employees, and specialist trades.[9] These identifiers are required at the time of order entry. While some of the member firms of the TSX have other lines of business as well, trades on account of these entities are identified as client orders. For example, an order placed for the portfolio management division of a member firm routed through the brokerage division of the same firm would be classified as a client order.[10] This eliminates problems associated with inappropriately assigning trades to intermediaries. We only include trades that occur during regular trading hours (9.30 a.m. – 4.00 p.m.)[11] We also use quotes data to sign trades as buys or sells. Quotes data include time stamped best bid and ask prices, with the associated quantities, available in the market.

[9] Identifiers also exist for trades by options specialists but such trades are rare and represent 0.33% of all buys and 0.39% of all sells in our data. We ignore these trades in our analysis.

[10] We are grateful to James Twiss, Chief Policy Counsel at Market Regulation Services Inc for clarifying order classifications.

[11] We also filter the data for data errors. Trades with prices or quantities equal to zero are deleted. Trades are also identified as errors if a particular trade is more than four standard deviations away from the average price of the stock for the particular day.

Our analysis focuses on price discovery contributions of market intermediaries versus that of their clients. Hence, in our estimations we use two trader types – clients, and market intermediaries (proprietary and employee trades of brokers, and specialist trades). The Hasbrouck (1995) methodology used in the analysis requires daily estimation of information shares. Accordingly, we include only those days where both trader types trade in the particular stock on the particular day. We also require a minimum of 50 trades in a particular stock in a trading day. Table 1 presents summary statistics for our sample, overall and divided into volume quartiles based on the daily average number of trades. We divide the sample into volume quartiles to analyze whether the trading volume of the stock affects price discovery contributions of market intermediaries. For example, on the NYSE, specialist behavior varies by liquidity of stocks (Cao, Choe and Hatheway (1997) document specialists using active stocks to subsidize trading in inactive stocks. Huang and Liu (2003) find support for cross-subsidization by individual specialists on the NYSE.)

Table 1, Panel A shows that out of a possible 1,900 "stock days" (each trading day for each sample stock that is included in the sample, defining the number of days for which estimations are run) in the sample (19 trading days for 100 stocks), 1,870 days meet our selection criteria. While 25 stocks fall in each of the quartiles, the number of stock days varies across the four quartiles. In the highest volume quartile (volume quartile 1) all of the possible 475 stock days (19 trading days each for 25 stocks) are included. The numbers drop as we go down the quartiles to 454 stock days for the lowest volume quartile. Statistics presented are the averages across the number of stock days. Given our selection criterion of the 100 most liquid stocks, we find that, in our sample, an average stock on an average day has 384 trades. However, there is wide variation across the volume quartiles – 129 trades per day for the lowest, to 892 trades daily for the highest volume quartile. The average prices across the four quartiles are similar - volume quartile 2 has the highest average price at C$32.5, and quartile 3 the lowest at C$23.4.

5. Methodology:

We use the Hasbrouck (1995) information shares technique for estimating the contribution of various trader types to price discovery on the Toronto Stock Exchange. Specifically, we divide trades into two categories: those initiated by market intermediaries, and those initiated by their clients. Trades initiated by member firms' for their own proprietary accounts (or their employees' accounts), and those initiated by specialists are classified as market intermediary initiated trades. We are comfortable in combining specialist and member firm initiated trades, as specialists, on the TSX, have no informational advantages over other traders. Further, these specialists work for the same member firms that act as broker-dealers, ruling out any firm-specific structural differences. We classify a trade as buyer or seller initiated using the Lee-Ready (1991) algorithm (trades above the midpoint of the bid and ask prices are classified as buyer initiated, below the midpoint as seller initiated. Trades at the midpoint are classified using the tick test. If the trade occurs on an uptick it is classified as buyer initiated, and as seller initiated if it occurs on a downtick.)[12] We then use the buyer and seller types associated with each trade to classify the trade as initiated by one of the two trader types. For example, a buyer initiated trade using the Lee-Ready algorithm, with a specialist or a member firm as the buyer in our data, would be classified as "market intermediary initiated" in our analysis. If more than one trade is reported in the same second for a particular trader type (client or intermediary initiated), we only keep the last trade at the time.[13]

Table 1, Panel B presents the proportion of trades and volume associated with each trader type. The proportions are calculated for each trading day for each stock. The numbers reported are the averages of these stock days. Client initiated trades form the majority of the sample on both dimensions. For the overall sample, clients initiate approximately 63% of trades, and 60% of volume. While client initiated trades form a majority in each of the volume quartiles (both by trading frequency and by volume), the proportions are highest for the lowest volume quartile. Approximately 70% of the trades, and 66% of the volume, in the lowest quartile is client initiated. The other 3 quartiles are fairly similar in terms of

[12] In matching trades and quotes, we lag the reported quote time by one second instead of the five seconds suggested by Lee and Ready (1991), since TSX is an automated market. Exchange officials suggested the lag and an examination of the data confirms the adjustment.
[13] This is necessary to implement the Hasbrouck (1995) information shares methodology.

proportion of trades, but the proportion of volume initiated by clients increases monotonically from quartile 1 (highest volume) to quartile 4 (lowest volume). Market intermediary initiated trades simply represent the rest of our sample, with the highest proportion of volume in the most liquid stocks (47% in volume quartile 1) and the lowest in the least liquid stocks (34% in quartile 4).

