Splitting and Shuffling: Institutional Trading Motives and Order Submissions Across Brokers *

Munhee Han

October 4, 2019

*Han (munhee.han@utdallas.edu) is at the University of Texas at Dallas, 800 West Campbell Road, Richardson, TX 75080. I wish to thank my adviser, Vikram Nanda, for his helpful comments, patience, and invaluable advice throughout all stages of Ph.D. dissertation research, and Kelsey Wei, Steven Xiao, Alejandro Rivera, Hugh Kim, Qinghai Wang, Michael Rebello, Harold Zhang, Jean-Marie Meier, and seminar participants at the University of Texas at Dallas for helpful comments. All errors are my own.

Splitting and Shuffling: Institutional Trading Motives and Order Submissions Across Brokers

Abstract

This paper studies order submission strategies by institutional investors when trading on private information. By merging institutional daily transactions with original/confidential 13F filings, I separate informed trades from uninformed ones. Informed large orders tend to be split across more brokers and over more days. While same brokers tend to work uninformed large orders over multiple days, the brokers who facilitated early parts of broken-up informed orders rarely receive the remaining parts of the same orders on later days. Institutional investors also provide camouflage for their informed orders by mixing an informed order with other uninformed orders simultaneously sent to the same broker. As a result, a higher degree of shuffling a portfolio of orders is associated with a larger share of informed trading volume. The splitting and shuffling strategies designed to conceal informed trades from brokers and other market participants tend to lower institutional trading costs, especially on informed orders.

Keywords: Institutional trading, informed trades, brokers, order submissions, trading costs

1 Introduction

In this paper I study the order submission strategies of institutional investors when they trade on private information. Motivating the analysis is a substantial literature that models the strategies that informed traders can use to conceal their trading motives and moderate the price impact of their trades. Informed traders can, for instance, engage in dynamic strategies, optimally splitting orders over time to better hide their trades among those of non-information motivated traders (e.g., Kyle (1985), Easley and O'Hara (1987)). Similarly, informed traders can seek opportunities and venues in which their trades can be better concealed (e.g., Admati and Pfleiderer (1988)). Since the presence of informed trading also raises the costs faced by uninformed traders, the uninformed traders have the incentive to certify (i.e., engage in sunshine trading) that their trades are not information-driven (see Admanti and Pfleiderer (1991)). Seppi (1990) suggests that uninformed traders can be screened and face lower costs trading blocks in the non-anonymous upstairs market; whereas informed traders would split up blocks into a series of smaller orders and trade downstairs anonymously. Consistent with the thrust of these models, this paper finds that informed institutional traders follow submission strategies intended to obscure their trading motives.

Institutions trading on information face the supplementary risk that their trading will be recognized and mimicked by other traders. Several recent empirical papers have raised concerns about the private information being detected by other market participants such as high-frequency traders (HFTs) or being leaked through brokers to other investors. Korajczyk and Murphy (2019) find that high-frequency traders submit more same-direction orders during institutional trade executions. Using Swedish equity data, Van Kervel and Menkveld (2019) present evidence that HFTs supply liquidity at the beginning, but eventually trade in the same direction as the institutions submitting the original orders. These papers argue that such back-running is costly to institutional investors trading on private information. Di Maggio et al. (2019) highlight the brokers' role in facilitating back-running by showing that brokers can extrapolate large informed trades from order flows and selectively leak this information to their important clients. Recently, Yang and Zhu (Forthcoming) propose that informed traders can randomize their order flows to prevent other traders from "back-running" on their fundamental information.

There are several key findings in the paper regarding the trading strategies of informed institutional investors. I find that institutional investors tend to spread out their orders across more brokers and over more days when they are trading on information. Institutional investors appear to randomize among brokers when submitting information-driven orders. Institutional investors not only shuffle their order flows across brokers, but also appear to submit their informed orders "camouflaged" among several other uninformed orders when submitting information-driven orders to a broker. Furthermore, this paper provides evidence that splitting and shuffling strategies, apparently designed to hide information from brokers and other investors, lead to lower trading costs as measured by implementation shortfall, especially on informed orders.

My empirical analysis requires separating informed trades from uninformed ones to analyze the order submission strategies of institutional investors trading on private information. Following Agarwal et al. (2013), I first identify "confidential" holdings by merging and comparing original 13F filings with the amendments to the original filings. 13F investors can request confidential treatment for certain holdings, which can be omitted in the original 13F filings and reported later in the amendment filings. Agarwal et al. (2013) show that confidential holdings of institutional investors (especially hedge funds) are information-motivated and tend to outperform the other holdings.[1] Next, I match the managers from the ANcerno institutional trading dataset with 13F managers based on the overlap between quarterly trades inferred from both datasets. I manually verify the potential matches based on manager names. Then, I merge ANcerno daily transactions with original/confidential 13F holdings reports. I identify any buy trades as informed (uninformed) if the stocks bought during a quarter can be matched with confidential (original) holdings for that quarter. This process allows for informed trades to be separated from uninformed ones. I only consider buys because it is not clear how 13F managers would try to "hide" informed sells by delaying reporting their positions in the original 13F filings.

[1]Section 13(f) of the Securities Exchange Act of 1934 requires institutional investors to disclose their quarterly portfolio holdings to public. Form 13F filings reports holdings information as of each calendar quarter-end and Form 13F must be filed within forty-five days of the report date. However, investment managers may request the confidential treatment for certain holdings and omitting those holdings from Form 13F filings would be allowed until SEC makes a decision on the request.

In order to examine how institutional investors break up large orders and spread them across brokers and over time, I stitch together all child orders that are part of the same parent orders following Anand et al. (2012). Specifically, I stitch together all orders on the same stock on the same side of the market (buy or sell) by the same manager on behalf of the same client (simply referred to as "fund" hereafter) that are executed through multiple brokers and over multiple consecutive trading days to construct parent orders. A parent order is a collection of child orders (tickets) on the same stock in the same direction that a fund may place with multiple brokers and over multiple days.

First, I find that investment managers tend to split their informed orders so that any given broker may not know the full size of the orders. Specifically, I find that informed trades are more spread out across more brokers over more trading days. This finding is more pronounced when I restrict my analysis to large orders (blocks), defined as parent orders with share volume greater than or equal to 10,000 shares or dollar volume greater than or equal to $200,000. Informed trades are not only spread out across a larger number of brokers, but they are also spread out more evenly across brokers, as measured by the Herfindahl-Hirschman Index (HHI) of dollar volume executed by each broker.[2] A concern is that my results could be driven by order size: informed orders could be larger and larger orders may simply require more brokers and take longer time to fill. However, my results are robust to controlling for order size.[3]

In addition to splitting and spreading out, I find that investment managers also camouflage their informed orders by mixing an informed order with other uninformed orders that are sent to the same broker on the same day. Interestingly, the brokers receiving informed orders not only receive orders on a larger number of stocks, they also receive a more evenly distributed volume of orders across stocks – so that informed orders are not distinguishable from uninformed ones. These results indicate that investment managers attempt to create their own noisy orders so as to hide their trading motives.

These order submission strategies appear to be designed to conceal informed trades at the individual

[2]HHI is calculated by squaring the portion of the traded aggregate dollar volumes through each broker for the fund and then summing the resulting numbers.