Our application of the information shares methodology assumes that prices of the trades initiated by the two trader types, for the same stock, share a common random walk component - also referred to as the efficient price. The information share of transactions initiated by a given trader type is then measured as that group's contribution to the total variance of innovations of this efficient price. Hasbrouck (1995, 2002 and 2003) thoroughly detail the methodology, as well as discuss various associated issues. We provide a brief description in the appendix.

We use actual trade prices to construct the price vector (with a time resolution of one second) used in the estimation. Although different from quotes used in Hasbrouck (1995), this approach is similar to the inclusion of last sale trade prices in the price vector in Hasbrouck (2003), transactions prices in Chakravarty, Gulen and Mayhew (2004), and trade prices of regular and E-mini S&P 500 and Nasdaq 100 futures contracts in Kurov and Lasser (2004). Kurov and Lasser (2004) also use the methodology on prices within a market (rather than across markets), in an application very similar to ours. Our estimations also suffer from the issue faced by Kurov and Lasser (2004), in that trader types are not available to the market in real time. As they discuss, the methodology is still appropriate to determine which trader types have a higher information share.

For each stock, we estimate the information share of client initiated, and intermediary initiated trades separately for each trading day (a "stock day" described in section 4). We follow Hasbrouck (1995, 2003) in not including any inter-day price changes. Thus, the analysis yields a set of information share estimates for each stock on each trading day that is included in the sample. We then calculate (and present in the tables) the means of these estimates, for the full sample, or partitioned into volume quartiles (the presentation of averages of estimates from different days follows Hasbrouck (2003)). The number of stock days in the tables provides the number of distinct estimates comprising each aggregation category.

6. Results:

6.1 Information shares estimates

We begin our analysis by estimating information shares associated with different trader types for each stock day in our sample. Table 2 presents the average information share for each trader type. We find that market intermediary initiated trades account for a majority of the information share in the overall sample, as well as across volume quartiles. For the overall sample, the information share of market intermediary initiated trades has a lower bound of 55.15% and an upper bound of 61.67%. The lower and upper bounds of information shares of client initiated trades are 38.33% and 44.85%. The differences between lower and upper bounds of the estimates occur due to the correlation across the innovations in the two price series. Our use of a fine time resolution (one second) in the construction of the price vector keeps the bounds reasonably close, so that the conclusions drawn from analyses of lower and upper bounds are similar. Across volume quartiles, market intermediary initiated trades have a higher information share for the higher volume quartiles (between 60.31% and 68.40% in the highest volume quartile, and between 58.04% and 64.36% in quartile 2) than in lower quartiles (between 52% and 57.8% in quartile 3, and between 49.95% and 55.8% in the lowest volume quartile). The results show that although market intermediaries are responsible for a majority of the price discovery in all volume quartiles, they contribute more in the higher volume stocks. Table 1, Panel B also shows that intermediaries initiate more trades in the more liquid stocks. Thus, the amount of trading and information share appear to be correlated.

6.2 Information shares - controlling for proportion of trades and volume

In this subsection, we follow Barclay and Warner (1993) and Kurov and Lasser (2004) in conditioning information shares on the proportion of trades and volume. In Table 3, we present the information shares for client and market intermediary initiated trades divided by their respective proportions of trades and volume. These ratios are also calculated on a daily basis, i.e. for each stock on each trading day, and then averaged across stock days for presentation. The intuition here is that if market intermediary initiated trades are more informed, then they will contribute a higher information share after controlling for their proportion of trades and volume in the total sample. We find that this is indeed the case. The ratios of information shares to the proportion of trades and volume are markedly lower for client initiated trades (almost half or less) than market intermediary initiated trades. Non-parametric z-tests for the equality of medians reject the hypothesis that the ratios are equal in all cases at the 1% level of significance.[14] We also compare the lower bound of the ratios for intermediary initiated trades to the upper bound of the ratios for client initiated trades. Once again, we find that the ratios are statistically significantly higher for intermediary initiated trades at the 1% level of significance. Thus, we find strong support for our earlier result that market intermediaries contribute more to price discovery than their clients. This suggests that market intermediaries are more informed than their clients. It should be noted here that trades of clients include all retail and institutional trading.

6.3: Do market intermediaries trade inappropriately?

A possible reason for observing a higher information share associated with market intermediary initiated trades is trading ahead or "frontrunning" of client orders by market intermediaries. Frontrunning of client orders involves violating priority of client market orders over member firm's market orders, or the priority of client limit orders over member firm's own limit orders at the same price. While such activities are strictly prohibited on the TSX, it is possible that intermediaries either trade in violation of the rule, or alternatively, step ahead of client orders. Examples of improper trading by market intermediaries include the investigation of intermediaries in Futures markets (in 1988, the FBI found evidence of wrongdoing by brokers on the CBOT and the CME) for trading ahead of client orders, and the more recent allegations of stepping ahead by specialists on the NYSE. Intermediaries step ahead of client orders by improving the price over that of a standing limit order. Such behavior is not prohibited by

[14] We use non-parametric tests to avoid making the normality assumptions required by t-tests. T-tests however, reject the hypothesis of equality at conventional levels of significance as well.

the exchange (making it distinct from frontrunning which is strictly forbidden), and forms the basis of allegations against the NYSE specialist firms. This activity, described as parasitic trading, by Harris (1996) involves buying (selling) ahead of large standing limit orders, thus positioning the parasitic traders to benefit from a potentially favorable price move, while limiting the downside due to the standing order. Parasitic trading might also force the limit order trader to demand immediacy later and buy (sell) from (to) the parasitic trader at a higher (lower) price. In this case, as in the case of frontrunning, we would see initiating trades by intermediaries precede those by their clients.

To analyze this issue, we use the regressions used in Chakravarty and Li (2003) to analyze the relation between client and intermediary initiated trades. These regressions draw on Granger (1980) in testing for causality. Specifically, Chakravarty and Li (2003) regress a market intermediary's (a "dual" trader in their context) customer trades within a time interval on intermediary's proprietary trades, and customer trades in lagged time intervals. The regression equation is of the form:

$$CustomerNetbuy_t = b_0 + b_1.CustomerNetbuy_{t-1} + b_2.CustomerNetbuy_{t-2} + b_3.CustomerNetbuy_{t-3} + b_4.CustomerNetbuy_{t-4} + b_5.CustomerNetbuy_{t-5} + c_1.\ IntermediaryNetbuy_{t-1} + c_2.\ IntermediaryNetbuy_{t-2} + c_3.\ IntermediaryNetbuy_{t-3} + c_4.\ IntermediaryNetbuy_{t-4} + c_5.\ IntermediaryNetbuy_{t-5} + v_t \quad (1)$$

where $CustomerNetbuy_t$ is the total number of shares bought minus the shares sold for all client initiated trades through a specific broker in time interval *t*. Similarly, $IntermediaryNetbuy_t$ is the total number of shares bought minus the total shares sold for initiating trades for the intermediaries' own account in time interval *t*. The rejection of $c_1=c_2=\ldots=c_5=0$ would suggest that market intermediaries' own trades precede those of their customers. We run these regressions separately for each broker for each stock. To allow for enough trades associated with a particular broker, we restrict our analysis to the top brokers (by volume) in each of the 100 stocks in our sample. Specifically, we estimate the regressions for the top broker alone, the top 5 brokers, and the top 10 brokers in each stock. While the objective here is to cover as much of the sample as possible, we also want to ensure that we do not obscure any behavior that might be exhibited by the biggest brokers only, hence the estimations for the top broker and the top 5 brokers as well. Table 4, Panel A gives the proportion of volume associated with the three broker cuts. The proportion of volume associated with a broker is calculated as the total number of shares in trades where the broker is involved (on the buy or the sell side) divided by twice traded volume. The average proportion of volume associated with the top broker over our sample stocks is around 25%. This proportion increases to 65% for the top 5 brokers and 85% for the top 10. Thus, our analysis covers a large proportion of total volume in the sample stocks. The choice of the length of the time interval, *t*, is the next decision. The time interval needs to be sufficiently small to capture any frontrunning behavior that might exist, and not so small as to have very few intervals in which any trades occur. Accordingly, we estimate the regressions using 1, 5 and 10 minute aggregations (Chakravarty and Li (2003) use 5 minute intervals in their estimations). The estimations are conducted using generalized method of moments (GMM), and the Wald Chi-squared test is used to test the hypothesis that $c_1=c_2=\ldots=c_5=0$. Given that there are 100, 500 and 1,000 estimates for the top broker, the top 5 brokers and the top 10 brokers respectively (top 1, 5 and 10 brokers in each of the 100 stocks), we only present the proportion of cases where the null hypothesis is rejected suggesting that intermediaries' own trades precede those of their clients. Table 4, Panel B shows that the null hypothesis is rejected in 7%, or less, of the cases (at the 10% level of significance) with 1 minute intervals, in 11%, or less, of the cases with 5 minute intervals and in only 16%, or less, of the cases even with 10 minute intervals.[15] Given that our tests are based on the assumption of short lived information (Holden and Subrahmanyam (1992)), the results do not support frontrunning, or stepping ahead, of clients' trades by market intermediaries, thus ruling out such improper trading as the source of intermediaries' informational advantage.[16]

[15] We also ran the regressions with an aggregation interval of 30 seconds. The rejection rate drops to about 5% for all the broker cuts.

[16] Chakravarty and Li (2003) also test for "piggybacking" by intermediaries in futures markets. Such follow on trading would not give us the higher information share that we see associated with intermediaries' trades in our analysis, and hence is not our focus here. However, we run those regressions and find no evidence of piggybacking by intermediaries on the TSX. Results are available from the author.

6.4 Robustness

6.4.1 Do intermediaries trade inappropriately – tests using order data

As a check of robustness of the results presented above, we test for frontrunning and stepping ahead, using actual orders instead of trades, again separately for each broker. The use of orders data gives us a much better handle on when the orders were submitted instead of when the trade occurred. This is especially important for the test of stepping ahead by intermediaries. If some client orders never execute due to intermediaries entering their own orders at a better price, an analysis of orders will still be able to capture this effect. Our order data includes every order submitted with the submission time and date, the price, the number of shares in the order, a buy or sell indicator, the trader type of the order submitter and an identifier for the broker that the order is routed though. The use of orders is however, complicated by the sheer volume of cancellations in electronic markets. For example, Hasbrouck and Saar (2002) document the frequent and almost immediate cancellations on the Island ECN. For our data as well, we find that 82% of all submitted orders are subsequently cancelled. Further, in our sample, 50% of the cancelled orders are cancelled in 29 seconds, 75% in a little over 2 minutes (125 seconds), and 90% in less than 9 minutes. Given our time aggregation intervals, it is then important how we treat these cancellations. If an order is submitted and cancelled within a few seconds, does that convey any information to the intermediary? Hasbrouck and Saar (2002) describe these orders as "fleeting" and emphasize that these are not passive orders but more likely to be a part of a dynamic trading strategy that seeks liquidity by continuously placing and cancelling orders until the order is filled. If intermediaries know that a cancelled order represents unfulfilled demand, then that might be valuable information for the intermediary. On the other hand, cancelled orders could provide information on changing market dynamics. To deal with this issue, we run our analysis twice – first, by simply ignoring cancellations and aggregating all orders placed by clients and intermediaries within 1, 5 and 10 minute intervals,[17] and second, by excluding any order that is cancelled within the same aggregation interval. That is, if we are

[17] We only include orders submitted within the time interval in the aggregation. Outstanding orders are not expected to contain any new information for the intermediaries to act upon.

using 5 minute aggregations, and in the 10:00:00 to 10:05:00 interval, an order is submitted after 10:00:00 and cancelled before 10:05:00, then that order is deleted from the data and not used in the estimation.