[3]I use three measures of the size of a parent order: the logarithm of the dollar volume of the parent order, the share volume of the parent order scaled by the daily trading volume reported in Center for Research in Security Prices (CRSP), and the share volume of the parent order scaled by the number of shares outstanding reported in CRSP.

order level – and lead to a high degree of randomization of order flows sent across brokers. To show this, I aggregate for each fund all buy orders executed on the same day to construct a portfolio of orders on a daily basis. Then, I measure the extent to which a set of brokers used by an investment manager tend to overlap with or deviate from the set of brokers that the investment manager has used in the past. I ask, for instance, if manager m submits her largest orders to broker b today, would broker b also receive the largest order flows from manager m tomorrow? Specifically, I measure the similarity between the set of brokers used by an investment manager today and the set of brokers used by the same manager one day later, two days later, up to twenty days later. Similarity is measured by cosine similarity based on the dollar volume of orders executed by each broker. I find that investment managers tend to submit their orders to a slightly different set of brokers each day. Looking across days that are further apart, the cosine similarity of the brokers used by each manager drops initially, but settles to a steady level in just a few days. This suggests that while shuffling order flows sent across brokers to hide informed orders, investment managers typically maintain close trading relationships with their core sets of brokers, that can facilitate liquidity provision when managers submit large liquidity-motivated orders (Han, Kim, and Nanda (2019)).

Next, I investigate why investment managers shuffle their orders across brokers each day. One reason for shuffling could be to maintain premium status with a large number of brokers and obtain valuable services such as access to sell-side research (Goldstein et al. (2009)). If a manager does not have a sufficient volume of orders to split on any given day, the manager may take turns submitting orders to different brokers in order to maintain close relationships with a large number of brokers. I argue that another important motive for shuffling is to conceal informed orders. To test this, I measure similarity between today's set of brokers used by a fund and the core set of brokers used by the same fund (from day t-25 to day t-6), as measured by cosine similarity based on the aggregate dollar volume of orders executed by each broker. I find that there is a larger share of informed trading volume on days when investment managers deviate from their core sets of brokers, that is, when the similarity is lower and the degree of shuffling is higher. This result is generally consistent with the randomization strategy envisioned in Yang and Zhu (Forthcoming) that investment managers can shuffle their order flows in order to conceal their informed orders. Again, a concern may be that investment managers trade larger quantities of shares on days with more informed trading volume and are naturally forced to send orders outside their core sets of brokers. However, the results are robust to controlling for trading volume, mitigating this concern.

So far my analysis has been limited to ANcerno client-manager pairs (funds) for which ANcerno managers could be matched with 13F managers and limited to manager-quarters with confidential 13F filings. As a robustness check, I extend this analysis to the full ANcerno sample. On every trading day I sort funds into quantiles based on cosine similarity of the dollar volume of orders executed by each broker between the set of brokers used today by a fund and the core set of brokers used by the same fund (from day t-25 to day t-6). Then I measure the value-weighted buy-and-hold return on the stocks bought on each day by each fund, weighted by dollar volume, for the next month (typically from day t+1 to day t+21). For each quintile, I first average value-weighted returns across funds and then use the time series average value-weighted returns for statistical inference in the spirit of Fama and MacBeth (1973), adjusting for serial correlation up to a lag of 20 trading days following Newey and West (1987). I find that the most dissimilar quintile outperforms the most similar quintile by 24 basis points per month (t = 3.20). When adjusting for DGTW-benchmark returns (Daniel et al. (1997)), the return spread remains large and statistically significant at 17 basis points per month (t = 2.77). This full sample result corroborates the previous finding that investment managers shuffle their orders across brokers more on days with a larger share of informed trading volume.

An important question that remains is whether the splitting and shuffling strategies work for investment managers in terms of reducing trading costs on informed orders. Following the literature (e.g., Anand et al. (2012), Anand et al. (2013)), I measure institutional trading costs using the implementation shortfall. First, I examine individual parent orders. I find that splitting orders across brokers on any given day tends to reduce trading costs for informed orders, whereas splitting uninformed orders has little effect on implementation shortfall. Since the first day portion of multi-day orders could be larger than the later day portions, which could affect both splitting decisions and trading costs, I control for order-sequence fixed-effects (first day, second day, etc.). The result is robust to controlling for various fixed-effects and order size. Next, I restrict my analysis to multi-day orders (about 30% of parent orders) in order to examine how orders are sequenced and shuffled across brokers over time. Investment managers may avoid sending remaining parts of the large informed orders to the brokers that executed early parts of the orders in order to conceal the underling informed trading motives. Consistent with this prediction, sending later parts of the orders to new brokers (i.e., shuffling) tends to lower trading costs on informed orders. Again, this result is robust to controlling for various fixed-effects and order size.

Finally, I examine portfolios of individual orders at a daily level to test whether the shuffling strategy works for investment managers. I find that dis-similarity between today's set of brokers used by a fund and the core set of brokers used by the same fund (one minus cosine similarity) is associated with lower trading costs as measured by implementation shortfall on informed orders. The result, combined with the earlier results, is consistent with investment managers shuffling informed orders across brokers and mixing informed orders with other orders sent to brokers, in order to mitigate trading costs on informed orders.

The remainder of the paper is organized as follows. Section 2 describes the data sources and the construction of my sample and variables. I report my results in Section 3. Section 4 concludes.

2 Data and Variable construction

My analysis of institutional order submission strategies of informed orders needs a detailed trade-level dataset that contains information on the institutional investors and brokers. Another important requirement of my analysis is identifying whether a trade is an informed trade or not. I describe the transaction data and how I construct my sample and then explain how I identify the informed trade.

I obtain proprietary data from institutional trading transactions via ANcerno data (also known as Abel/Noser or Abel Noser data).[4] The data cover equity trading transactions made by a large sample of institutions from January 1999 through December 2009. For each transaction, the data include the date

[4] Other recent studies using ANcerno data to examine the behavior of institutional investors include Chemmanur, He, and Hu (2009); Chemmanur, Hu, and Huang (2010), Goldstein et al. (2009), Goldstein, Irvine, and Puckett (2011), Puckett and Yan (2011), Anand et al. (2012, 2013), Hu et al. (2014), Jame (2018), Barbon et al. (Forthcoming), Ben-Rephael and Israelsen (2018), and Di Maggio et al. (2019).

of transaction, the stock traded (identified by both tickers and CUSIPs), the number of shares traded, transaction prices, and market prices at the times of the transactions, and whether it is a buy or sell by the institution. Additionally, ANcerno provided names of investment managers who order the transactions. This enables me to use the names to merge the ANcerno data with the 13F holdings data and identify whether the trading is informed trading or not based on holdings on confidential filings.[5]

My analysis is carried out at two different levels. At the parent-order level, I stitch all trades on the same stock on the same side of the market (buy or sell) by the same manager on behalf of the same client executed through multiple brokers and over consecutive multiple days. At the ticket level, I aggregate all trades on the same stock on the same side of the market (buy or sell) by the same manager and same client and executed through the same broker on the same day.

Next, I explain how I identify an informed trade based on "confidential filing". My purpose is to identify whether an order is driven by informed motivation – in other words, if institutional investors intentionally hide their holdings, as placing an order for hidden securities is more likely to be driven by information-motivated reasons.

I follow Agarwal et al. (2013) to verify confidential holdings for my sample. Basically, I directly retrieve both the original and the amended 13F filings (forms 13F-HR and 13F-HR/A[6]) dated between March 1999 and December 2009 from the SEC's EDGAR database. For the better comparison between the paired filings, I retrieve information about the original filings from the SEC rather than using Thomson Reuters' database. Amendments to 13F filings provide two types of information. The first type of information is disclosing the difference between the position that was previously filed and the position included on the amendment. The second type of information is a newly added position that was not previously filed. I define a holding as confidential if it is in a positive position on the amendment compared to the original filing or a newly added position on the amendment. Consequently, for each quarter, I know which

[5]I follow the matching process from Hu et al. (2018) and Choi et al. (2016). I matched institutional managers in ANcerno and 13F holdings data by comparing both the quarterly changes in holdings computed from the two datasets and the manager names.

[6]Form 13F-HR/A includes an indication on its cover regarding whether it is an "amendment" (i.e., whether it adds new holdings) or a "restatement." I only include forms with the "amendment" box checked.

securities are confidential holdings or not. Figure 1 provides a timeline of the original and confidential 13F filings. I define a trade between the quarter start and the quarter end as an informed motivated trade if it is placed for a security regarded as a confidential holding at the quarter end.

[Insert Figure 1]

The manager matching processed dataset gives rise to 129 managers among unique institutional managers in the ANcerno database. I merge the data with confidential holdings, resulting in matches with 61 managers. About 6 % of orders are considered informed orders following this process. As previously stated, I consider the parent-order level and daily transaction level. Panel A of Table 1 describes the characteristics of the parent-order level. The average of number of days of parent order is 1.38 days, and the average number of brokers who executed each order is 1.74. Panel B of Table 1 describes the characteristics of the daily order level.

[Insert Table 1]

Next, I examine how institutional investors place orders across brokers by constructing a measure that describes how similarly institutional investors submit their orders to their brokers based on daily transaction level data. I aggregate the dollar volume of orders across the same manager, client, and broker on the same day. I measure similarity between today's set of brokers used by a fund and the core set of brokers used by the same fund (from $t-25$ to day $t-6$), as measured by cosine similarity based on the dollar volume of orders executed by each broker. This proxy captures how similar today's set of brokers used by an investment manager (manager-client) is to the set of brokers used by the same manager in the past 20 days (I skip the 5 days immediately preceding the current day.)

For a robustness check, I construct two more measures to capture patterns of institutional investors' brokerage usage. First, like the first measure, I aggregate the dollar volume of orders across the same fund, as well as the broker, on the same day. I then consider the overlapping dollar volumes through brokers today and in the past days (from $t-25$ to day $t-6$). Second, I consider only one broker who executed the largest dollar volume for each fund during the past 20 days ($t-25$ to day $t-6$). I then use the proxy for determining how many orders the fund places through the top broker today in terms of the fraction of the aggregate dollar trading volumes. Ultimately, the main goal of all three measures is to capture how similarly the institutional investors submit their orders to brokers.

In order to examine how the randomization strategy across brokers affects trading costs, I construct the implementation shortfall measure (the percentage of change between the execution price and the benchmark price) at the daily level. In other words, I aggregate all trades on the same stock on the same side of market (buy or sell) by the same manager and same client and executed through same broker on the same day. Following Keim and Madhavan (1997) and Anand et al. (2012, 2013), I calculate implementation shortfall as follows:

$$\frac{P_1(t) - P_0(t)}{P_0(t)} \times D(t) \tag{1}$$

where $P_1(t)$ is the volume-weighted execution price, $P_0(t)$ is the price at prevailed in the market, and $D(t)$ is the sign of the trade.

3 Order Submission Strategies

In this section, I begin my empirical analysis at the parent-order level in Section 3.1, first showing how institutional investors place orders differently when trading on private information and then examining whether the order submission strategy across brokers can affect trading costs. In Section 3.2, I demonstrate that informed trades are more likely to show a higher degree of randomization across brokers. This randomization strategy also brings about better performance for institutional investors in terms of higher returns; the strategy can also reduce trading costs measured by the implementation shortfall when trading on information.

3.1 Parent-Order Level Analysis

Following Anand et al. (2012), I stitch together all orders on the same stock on the same side of the market (buy or sell) by the same fund that are executed through multiple brokers and over consecutive multiple days to construct parent orders. A parent order is a collection of tickets (child orders) on the same stock in the same direction that a fund may place with multiple brokers and over multiple days. As most institutions in ANcerno are long-only investors, I focus only on the buy side in this paper. I contend that, when trading on information, institutional investors can spread their orders over multiple days and split them across multiple brokers in order to avoid information leakage.

In order to test this hypothesis, I identify information-motivated orders by combining ANcerno transaction data with 13F filings (both original and confidential). Specifically, a parent buy order for which an investment manager has sought confidentiality treatment is identified as an informed trade. In other words, informed buy orders appear on the manager's confidential 13F filings, but not on the original 13F filings (or a more long position on the amended files compared to the original one), whereas uninformed buy orders appear on the manager's original 13F filings at the quarter end. Intuitively, an attempt to hide their positions by seeking confidentiality implies that managers have traded those stocks acting on their superior private information. According to Agarwal et al. (2013), confidential holdings outperform original holdings. Having separately identified information-motivated orders and uninformed ones, I estimate the following linear regression model:

$$Number\ of\ Days_{i,k,s,t} = \beta \times \mathbb{1}(Informed\ Motivated_{i,k,s,t}) + \alpha_{i,k} + \rho_s + \theta_t + \varepsilon_{i,k,s,t} \tag{2}$$

where i indexes managers, k indexes clients, s indexes stocks, and t indexes first trading days on which parent orders were placed. The dependent variable is the number of days for which parent orders are extended ($Number\ of\ Days_{i,k,s,t}$) or the number of brokers to which parent orders are sent ($Number\ of\ Brokers_{i,k,s,t}$) or the broker's Herfindahl-Hirschman Index ($Broker\ HHI_{i,k,s,t}$), which measures the degree of concentration across brokers to which a parent order is sent, calculated by squaring a fraction of the dollar volume executed by each broker and summing it over all brokers to which a parent order is submitted. The independent variable of interest, $\mathbb{1}(\textit{Informed Motivated}_{i,k,s,t})$, is an indicator variable that takes the value of one if a parent order is motivated by information and zero otherwise. Depending on the specification, the regression includes manager × client (fund) fixed effects ($\alpha_{i,k}$) and stock fixed effects (ρ_s). All regressions include time fixed effects (θ_t), and standard errors are clustered by manager × client (fund).

The baseline regression results are presented in Panel A of Table 2. In columns (1) and (2), the dependent variable is $\textit{Number of Days}_{i,k,s,t}$. In column (1), the coefficient of $\mathbb{1}(\textit{Informed Motivated}_{i,k,s,t})$ is positive and statistically significant at the 1% level. This is consistent with my hypothesis that institutional investors are more likely to split their order across multiple days when trading on private information. The result remains largely unchanged when controlling for stock fixed effects in column (2). In columns (3) and (4), the dependent variable is $\textit{Number of Brokers}_{i,k,s,t}$. In column (3), the coefficient of $\mathbb{1}(\textit{Informed Motivated}_{i,k,s,t})$ is positive and statistically significant at the 1% level. This suggests that institutional investors tend to place orders across more brokers when the orders are motivated by private information. Also, the result remains robust when controlling for stock fixed effects. In columns (5) and (6), I replace $\textit{Broker HHI}_{i,k,s,t}$ with $\textit{Broker HHI}_{i,k,s,t}$ in the above linear regression model. *Broker HHI* measures the degree of concentration across brokers to which a parent order is sent and is calculated by squaring a fraction of the dollar volume executed by each broker and summing it over all brokers to which a parent order is submitted. A lower *Broker HHI* implies that a parent order is spread out across a more diversified set of brokers. I contend that informed orders are more likely to be sent across a diversified set of brokers in order to hide the information content of the order from the brokers and other market participants. From columns (5) and (6), $\mathbb{1}(\textit{Informed Motivated})$ is negatively associated with *Broker HHI* and statistically significant at 1%, controlling for stock fixed effects. This implies that informed orders are sent in a scattered way across brokers when compared to uninformed orders.

One might be concerned that my results could be driven by order sizes. In other words, informed orders could be larger, and larger orders may require more days and more brokers to fill. In order to address this concern, I re-run the tests in a sub-sample of "block orders" (large orders of more than 10,000 shares or with volumes of more than 200,000 dollars). The regression results are presented in Panel B of Table 2. I continue to find qualitatively similar, even stronger, results that large informed orders are spread over more days and sent across more brokers than large uninformed orders. I similarly control for several proxies' capture effects of order size-namely, log of dollar volumes, percentage of shares traded in CRSP volume, and percentage of shares traded in outstanding shares. The results are reported in Panel C, Panel D, and Panel E, respectively. I continue to obtain qualitatively similar results.