Table 5, Panel A presents the results for the two estimations of equation 1.[18] In the context of orders, equation 1 tests whether intermediaries submit their orders based on the knowledge of *impending* client orders. We find no evidence that such is the case. The results for 1 minute aggregations indicate that the null of no trading based on knowledge of impending orders, is rejected in only 12% of the cases for the top broker, and in less than 10% of the cases for the top 5 and top 10 brokers, without taking cancellations into account. The results are similar when we delete the orders cancelled within the relevant time aggregation interval. With 5 and 10 minute aggregations the null is rejected in no more than 15% of the estimations. The level of significance used in these tests is again 10%.

Table 5, Panel B tests for frontrunning and stepping ahead by intermediaries. This test is similar to the test of "piggybacking" in Chakravarty and Li (2003), but with the use of order data, it tests if intermediaries submit their orders after client orders have been submitted. If intermediaries step ahead of clients' orders, by improving the price offered (or bid) by a client order (parasitic trading described above), then we would see intermediary orders follow client orders. Similarly, frontrunning will show up as intermediary orders following client orders. Here, the regressions take the form:

$$IntermediaryNetbuy_t = b_0 + b_1.CustomerNetbuy_{t-1} + b_2.CustomerNetbuy_{t-2} + b_3.CustomerNetbuy_{t-3} + b_4.CustomerNetbuy_{t-4} + b_5.CustomerNetbuy_{t-5} + c_1.\ IntermediaryNetbuy_{t-1} + c_2.\ IntermediaryNetbuy_{t-2} + c_3.\ IntermediaryNetbuy_{t-3} + c_4.\ IntermediaryNetbuy_{t-4} + c_5.\ IntermediaryNetbuy_{t-5} + v_t \quad (2)$$

In the estimation with orders, $CustomerNetbuy_t$ is the total number of shares in buy orders minus the total number of shares in sell orders for all client orders through a specific broker in time interval *t*. Similarly, $IntermediaryNetbuy_t$ is the total number of shares in buy orders minus the total number of

[18] For the estimation of equation 1 with orders, *Netbuy* is defined as the difference between the number of shares in orders to buy and sell, for a specific trader type, associated with a particular broker.

shares in sell orders for the intermediaries' own account in time interval *t*. The rejection of $b_1=b_2=\ldots=b_5=0$ in equation 2 would provide evidence of stepping ahead, or frontrunning, by intermediaries. As can be seen in Table 5, Panel B, there is little evidence of intermediaries' orders following client orders. The null of no stepping ahead (and no frontrunning) is rejected in 13% of the cases for 1 minute aggregations for the top broker, in 9.2% of the cases for the top 5 brokers, and in 11.1% of the cases for the top 10 brokers, when cancelled orders are ignored. Deleting orders cancelled within the relevant aggregation interval yields substantially similar results (the numbers in this estimation, for 1 minute aggregations are 13%, 9.4% and 10% for the top broker, the top 5 brokers and the top 10 brokers, respectively). The results for 5 and 10 minute intervals also indicate that intermediaries' information share is not simply a manifestation of frontrunning, or of parasitic trading.

Thus, our analysis of order data confirms the results from our analysis of trades in Section 6.3 that intermediaries on the TSX do not trade improperly by either frontrunning, or stepping ahead, of client orders.

6.4.2: Information shares-different sample period:

To analyze whether our information share results are sample specific, we estimate information shares of client and intermediary initiated trades for 19 days in the month of March 2002 (March 1, 2002 – March 27, 2002).[19] We apply the same sample selection criteria to this period as well. We analyze the 100 most liquid stocks, and for a stock day to be included, require trades by each trader type during the day, and at least 50 trades during the day for the stock. This leaves us with 1,850 stock days (out of a possible 1,900), with 474 in the most liquid quartile to 450 in the least liquid. Table 6, Panel A presents the summary characteristics. This sample is similar to our September 2002 sample, with around 430 trades for an average stock day, ranging from approximately 929 in the most liquid quartile to 174 in the least liquid. Table 6, Panel B also shows that, similar to our earlier sample, most of the trades and volume

[19] The March trades data is similar to the September data in its detail, with the exception that it doesn't contain identifiers for submitting brokers. Fortunately, this does not affect our information shares analysis.

tends to be initiated by clients. Clients initiate more than 70% of trades and 63% of volume in the overall sample. In a trend we saw in Table 1, client initiated trades and volume are higher for the lower volume quartiles.