[Insert Table 2]

Next, I examine day-to-day observations of each parent order. From the previous analysis of the parent-order level, I find that institutional investors tend to spread their informed orders out more over time and across multiple brokers. Possible questions include how the splitting strategy across days and brokers works better for institutional investors. I argue that the strategy can reduce trading costs because spreading out the orders can deter brokers from detecting the intention of the original orders, thereby reducing the risk of leaking information to others. I compute the implementation shortfall, which is the percentage change between the execution price and a benchmark price as the transaction cost. (Keim and Madhavan (1997) and Anand et al. (2012, 2013)) In order to test the idea, I estimate the following linear regression model:

$$\begin{aligned} \textit{Implementation Shortfall}_{i,k,s,b,t} &= \delta \times \mathbb{1}(\textit{Informed Motivated}_{i,k,s,t}) \times \textit{Split Brokers}_{i,k,s,t} \\ &+ \beta \times \mathbb{1}(\textit{Informed Motivated}_{i,k,s,t}) + \lambda \times \textit{Split Brokers}_{i,k,s,t} \\ &+ \alpha_{i,k} + \gamma_d + \rho_s + \eta_b + \theta_t + \varepsilon_{i,k,s,b,d,t} \end{aligned} \tag{3}$$

where i indexes managers, k indexes clients, s index stocks, b indexes brokers, and t indexes trade date. The dependent variable $\textit{Implementation Shortfall}_{i,k,s,b,t}$, is the percentage difference between the execution price and a benchmark price. The independent variable of interest, $\mathbb{1}(\textit{Informed Motivated}_{i,k,s,t})$ denotes whether the order is driven by informed reason or not and $\textit{Split Brokers}_{i,k,s,t}$ is an indicator variable that takes the value of one if an institutional investor hires more than one broker, zero otherwise. (Daily Basis) Depending on the specification, the regression includes manager × client (fund) fixed effects ($\alpha_{i,k}$), stock fixed effects (ρ_s),broker fixed effects (η_b), and time fixed-effects (θ_t). All regressions include order sequence fixed effects (γ_d) and standard errors are clustered by manager $\times$ client.

The regression results are presented in Panel A of Table 3. All regressions include order-sequence fixed effects to control for unobserved characteristics of first day, second day, and so on of the parent order. Consistent with my prediction, working through multiple brokers to execute an order can bring lower trading costs to institutional investors, especially when they are trading on information. As trading volumes can impact price, I control for log of dollar volumes and then continue to find qualitatively similar results in Panel B. Furthermore, in Panels C and D, I use a sub-sample that includes multi-day orders, which occur when the duration of the parent order is more than one day. I find the coefficient of the interaction term between $\mathbb{1}(\textit{Informed Motivated}_{i,k,s,t})$ and $\textit{Split Brokers}_{i,k,s,t}$ to be much stronger than the one from the regression of the entire sample. This result suggests that the effect of spreading out informed orders across brokers on the trading costs is more pronounced when the order is a multi-day order.

[Insert Table 3]

I show that submitting orders to multiple brokers can help institutional investors reduce the trading costs when trading on private information. The remaining question at the parent-order level can be whether switching brokers over time to complete orders affects the trading costs or not. I contend that it can reduce the trading costs when institutional investors submit orders after the first day to new brokers who are different from the broker to whom investors submitted orders on the first day. To test my argument, I estimate the following linear regression model:

$$\begin{aligned}\textit{Implementation Shortfall}_{i,k,s,b,t} = {}& \delta \times \mathbb{1}(\textit{Informed Motivated}_{i,k,s,t}) \times \mathbb{1}(\textit{Order Sent to Different Broker}_{i,k,s,t}) \\ & + \beta \times \mathbb{1}(\textit{Informed Motivated}_{i,k,s,t}) + \lambda \times \mathbb{1}(\textit{Order Sent to Different Broker}_{i,k,s,t}) \\ & + \alpha_{i,k} + \gamma_d + \rho_s + \eta_b + \theta_t + \varepsilon_{i,k,s,b,d,t}\end{aligned} \tag{4}$$

where i indexes managers, k indexes clients, s index stocks, b indexes brokers, d indexes order-sequence, and t indexes trade date. The dependent variable $\textit{Implementation Shortfall}_{i,k,s,b,t}$, is the percentage difference between the execution price and a benchmark price. The independent variable of interest, $\mathbb{1}(\textit{Informed Motivated}_{i,k,s,t})$ denotes whether the order is driven by informed reason or not and $\mathbb{1}(\textit{Order Sent to Different Broker}_{i,k,s,t})$ is one if an order is sent to different brokers after the first day, zero otherwise. Depending on the specification, the regression includes manager × client (fund) fixed effects ($\alpha_{i,k}$), stock fixed effects (ρ_s), broker fixed effects (η_b), and time fixed-effects (θ_t). All regressions include order-sequence fixed effects (γ_d) and standard errors are clustered by manager × client.

The regression results are presented in Table 4. I show that switching brokers from the first day to the remaining days in a parent order can decrease the trading costs. Like the previous test, I control for order-sequence fixed effects. Moreover, I control for trading volumes and find qualitatively similar results, as shown in Panel B.

[Insert Table 4]

3.2 Evidence from Aggregate Daily Transactions

In the previous section, I examined institutional order submission strategies when trading on private information at the parent-order level. In this section, I analyze the order submission strategies for daily observations. First, I provide a camouflage strategy that can serve as a bridge between the analysis of parent order and analysis of daily observations. Second, I demonstrate that institutional informed orders bring a higher degree of randomization across brokers as part of their trading strategies.

My analysis of the parent order provides evidence that institutional investors are more likely to spread their orders over time and across multiple brokers, especially when they trade on information. Moreover, the splitting strategies can reduce trading costs. The mechanism behind the result is that spreading orders across brokers and over time can make information diffusion through brokers to other investors harder. Furthermore, I examine whether institutional investors might try to hide their informed intention by submitting a bunch of uninformed orders together when they place informed orders, which I call the camouflage strategy. In order to test this argument, I estimate the following linear regression model:

$$\textit{log(Number of Stocks)}_{i,k,b,t} = \beta \times \mathbb{1}(\textit{Informed Order to Broker}_{i,k,b,t}) + \alpha_{i,k} + \eta_b + \theta_t + \varepsilon_{i,k,b,t} \quad (5)$$

where i indexes managers, k indexes clients, b indexes brokers, and t indexes time. The dependent variable $\textit{log(Number of Stocks)}_{i,k,b,t}$, is the number of stocks which are executed through broker at manager-client-tradedate-broker level. The independent variable of interest, $\mathbb{1}(\textit{Informed Order to Broker}_{i,k,b,t})$ is an indicator variable which is equal to one if a investor submit at least one informed order to a broker, zero otherwise. Depending on the specification, the regression includes includes manager $\times$ client (fund) fixed effects ($\alpha_{i,k}$), broker fixed effects (η_b), and time fixed-effects (θ_t). Standard errors are clustered by manager $\times$ client.

The results are reported in Table 5. In columns (1)-(3) of Panel A, I show that institutional investors are more likely to submit a greater number of stocks to a broker to whom they have been simultaneously submitting informed orders. Moreover, I find robust results after controlling for log of volumes in columns (4)-(6). As a robustness check, I replace $\textit{log(Number of Stocks)}_{i,k,b,t}$ with $\textit{Stock HHI}_{i,k,b,t}$. $\textit{Stock HHI}_{i,k,b,t}$ is calculated by squaring the portion of the traded aggregate volumes of each stock traded through a broker for investors and then summing the results. According to the results, I can infer that institutional investors are not only submitting a greater number of stocks, but also evenly placing orders across the stocks.