Table 6, Panel C presents the information share contributions of trades initiated by the two trader types. For the overall sample, the lower and upper bounds, for the information share of client initiated trades, are 36.55% and 42.86%, and that of market intermediary initiated trades are 57.14% and 63.45%. While the information share of intermediary initiated trades is lower for the lower volume quartiles, it is still higher than that of client initiated trades. Thus, in the lowest volume quartile, market intermediary initiated trades make up only 29% of volume but contribute between 55% and 60% of the information share. Although we do not present the ratios of information shares to the proportion of trades and volume in this section, the ratios are significantly higher (at the 1% level of significance using the non-parametric linear rank sum z-statistic) for intermediary initiated trades compared to client initiated trades. We thus find support for our earlier result that intermediaries provide a majority of the price discovery on the TSX.

6.4.3 Internalized trades:

Smith, Turnbull and White (2001) find that intermediaries effectively screen out informed trades in the upstairs market. Thus, trades in the upstairs market tend to be uninformed, and are internalized by the member firm, with the firm crossing the trade – either as principal, or with another client's order (page 1727, Smith, Turnbull and White (2001)). For our analysis this raises the possibility that the information share of our sample of client initiated trades is biased downward due to the mixing up of these clearly uninformed trades with other client initiated trades. Although that doesn't change the conclusion that intermediaries are more informed in aggregate, since they contribute most to price discovery, it could affect the results for the ratios of information shares to proportion of trades and volume. To analyze this issue, we divide all trades for the September sample into three categories- client initiated, non-internalized trades; intermediary initiated, non-internalized trades; and all internalized trades (client or intermediary initiated). A trade is classified as internalized if the same member firm is on the buy as well as the sell side. We find that for the overall sample, internalized trades represent 13.87% of all trades but 39.46% of the total volume.[20] As expected, internalized trades contribute less to price discovery than other categories. Specifically, the information share for internalized trades ranges between 9.67% and 14.72%, whereas intermediary initiated trades have lower and upper bounds of 46.94% and 54.65%, and client initiated trades of 33.13% and 41.34%. We also find that the ratios of information shares to the proportion of trades as well as volume are still significantly higher (at the 1% level of significance using the z-statistic described earlier) for intermediary initiated trades than client initiated trades. Thus, our results are not affected by trading in the upstairs market.[21]

7. Conclusion:

Given the lack of any clear evidence on the informational contributions of market intermediaries vis-à-vis their clients, we analyze whether market intermediaries are informed traders, and whether they trade ahead of their clients to buttress their profits.

Using confidential trades and orders data from an electronic equity market, we find that intermediaries contribute more to price discovery than their clients. Specifically, our results show that intermediaries' share of price discovery ranges between 55% and 62%, while their share of initiated volume is around 40%. Adjusted for the proportion of trades and volume, intermediaries' information share is significantly higher than that of their clients. This indicates a dominant informational role for intermediaries on the Toronto Stock Exchange – an electronic equities market. Further, we find that this result is not driven by illegal or improper trading behavior, such as frontrunning or stepping ahead of client orders.

Our results also have implications for the debate between proponents of the floor and the screen based systems. Both sides seem to believe that such a transition would make intermediaries irrelevant

---

[20] In the Smith, Turnbull and White (2001) sample, upstairs trades account for 55% of the volume. However, they note that these trades are more common in illiquid securities. Thus, our sample selection criterion of picking the 100 most liquid firms on the TSX might account for the difference.

[21] Results for the volume quartiles are similar, and are available from the author.

from an informational viewpoint. We clearly show that this is not the case. Intermediaries contribute more to price discovery, and hence tend to be more informed, even in a transparent electronic market.

References:

Anand, A. and Chakravarty, S., 2004, Stealth trading in options markets, Working paper, Syracuse University

Barclay, M., Hendershott, T. and McCormick, T., 2003, Competition Among Trading Venues: Information and Trading on Electronic Communications Networks, *Journal of Finance*, 58, 2637-2666

Barclay, M., and Warner, J., 1993, Stealth and volatility: which trades move prices? *Journal of Financial Economics*, 34, 281-306.

Cao, C., Choe, H., and Hatheway, F., 1997, Does the specialist matter? Differential execution costs and inter-security subsidization on the New York Stock Exchange, *Journal of Finance*, 52, 1615-1640

Chae, J., and Wang, A., 2003, Who Makes Markets? Do Dealers Provide or Take Liquidity? Working paper, State University of New York, Buffalo

Chakravarty, S., 2001. Stealth Trading: Which traders' trades move stock prices? *Journal of Financial Economics*, 61, 289-307

Chakravarty, S., Gulen, H. and Mayhew, S., 2004. Informed trading in stock and options markets, *Journal of Finance*, 59, 1235-1257

Chakravarty, S., and Li, K., 2003, An examination of own account trading by dual traders in futures markets, *Journal of Financial Economics*, 69, 375-397

Charitou, A., and Panayides, M., 2004, The role of the market maker in international capital markets: challenges and benefits of implementation in emerging markets, Working paper, University of Utah

Chiang, R., and Venkatesh, P., 1988. Insider holdings and perceptions of information asymmetry: A note. *Journal of Finance*, 43, 1041-1048.

Dennis, P.J., and Weston, J., 2001, Who's Informed? An Analysis of Stock Ownership and Informed Trading, Working paper, University of Virginia

Fishman, M. J., and Longstaff, F.A., 1992, Dual trading in futures markets, *Journal of Finance*, 47, 643-671.

Harris, L., 1996, Does a Large Minimum Price Variation Encourage Order Display? Working paper, Marshall School of Business at USC, October 1996.

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

Hasbrouck, J., 2002. Stalking the efficient price in empirical microstructure specifications, *Journal of Financial Markets*, 5, 329-339

Hasbrouck, J., 2003. Intraday price formation in U.S. equity index markets, *Journal of Finance*, 58, 2375-2399.