[Insert Table 5]

Such order submission strategies designed to hide informed trades at the individual order level lead to a large degree of randomization of order flows across brokers at the level of portfolios of individual orders. To see this, I aggregate all buy orders on the same stock by the same fund executed through the same broker on the same day. Then I measure the extent to which a set of brokers an investment manager uses overlap with or deviate from the set of brokers the same investment manager has used in the past. Specifically, I measure the similarity between today's set of brokers used by an investment manager and the set of brokers used by the same manager one day later, two days later, all the way through twenty days later, as measured by cosine similarity based on the dollar volume of orders executed by each broker. Figure 2 shows that investment managers tend to submit their orders to a slightly different set of brokers every day. Looking across days that are further apart, the cosine similarity of the set of brokers used by each manager further drops, but at a decreasing rate and starts to converge in just a few days. This suggests that while shuffling order flows sent across brokers to hide informed orders, investment managers typically maintain close trading relationships with their core sets of brokers, who can facilitate liquidity provision when managers submit large liquidity-motivated orders (Han, Kim, and Nanda (2019)).

[Insert Figure 2]

An important question at this point is why investors shuffle their order flows across brokers everyday. One reason could be to maintain premium status with a large number of brokers and receive valuable services such as access to sell-side research (Goldstein et al. (2009)) Another important motive for shuffling could be hiding informed orders. Thus, I examine whether the investors' order submission strategy across brokers is different depending on the existence of information in order flows. I speculate that institutional investors shuffle more order flows through brokers when trading on their private information to deter other investors from mimicking their trades based on leaked information through brokers. In order to test this hypothesis, I measure similarity between today's set of brokers used by a fund and the core set of brokers used by the same fund (from $t-25$ to day $t-6$), as measured by cosine similarity based on the dollar volume of orders executed by each broker. I regress dissimilarity (1-cosine similarity) measure on the an indicator variable for identifying the day when informed trading happens enough at manager-client level in the following linear regression model:

$$Dissimilarity_{i,k,t} = \beta \times \mathbb{1}(Informed\ Shares > \ 0.5_{i,k,t}) + \alpha_{i,k} + \theta_t + \varepsilon_{i,k,t} \tag{6}$$

where i indexes managers, k indexes clients, and t indexes time. The dependent variable $Dissimilarity_{i,k,t}$, is 1−cosine similarity between the set of manager-client's aggregate trading dollar volume through its broker at today and that from 25 days before to 5 days before today. The independent variable of interest,

$\mathbb{1}(\textit{Informed Shares} > \textit{0.5}_{i,k,t})$ is an indicator variable identifying the day when enough informed trades happen (the portion of informed trades out of the entire executed order a day is more than 0.5). Depending on the specification, the regression includes manager × client (fund) fixed effects ($\alpha_{i,k}$), and time fixed-effects (θ_t) and standard errors are clustered by manager × client.

Panel A of Table 6 shows the results from regressing $\textit{Dissimilarity}_{i,k,t}$ on $\mathbb{1}(\textit{Informed Shares} > \textit{0.5}_{i,k,t})$, including various fixed effects. In column (1), the coefficient of $\mathbb{1}(\textit{Informed Shares} > \textit{0.5}_{i,k,t})$ is positive and statistically significant at the 1% level, which is consistent with my prediction. This result remains robust when controlling for time-fixed effects in column (2). A possible alternative of the results is that large orders can make investors order through various brokers, which can lead to higher dissimilarity. I control for log of dollar volumes, the number of shares traded, and the number of stocks traded on one day to mitigate the concerns about the alternative conjecture. I find qualitatively similar results after controlling for these proxies for the order size in columns (4)-(6).

Similarly, I provide additional robustness checks by constructing more measures of dissimilarity. The first measure is one minus the overlapped dollar volumes traded through the set of brokers used by an investment manager and the set of brokers used by the same manager in the past. For the third measure, I compute how much the manager-client trades through their top broker, which is chosen based on the trading information from 25 to 5 days before the current day (i.e., today). I replace cosine similarity with the computed measure. These results are reported in Panel B and Panel C of Table 6, respectively. My results are robust when I use different measures of dissimilarity.

[Insert Table 6]

By research design, my analysis is limited to ANcerno manager-client pairs (funds) where ANcerno managers can be matched with 13F managers and manager-quarters when confidential filings. I extend the previous analysis to the full ANcerno sample. I examine the effect of the randomization strategy across brokers on performance. Every trading day I sort funds into quantiles based on similarity based on cosine similarity based on the dollar volume of orders executed by each broker between today's set of brokers used by a fund and the core set of brokers used by the same fund (from day t-25 to day t-6). Then I measure the value-weighted buy-and-hold return on the stocks bought on each day by each fund, weighted by dollar volume, for the next month (typically from day t+1 to day t+21). For each quintile, I first average value-weighted returns across funds then use the time series average value-weighted returns for statistical inference in the spirit of Fama and MacBeth (1973) and adjust for serial correlation up to a lag of 20 trading days following Newey and West (1987). I find that the most dissimilar quintile outperform the most similar quintile by 24 basis points per month (t = 3.20). When adjusting for DGTW-benchmark returns (Daniel et al. (1997)), the return spread remains large and statistically significant at 17 basis points per month (t = 2.77). This full sample result corroborates the previous finding that investment managers shuffle their orders across brokers more on days with a larger share of informed trading volume.

[Insert Table 7]

When institutional investors are trading on their private information, they believe that the information will eventually be publicly disclosed and the price will reflect the information, accordingly. However, they want to maximize their possible profits from this information. The primary prediction that I can derive from my hypothesis is that a randomization strategy across brokers can deter brokers from leaking private information, thereby reducing trading costs on informed orders. I compute implementation shortfall which is the percentage difference between the execution price and a benchmark price as transaction costs. (Keim and Madhavan (1997) and Anand et al. (2012, 2013)) This measure is at manager-client-stock-broker-day. Then, I interact dissimilarity measure with the indicator variable for identifying information motivated trade and estimate the following linear regression model:

$$\begin{aligned} \textit{Implementation Shortfall}_{i,k,s,b,t} &= \delta \times \mathbb{1}(\textit{Informed Motivated}_{i,k,s,t}) \times \textit{Dissimilarity}_{i,k,t} \\ &+ \beta \times \mathbb{1}(\textit{Informed Motivated}_{i,k,s,t}) + \lambda \times \textit{Dissimilarity}_{i,k,t} \\ &+ \alpha_i + \gamma_k + \rho_s + \eta_b + \theta_t + \varepsilon_{i,k,s,b,t} \end{aligned} \tag{7}$$

where i indexes managers, k indexes clients, s index stocks, b indexes brokers, and t indexes time in day. The dependent variable $\textit{Implementation Shortfall}_{i,k,s,b,t}$, is the percentage difference between the execution price and a benchmark price. The independent variable of interest, $\mathbb{1}(\textit{Informed Motivated}_{i,k,s,t})$ denotes whether the order is driven by informed reason or not and $Dissimilarity_{i,k,t}$ is 1−cosine similarity between the set of manager-client's aggregate trading dollar volume through its broker at today and that from 25 days before to 5 days before today. Depending on the specification, the regression includes manager × client (fund) fixed effects ($\alpha_{i,k}$), stock fixed effects (ρ_s), and broker fixed effects (η_b). All regressions include time fixed-effects (θ_t) and standard errors are clustered by manager × client.