Hasbrouck, J., and Sofianos, G., 1993, The trades of market makers: An empirical analysis of NYSE specialists, *Journal of Finance*, 48, 1565-1593

Hasbrouck, J., and Saar, G., 2002, Limit Orders and Volatility in a Hybrid Market: The Island ECN, working paper, New York University

Heflin, F. and Shaw, K., 2000, Blockholder ownership and market liquidity, *Journal of Financial and Quantitative Analysis*, 35, 621-633

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

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

Huang, R., and Liu, J., 2003, Do Individual NYSE Specialists Subsidize Illiquid Stocks? Working paper, University of Notre Dame

Kothare, M., and Laux, P., 1995, Trading costs and the trading systems for NASDAQ stocks, *Financial Analysts Journal*, 42-53.

Kurov, A., and D. Lasser, 2004, Price dynamics in the regular and E-Mini futures markets, *Journal of Financial and Quantitative Analysis*, 39, 365-384

Lee, C., and Ready, M., 1991, Inferring Trade Direction from Intraday Data, *Journal of Finance*, 41, 733-746

Madhavan, A. and Panchapagesan V., 2000, Price discovery in auction markets: a look inside the black box, *Review of Financial Studies*, 13, 627-658

Madhavan, A. and Smidt, S., 1993, An analysis of changes in specialist inventories and quotations, *Journal of Finance*, 48, 1595-1628

Saar, G., 2001, Investor Uncertainty and Order Flow Information, Working paper, New York University.

Sarin, A., Shastri, K., and Shastri, K., Nov.2000, Ownership structure and stock market liquidity, Working paper, University of Pittsburgh.

Smith, B., Turnbull, A., and White, R., 2001, Upstairs market for principal and agency trades: analysis of adverse information and price effects, *Journal of Finance*, 56, 1723-1746

Subrahmanyam, A., 1991, A theory of trading in stock index futures. *Review of Financial Studies,* 4, 17–51.

Table 1 : Summary Statistics

This table summarizes the sample used in the study. We use trades in the 100 most liquid stocks on the TSX between September 3, 2002 and September 27, 2002. The numbers are first calculated for each stock each day and then averaged for presentation here. Statistics are also presented separately for volume quartiles. A stock is assigned to a volume quartile based on the average daily number of trades in the stock. Panel A presents characteristics of stocks included in the sample, and Panel B presents the proportion of trades and volume initiated by the two trader types.

Panel A: Characteristics of stocks in the sample

Quartiles	Number of underlying stocks	Number of stock days estimated	Average daily number of trades	Average Price	Average trade size	Average daily share volume
Overall	100	1870	384.2	27.0	1,594.4	728,761.9
(Highest) 1	25	475	892.1	27.9	1,896.9	1,940,420.0
2	25	475	304.1	32.5	1,132.9	353,627.0
3	25	466	196.4	23.4	1,801.6	380,588.2
(Lowest) 4	25	454	129.4	23.8	1,548.1	210,921.0

Panel B: Proportion of trades and volume by trader type

			Client initiated		Market intermediary initiated	
Quartiles	**Number of underlying stocks**	**Number of stock days estimated**	**Proportion of trades**	**Proportion of volume**	**Proportion of trades**	**Proportion of volume**
Overall	100	1870	62.56%	59.81%	37.44%	40.19%
(Highest) 1	25	475	60.09%	53.36%	39.91%	46.64%
2	25	475	59.66%	58.52%	40.34%	41.48%
3	25	466	61.18%	61.84%	38.82%	38.16%
(Lowest) 4	25	454	69.60%	65.81%	30.40%	34.19%

Table 2: Information shares by trader type

This table summarizes the results of Hasbrouck (1995) information share estimations. We use trades in the 100 most liquid stocks on the TSX between September 3, 2002 and September 27, 2002. Information shares are calculated for each stock on each day included in the sample. These estimates are then averaged across stock days and presented here. Results are also presented separately for volume quartiles. A stock is assigned to a volume quartile based on the average daily number of trades in the stock.

Quartiles	Number of stocks	Number of stock days estimated	Client initiated		Market Intermediary initiated	
			Min	Max	Min	Max
Overall	100	1870	38.33%	44.85%	55.15%	61.67%
(Highest) 1	25	475	31.60%	39.69%	60.31%	68.40%
2	25	475	35.64%	41.96%	58.04%	64.36%
3	25	466	42.20%	48.00%	52.00%	57.80%
(Lowest) 4	25	454	44.20%	50.05%	49.95%	55.80%

Table 3: Information shares-controlling for proportion of trades and volume

This table summarizes the results of Hasbrouck (1995) information share estimates divided by the proportion of trades and volume. We use trades in the 100 most liquid stocks on the TSX between September 3, 2002 and September 27, 2002. The ratios are calculated for each stock on each day included in the sample. These estimates are then averaged across stock days and presented here. Results are also presented separately for volume quartiles. A stock is assigned to a volume quartile based on the average daily number of trades in the stock.

			Information share/Proportion of trades				Information share/Proportion of volume			
Quartiles	Number of stocks	Number of stock days estimated	Client initiated		Market Intermediary initiated		Client initiated		Market Intermediary initiated	
			Min	Max	Min	Max	Min	Max	Min	Max
Overall	100	1870	0.62	0.74	1.79*,[a]	1.98*	0.72	0.85	1.94*,[a]	2.14*
(Highest) 1	25	475	0.53	0.68	1.60*,[a]	1.82*	0.62	0.78	1.42*,[a]	1.61*
2	25	475	0.61	0.72	1.59*,[a]	1.76*	0.67	0.79	1.78*,[a]	1.96*
3	25	466	0.70	0.80	1.88*,[a]	2.05*	0.82	0.95	2.06*,[a]	2.25*
(Lowest) 4	25	454	0.66	0.75	2.09*,[a]	2.31*	0.77	0.88	2.52*,[a]	2.75*

*linear rank sum z-statistic for the two sample test of difference of medians of the ratios for client and intermediary initiated trades significant at the 1% level.

[a] linear rank sum z-statistic for the two sample test of difference of medians of the upper bound of the ratios for client initiated trades and lower bound of the ratios for intermediary initiated trades significant at the 1% level.

Table 4: Test of intermediary trades preceding client trades

The table presents results of tests of the hypothesis of frontrunning or stepping ahead of client orders by intermediaries. Panel A presents the proportion of sample represented in the tests for top broker alone, the top 5 brokers and the top 10 brokers. Panel B presents the proportion of cases where the null of $c_1=c_2=\ldots=c_5=0$ is rejected in equation 1, at the 10% level of significance. A rejection implies frontrunning of client orders by intermediaries. Equation 1 is estimated using trades data for the 100 most liquid stocks on the TSX between September 3, 2002 and September 27, 2002. The regressions are estimated separately for each broker, using GMM. The null hypothesis is tested using the Wald Chi-square test statistic. Results are presented for the top broker alone, for the top 5 brokers and the top 10 brokers.

Panel A: Proportion of initiated volume in which the top broker, the top 5, and top 10 brokers are involved

	Proportion of Volume		
	Mean	**Min**	**Max**
Top Broker	25.0%	10.1%	67.2%
Top 5 Brokers	64.6%	44.4%	95.5%
Top 10 Brokers	85.1%	68.6%	98.1%

Panel B: Proportion of cases where the null of no front-running is rejected

	Aggregation interval (in minutes)	**Proportion of cases rejected**
Top Broker	1	7.0%
Top Broker	5	10.0%
Top Broker	10	16.0%
Top 5 Brokers	1	6.2%
Top 5 Brokers	5	11.0%
Top 5 Brokers	10	15.2%
Top 10 Brokers	1	6.5%
Top 10 Brokers	5	11.1%
Top 10 Brokers	10	15.8%

Table 5: Test of frontrunning and stepping ahead using order data

The table presents the results of tests for stepping ahead or frontrunning by intermediaries using order data for the 100 most liquid stocks on the TSX between September 3, 2002 and September 27, 2002. Panel A presents the proportion of cases where the null of $c_1=c_2=\ldots=c_5=0$ is rejected in equation 1, at the 10% level of significance. A rejection implies that intermediaries place their orders based on knowledge of impending client orders. Panel B presents proportion of cases where the null of $b_1=b_2=\ldots=b_5=0$ is rejected in equation 2, at the 10% level of significance. A rejection implies frontrunning or stepping ahead of client orders by intermediaries. Results are presented separately depending on the treatment of cancellations. Results in the first column are drawn from tests where we ignore cancellations and include all orders submitted within the appropriate time interval. Results in the second column are of tests where orders placed and cancelled within a time aggregation interval (1, 5 or 10 minutes) are deleted from the analysis. The regressions are estimated separately for each broker, using GMM. The null hypothesis is tested using the Wald Chi-square test statistic. Results are presented for the top broker alone, for the top 5 brokers and the top 10 brokers.

Panel A: Test of intermediary order submission preceding client order submission

		Proportion of cases rejected	
	Aggregation interval (in minutes)	All orders	All order except cancelled orders
Top Broker	1	12.0%	11.0%
Top Broker	5	9.0%	13.0%
Top Broker	10	15.0%	12.0%
Top 5 Brokers	1	9.8%	9.0%
Top 5 Brokers	5	12.4%	10.8%
Top 5 Brokers	10	12.0%	12.4%
Top 10 Brokers	1	8.7%	8.3%
Top 10 Brokers	5	11.0%	10.5%
Top 10 Brokers	10	11.6%	10.7%

Panel B: Test of intermediary order submission following client order submission

		Proportion of cases rejected	
	Aggregation interval (in minutes)	All orders	All orders except cancelled orders
Top Broker	1	13.0%	13.0%
Top Broker	5	11.0%	14.0%
Top Broker	10	17.0%	17.0%
Top 5 Brokers	1	9.2%	9.4%
Top 5 Brokers	5	9.4%	12.6%
Top 5 Brokers	10	11.0%	11.2%
Top 10 Brokers	1	11.1%	10.0%
Top 10 Brokers	5	10.3%	12.0%
Top 10 Brokers	10	11.2%	10.7%

Table 6: Information shares – March 2002

This table summarizes the March 1, 2002 and March 27, 2002 sample, and the results of Hasbrouck (1995) information share estimations for this sample. We use trades in the 100 most liquid stocks between March 1, 2002 and March 27, 2002. Panel A summarizes stock characteristics. Panel B presents the proportion of trading initiated by clients and intermediaries, and Panel C summarizes the information share estimates. Information shares are calculated for each stock on each day included in the sample. These estimates are then averaged across stock days and presented here. Results are also presented separately for volume quartiles. A stock is assigned to a volume quartile based on the average daily number of trades in the stock.