Panel A of Table 8 reports the results. Consistent with my prediction, I find a negative and statistically significant coefficient on $\mathbb{1}(Informed\ Motivated_{i,k,s,t}) \times Dissimilarity_{i,k,t}$ and I continue to find the robust results after using different specifications including several fixed effects in columns (1)-(3). The estimates are also economically significant as the implementation shortfall decrease at around 12 bps. In contrast, the coefficients on $\mathbb{1}(Informed\ Motivated_{i,k,s,t})$ are positive and statistically significant. The results suggest that more randomizing order flows across brokers can reduce the execution cost by hiding the true intention of order. Moreover, even though the coefficient on $Dissimilarity_{i,k,t}$ is not statistically insignificant, it is positive. Potentially, it can be interpreted as that more shuffling order flows across brokers can be relatively more expensive, thereby increase transaction costs. Moreover, I re-run the regression by using sub-sample which includes block orders (the aggregate volume of trade is greater than equal to 10,000 shares or the aggregate dollar volumes of trade is greater than equal to $200,000). Then I continue to obtain qualitative similar but much stronger magnitude of coefficients in columns (4)-(6).

Since trading volume can not only impact on price but also can contain information itself, I control for order sizes by using several proxies capture effects of order size. They are log of dollar volumes, log of volumes, percentage of shares traded in CRSP volume, and percentage of shares traded in outstanding shares. The results are reported in Panel B and C respectively. I continue to obtain qualitatively similar results.

[Insert Table 8]

3.3 Robustness Checks

In this section, I demonstrate the robustness of my main results regarding splitting and shuffling strategy by extending sample to the full ANcerno sample. For the main results, I separate information-motivated orders from uninformed ones by matching ANcerno transaction data with confidential 13F filings. Unfortunately, this process results in a large loss of data, as it requires matching ANcerno managers and 13F managers, yet not all ANcerno-13F matched managers seek confidentiality treatment on their holdings. Therefore, I extend the analysis to the entire ANcerno dataset to corroborate my main findings by using a different proxy to identify informed orders. Form 13F requires that all investment managers who have investment discretion over \$100 million or more in Section 13(f) securities to disclose their quarter-end holdings in these securities (Agarwal et al. (2013)). If private information arrives near the quarter end, an investment manager might want to delay building a large position until the new quarter starts in order to avoid prematurely reporting a new informed position. In addition, investment managers may exert more effects into acquiring private information early in the quarter for the same reason. Hence, investment managers are likely to trade more heavily on private information earlier in the quarter. Based on this understanding, I estimate the following linear regression model:

$$\textit{Number of Days}_{i,k,s,t} = \beta \times \textit{Days to Quarter-end}_{i,k,s,t} + \alpha_{i,k} + \rho_s + \theta_t + \varepsilon_{i,k,s,t} \tag{8}$$

where i indexes managers, k indexes clients, s indexes stocks, and t indexes first trading days on which parent orders were placed. The dependent variable is the number of days for which parent orders extend ($\textit{Number of Days}_{i,k,s,t}$) or the number of brokers to which parent orders are sent ($\textit{Number of Brokers}_{i,k,s,t}$) or broker Herfindahl-Hirschman Index ($\textit{Broker HHI}_{i,k,s,t}$) which measures the degree of concentration across brokers to which a parent order is sent, calculated by squaring a fraction of the dollar volume executed by each broker and summing it over all brokers to which a parent order is submitted. The independent variable of interest, Days to Quarter-end$_{i,k,s,t}$, is the number of days from the first day on which a parent order is submitted to the calendar quarter-end. Depending on the specification, the regression includes manager×client (fund) fixed-effects ($\alpha_{i,k}$), stock fixed-effects (ρ_s) and quarter fixed-effects (θ_t). Standard errors are clustered by manager × client (fund) for all regressions.

The results are reported in Panel A of Table A1. In columns (1) and (2), the coefficient of Days to Quarter-end$_{i,k,s,t}$ is positive and statistically significant at the 1% level, controlling for various fixed effects. This suggests that trade orders executed early in the quarter (presumably informed orders) require more days to complete. In columns (3) and (4), the coefficient of Days to Quarter-end$_{i,k,s,t}$ is positive and statistically significant, implying that orders tend to be placed across more brokers earlier in the quarter. From columns (5) and (6), I replace $Broker\ HHI_{i,k,s,t}$ with $Broker\ HHI_{i,k,s,t}$. I find that trades in the earlier period of the quarter are more likely to be sent across a diversified set of brokers. In order to mitigate concerns that my results could be driven by order size, I re-run the same linear regression in a sub-sample of large orders. However, the results reported in Panel B of Table A1 remain qualitatively similar and somewhat stronger. Overall, the results suggest that informed orders identified as those executed early in the quarter are likely spread over more days across more brokers and consistent with my prediction that investment managers attempt to conceal information content of orders from the brokers. As a robustness check, I control for several measures of order size: log of dollar volumes, percentage of shares traded in CRSP volume, and percentage of shares traded in outstanding shares. The results are reported in Panel C, Panel D, and Panel E, respectively. I continue to obtain qualitatively similar results.

[Insert Table A1]

Examining trades executed during the earlier period of quarter is allowing me not to lose lots of observations. Hence, I regress $Dissimilarity_{i,k,t}$ on $Days\ to\ Quarter\text{-}end_{i,k,t}$ including various fixed effects in the following linear regression model:

$$Dissimilarity_{i,k,t} = \beta \times Days\ to\ Quarter\text{-}end_{i,k,t} + \alpha_{i,k} + \theta_t + \varepsilon_{i,k,t} \quad (9)$$

where i indexes managers, k indexes clients, and t indexes time. The dependent variable $Dissimilarity_{i,k,t}$, is 1−cosine similarity between the set of manager-client's aggregate trading dollar volume through its broker at today and that from 25 days before to 5 days before today. The independent variable of interest, $\textit{Days to Quarter-end}_{i,k,t}$ is the number of days from the first day of order to quarter-end. Depending on the specification, the regression includes manager × client (fund) fixed effects ($\alpha_{i,k}$), and time fixed-effects (θ_t) and standard errors are clustered by manager × client. The regression results are presented in Table A2. For all three proxies, I find that trades executed early in the quarter (presumably informed orders) tend to randomize their order flows across brokers. This evidence supports my hypothesis that investors are more likely to shuffle order flows across brokers when they are trading on their private information. This result is robust after controlling for various fixed effects and controlling daily aggregate dollar volumes, the number of shares traded, and the number of stocks traded.

[Insert Table A2]

4 Conclusion

In this paper, I study order submission strategies by institutional investors when trading on private information by exploiting proprietary data on institutional daily transactions. I separate informed orders from uninformed order by using 13F confidential holdings following Agarwal et al. (2013). I find that institutional investors tend to split orders over more days to complete the order and across more brokers when they are trading on private information. Furthermore, while institutional investors place uninformed large orders to the same brokers over multiple days, they tend to submit informed orders after the first day to new brokers who are different from the broker to whom they submitted orders on the first day.

Institutional investors not only shuffle their orders across brokers over time, but they also provide camouflage for their informed orders by mixing an informed order with other uninformed orders simultaneously sent to the same broker. As a result, a higher degree of shuffling a portfolio of orders is associated with a larger share of informed trading volume. The splitting and shuffling strategies designed to conceal informed trades from brokers and other market participants tend to lower institutional trading costs as measured by implementation shortfall, especially on informed orders.

References

Admanti, Anat R, and Paul Pfleiderer, 1991, Sunshine trading and financial market equilibrium, *The Review of Financial Studies* 4, 443–481.

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

Agarwal, Vikas, Wei Jiang, Yuehua Tang, and Baozhong Yang, 2013, Uncovering hedge fund skill from the portfolio holdings they hide, *Journal of Finance* 68, 739–783.

Anand, Amber, Paul Irvine, Andy Puckett, and Kumar Venkataraman, 2012, Performance of institutional trading desks: An analysis of persistence in trading costs, *Review of Financial Studies* 25, 557–598.