Panel A: Characteristics of stocks in the sample

Quartiles	Number of underlying stocks	Number of stock days estimated	Average number of trades	Average Price	Average trade size	Average daily share volume
Overall	100	1850	431.8	30.2	1588.2	649,162.8
(Highest) 1	25	474	928.9	33.7	1317.5	1,270,743.7
2	25	467	369.6	37.2	1548.3	533,216.6
3	25	459	234.8	26.9	2183.8	524,724.8
(Lowest) 4	25	450	173.9	22.8	1307.2	241,684.0

Panel B: Proportion of trades and volume by trader type

			Client initiated		Market intermediary initiated	
Quartiles	Number of underlying stocks	Number of stock days estimated	Proportion of trades	Proportion of volume	Proportion of trades	Proportion of volume
Overall	100	1850	70.23%	63.55%	29.77%	36.45%
(Highest) 1	25	474	66.62%	56.25%	33.38%	43.75%
2	25	467	68.51%	62.04%	31.49%	37.96%
3	25	459	71.03%	65.33%	28.97%	34.67%
(Lowest) 4	25	450	74.98%	71.00%	25.02%	29.00%

Panel C: Information shares by trader type

Quartiles	Number of stocks	Number of stock days estimated	Client initiated		Market Intermediary initiated	
			Min	Max	Min	Max
Overall	100	1850	36.55%	42.86%	57.14%	63.45%
(Highest) 1	25	474	33.28%	41.20%	58.80%	66.72%
2	25	467	35.28%	40.88%	59.12%	64.72%
3	25	459	37.95%	44.16%	55.84%	62.05%
(Lowest) 4	25	450	39.91%	45.35%	54.65%	60.09%

Appendix

The Hasbrouck (1995) information shares methodology:

We use the Hasbrouck (1995) methodology to examine the information content inherent in trades of various market participants. Specifically, we consider two kinds of market participants: market intermediaries, and their clients. This approach assumes that prices of trades initiated by the two trader types for the same stock share a common random walk component - commonly referred to as the efficient price. The information share of transactions for a trader type is then measured as that group's contribution to the total variance of the random walk component.

Specifically, if we denote a price vector *p* that includes p_{ct}, the transaction price of a trade initiated by a client, and p_{it}, the transaction price of a trade initiated by an intermediary– for the same stock during a trading session - then we can express p_t as

$$p_t = \begin{bmatrix} p_{ct} \\ p_{it} \end{bmatrix} = \begin{bmatrix} V_t + e_{c,t} \\ V_t + e_{i,t} \end{bmatrix} \tag{A.1}$$

The common efficient price V_t is assumed to follow a random walk given by

$$V_t = V_{t-1} + u_t \tag{A.2}$$

where $E(u_t) = 0$, $E(u_t^2) = \sigma_u^2$, and $E(u_t u_s) = 0$ for $t \neq s$. Then, by the Granger Representation Theorem (see Engle and Granger (1987)), these cointegrated prices can be expressed as a vector error correction model of order *N* as:

$$\Delta p_t = A_1 \Delta p_{t-1} + ... + A_N \Delta p_{t-n} + \gamma(z_{t-1} - \mu_z) + \varepsilon_t \tag{A.3}$$

where p_t is a 2x1 vector of prices; A_i is a 2x2 matrix of autoregressive coefficients corresponding to lag *i*; $\gamma(z_t - \mu_z)$ is the error correction term, and $\mu_z = E(z_t)$. The covariance matrix of the error term above can be expressed as

$$Cov(\varepsilon_t) = E\varepsilon_t \varepsilon_t' = \Omega \tag{A.4}$$

The vector moving average representation of the model given by equation (1) is:

$$\Delta p_t = B_0\varepsilon_t + B_1\varepsilon_{t-1} + B_2\varepsilon_{t-2} + ..., \, where \; B_0 = I \qquad \text{(A.5)}$$

where ε is a 2x1 vector of zero-mean innovations with variance-covariance matrix given by Ω which is of order 2x2. Also notice that in (A.5), the identity matrix *I* is 2x2, and the sum of all moving average coefficients $B(1) = I + B_1 + B_2 + ...$ has identical rows *B*. And since *B* reflects the impact of innovations on the permanent price component rather than transitory components, the total variance of the implicit efficient price changes can be expressed as $\sigma_T^2 = B\Omega B^{'}$. Following Hasbrouck (1995), the contribution to price discovery by each trader type is expressed as that trader type's contribution to this total innovation variance.

Following Hasbrouck (1995), the information share of an intermediary initiated trade (say) is defined as:

$$IS_i = \frac{B_i^2 \sigma_i^2}{\sigma_T^2}$$

where B_i represents the element corresponding to intermediary initiated trades in the *B* vector.

However, if price innovations across client and intermediary initiated trades are correlated, then IS_i is not uniquely defined as above. Instead we estimate the upper and lower bounds, given by $\overline{IS_i}$ and $\underline{IS_i}$, by performing Cholesky factorizations of all possible rotations (i.e., permutations) of the disturbance term, ε. The information shares for client initiated trades, are estimated similarly.

The estimation procedure involves building a second by second series of prices for client and intermediary initiated trades for each stock every trading day. We follow Hasbrouck in not including any inter-day price changes. Using this price vector, we estimate the VECM model in equation A.3 using a lag of 5 minutes (300 seconds). We set the number of periods in the VMA representation (to compute the impulse response functions) as 1,800 seconds. Hasbrouck suggests setting this number sufficiently high to allow forecasted prices to converge. We find that 30 minutes is sufficient for our data.