Anand, Amber, Paul Irvine, Andy Puckett, and Kumar Venkataraman, 2013, Institutional trading and stock resiliency: Evidence from the 2007-2009 financial crisis, *Journal of Financial Economics* 108, 773–797.

Barbon, Andrea, Marco Di Maggio, Francesco A. Franzoni, and Augustin Landier, Forthcoming, Brokers and Order Flow Leakage: Evidence from Fire Sales, *Journal of Finance* .

Ben-Rephael, Azi, and Ryan D. Israelsen, 2018, Are some clients more equal than others? An analysis of asset management companies' execution costs, *Review of Finance* 22, 1705–1736.

Chemmanur, Thomas J., Shan He, and Gang Hu, 2009, The role of institutional investors in seasoned equity offerings, *Journal of Financial Economics* 94, 384–411.

Chemmanur, Thomas J., Gang Hu, and Jiekun Huang, 2010, The role of institutional investors in initial public offerings, *Review of Financial Studies* 23, 4496–4540.

Choi, Jaewon, Ji Min Park, Neil D. Pearson, and Shastri Sandy, 2016, A First Glimpse into the Short Side of Hedge Funds, *Ssrn* 1.

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

Di Maggio, Marco, Francesco Franzoni, Amir Kermani, and Carlo Sommavilla, 2019, The relevance of broker networks for information diffusion in the stock market, *Journal of Financial Economics* .

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

Fama, Eugene F, and James D MacBeth, 1973, Risk, Return, and Equilibrium: Empirical Tests, *Journal of Political Economy* 84, 407–409.

Goldstein, Michael A., Paul Irvine, Eugene Kandel, and Zvi Wiener, 2009, Brokerage commissions and institutional trading patterns, *Review of Financial Studies* 22, 5175–5212.

Goldstein, Michael A., Paul Irvine, and Andy Puckett, 2011, Purchasing IPOs with commissions, *Journal of Financial and Quantitative Analysis* 46, 1193–1225.

Han, Munhee, Sanghyun (Hugh) Kim, and Vikram K. Nanda, 2019, Institutional Brokerage Networks: Facilitating Liquidity Provision, *SSRN Electronic Journal* .

Hu, Gang, Bin Ke, and Yong Yu, 2018, Can transient institutions correctly interpret small negative earnings surprises in the absence of access to management's private information?, *Journal of Accounting, Auditing and Finance* 33, 3–33.

Hu, Gang, R. David McLean, Jeffrey Pontiff, and Qinghai Wang, 2014, The year-end trading activities of institutional investors: Evidence from daily trades, *Review of Financial Studies* 27, 1593–1614.

Jame, Russell, 2018, Liquidity Provision and the Cross Section of Hedge Fund Returns Liquidity Provision and the Cross Section of Hedge Fund Returns, *Management Science* .

Keim, Donald B., and Ananth Madhavan, 1997, Transactions costs and investment style: An inter-exchange analysis of institutional equity trades, *Journal of Financial Economics* 46, 265–292.

Korajczyk, Robert A., and Dermot Murphy, 2019, High Frequency Market Making to Large Institutional Trades, *Review of Financial Studies* .

Kyle, Albert S, 1985, Continuous auctions and insider trading, *Econometrica: Journal of the Econometric Society* 1315–1335.

Newey, Whitney K, and Kenneth D West, 1987, A Simple , Positive Semi-Definite , Heteroskedasticity and Autocorrelation Consistent Covariance Matrix, *Econometrica* 55, 703–708.

Puckett, Andy, and Xuemin Sterling Yan, 2011, The Interim Trading Skills of Institutional Investors, *Journal of Finance* 66, 601–633.

Seppi, Duane J., 1990, Equilibrium Block Trading and Asymmetric Information, *Journal of Finance* 45, 73–94.

Van Kervel, Vincent, and Albert J. Menkveld, 2019, High-Frequency Trading around Large Institutional Orders, *Journal of Finance* 74, 1091–1137.

Yang, Liyan, and Haoxiang Zhu, Forthcoming, Back-Running: Seeking and Hiding Fundamental Information in Order Flows, *Review of Financial Studies* .

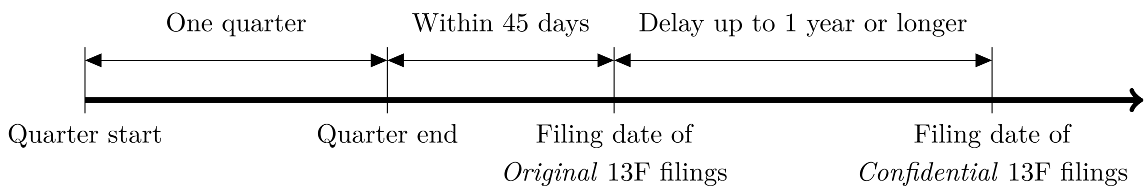

Figure 1: Timeline of the original and confidential 13F filings

This figure depicts timeline of the original and confidential 13F filings.

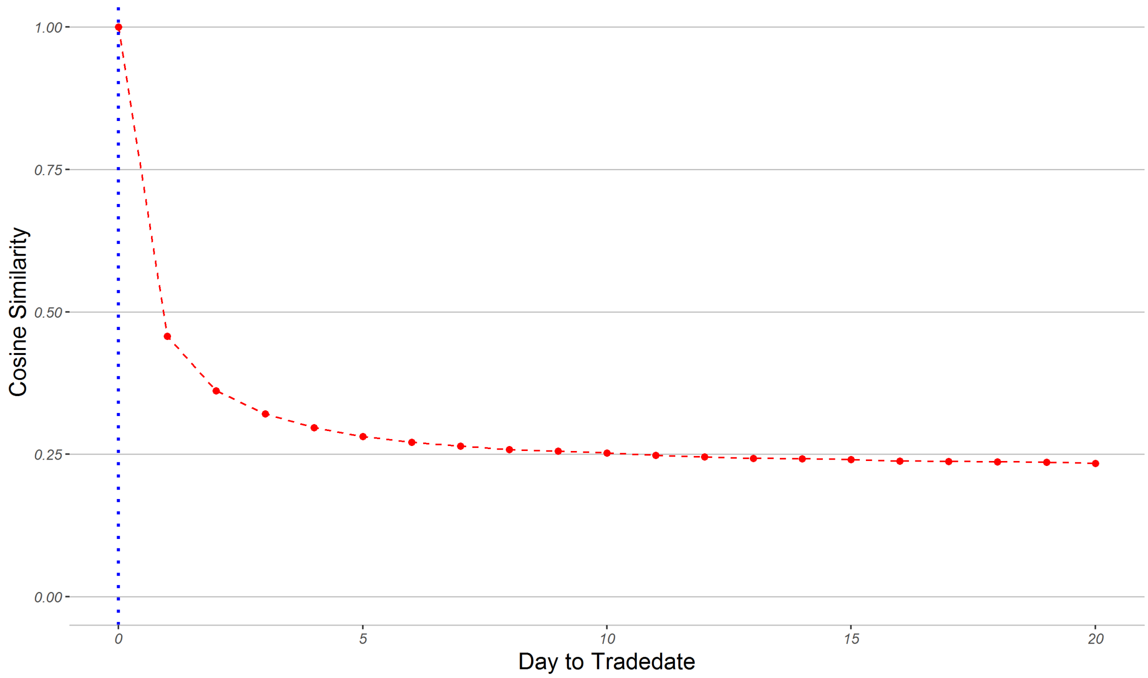

Figure 2: Shuffling order flows across brokers

This figure shows how different order flows across brokers day-to-day. I aggregate all buy side trades, by the same manager and client (fund), executed through the same broker, on the same day. Then I compute cosine similarity between today's set of brokers used by a manager-client (fund) and the set of brokers used by the same fund one day later, two days later, all the way through twenty days later based on the dollar volumes of orders executed by each broker.

Table 1: Summary Statistics

This table reports the summary statistics. Panel A describes the characteristics of parent-order level (based on confidential holdings). I stitch all trades on the same stock, on the same side of market (focus on buy side), by the same manager and same client, continued day for the parent-order level. *nBrokers* is the number of brokers who worked for each parent order. *nDays* is how many days a parent order completed. *Broker HHI* measures the degree of concentration across brokers to which a parent order is sent, calculated by squaring a fraction of the dollar volume executed by each broker and summing it over all brokers to which a parent order is submitted. *Dollar Volume* is the total executed dollar volumes for a parent order and *volume (# of shares)* is the total number of shares executed through a parent order. *% in Trading Volume (CRSP)* (*% in Outstanding Volume (CRSP)*) is the executed volumes on the first day of each parent order scaled by CRSP trading volume (CRSP outstanding share) of the first day. $\mathbb{1}$(Informed Motivated) is an indicator variable if the executed stock of a parent order is contained in a confidential file or not. Panel B presents variables on daily order level. (Only buy side) Cosine similarity is calculated between the set of manager-client (fund)'s aggregate dollar volumes through its broker at today and that from 25 days before to 5 days before today (from $t-25$ to $t-6$). In a similar way, similarity (Top 1) is computed as how much manager-client (fund) trades through their top broker which is chosen based on trading information from 25 days before to 5 days before today (from $t-25$ to $t-6$). *% in Trading Volume (CRSP)* (*% in Outstanding Volume (CRSP)*) is the executed volumes on a day scaled by CRSP trading volume (CRSP outstanding share) of the day. Implementation shortfall is the percentage difference between the execution price and benchmark price. $\mathbb{1}$(Informed Motivated) is an indicator variable if the executed stock is contained in a confidential file or not. All the reported samples are merged with the sample of confidential holdings and I follow Agarwal et al. (2013) to identify confidential holdings.

Variable	N	Mean	St. Dev.	Q_1	Median	Q_3
Panel A: Parent-order (Confidential Filings)						
nBrokers	214,250	1.38	0.88	1.00	1.00	1.00
nDays	214,250	1.74	1.84	1.00	1.00	2.00
Broker HHI (x 100)	214,250	91.42	18.35	100.00	100.00	100.00
Dollar Volume ($ 000)	214,250	1,756.51	6,910.29	32.98	154.75	802.07
Volume (# of shares)	214,250	72,035.19	330,360.10	1,100.00	5,600.00	31,800.00
% in Trading Volume (CRSP)	214,242	2.31	7.21	0.04	0.24	1.38
% in Outstanding Share (CRSP)	214,250	1.50	5.46	0.03	0.17	0.93
$\mathbb{1}$(Informed Motivated)	214,250	0.04	0.21	0.00	0.00	0.00
Panel B: Daily-order (Confidential Filings)						
Cosine Similarity	279,758	0.74	0.27	0.68	0.87	0.91
Similarity(Top 1 Broker)	279,758	0.12	0.15	0.05	0.10	0.10
Dollar Volume ($ 000)	279,758	1,053.63	2,337.25	29.27	172.48	862.83
Volume (# of shares)	279,758	43,225.40	91,528.08	1,200.00	9,000.00	40,000.00
% in Trading Volume (CRSP)	279,758	6.50	9.95	0.16	1.79	9.11
% in Outstanding Share (CRSP)	279,758	4.55	7.76	0.11	1.19	5.58
Implementation Shortfall (%)	279,758	0.10	1.49	−0.43	0.01	0.61
$\mathbb{1}$(Informed Motivated)	279,758	0.08	0.28	0.00	0.00	0.00

Table 2: Split or Concentrate Large Orders? Evidence from Confidential Holdings

This table examines whether the informed trades tend to split orders across more days and more brokers. Specifically, I regress $\textit{Number of Days}_{i,k,s,t}$ on $\mathbb{1}(\textit{Informed Motivated}_{i,k,s,t})$ in my specification as follows:

$$\textit{Number of Days}_{i,k,s,t} = \beta \times \mathbb{1}(\textit{Informed Motivated}_{i,k,s,t}) + \alpha_{i,k} + \rho_s + \theta_t + \varepsilon_{i,k,s,t} \quad (10)$$

where i indexes managers, k indexes clients, and t indexes the first day in parent order. The dependent variable, $\textit{Number of Days}_{i,k,s,t}$, is the number of days in each parent order. In column (3) through (4), $\textit{Number of Days}_{i,k,s,t}$ is replaced by $\textit{Number of Days}_{i,k,s,t}$, which is the number of brokers in each parent order. In column (5) through (6), the dependent variable is replaced by $\textit{Broker HHI}_{i,k,s,t}$ which is the measure of institutional investors' concentration on usage of brokers. The independent variable of interest, $\mathbb{1}(\textit{Informed Motivated}_{i,k,s,t})$ denotes whether the order is driven by informed reason or not. Panel A presents the baseline results. In Panel B, I examine block trades which include a trade's volume is greater than or equal to 10,000 or its dollar volume is greater than or equal to 200,000. Panel C presents results when controlling for log(Dollar Volume). In Panel D, I control for trading volume, % in CRSP volume and I also control for trading volume, % in outstanding share in Panel E. Depending on the specification, the regression includes manager × client fixed effects ($\alpha_{i,k}$), and stock fixed effects (ρ_s). All regressions include time fixed-effects (θ_t). Standard errors are clustered by manager × client and the resulting t-statistics are reported in parentheses. Statistical significance at the 10%, 5%, and 1% level is indicated by *, **, and ***, respectively.

Panel A: Baseline						
Dependent variable:	Number of Days		Number of Brokers		Broker HHI	
	(1)	(2)	(3)	(4)	(5)	(6)
$\mathbb{1}$(Informed Motivated)	0.64***	0.48***	0.18***	0.18***	−3.53***	−3.14***
	(3.49)	(4.05)	(3.79)	(3.60)	(−3.65)	(−4.13)
Time Fixed-effects	Yes	Yes	Yes	Yes	Yes	Yes
Stock Fixed-effects	No	Yes	No	Yes	No	Yes
Manager × Client Fixed-effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	214,250	214,250	214,250	214,250	214,250	214,250
Adjusted R^2	0.12	0.17	0.13	0.16	0.12	0.15
Panel B: Block Trade (Volume ≥ 10K or Dollar volume ≥ 200K)						
Dependent variable:	Number of Days		Number of Brokers		Broker HHI	
	(1)	(2)	(3)	(4)	(5)	(6)
$\mathbb{1}$(Informed Motivated)	0.89***	0.61***	0.21***	0.22***	−3.83***	−3.28***
	(8.79)	(9.56)	(10.85)	(7.77)	(−7.27)	(−6.68)
Time Fixed-effects	Yes	Yes	Yes	Yes	Yes	Yes
Stock Fixed-effects	No	Yes	No	Yes	No	Yes
Manager × Client Fixed-effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	104,522	104,522	104,522	104,522	104,522	104,522
Adjusted R^2	0.14	0.23	0.14	0.18	0.12	0.15

Table 2–*Continued*

Panel C: Robustness checks						
Dependent variable:	Number of Days		Number of Brokers		Broker HHI	
	(1)	(2)	(3)	(4)	(5)	(6)
$\mathbb{1}$(Informed Motivated)	0.57***	0.36***	0.14***	0.11***	−2.82***	−1.87***
	(4.28)	(7.24)	(7.52)	(9.54)	(−6.25)	(−8.86)
log(Dollar Volume)	0.28***	0.33***	0.17***	0.18***	−3.00***	−3.29***
	(14.16)	(17.51)	(9.33)	(12.22)	(−11.65)	(−16.18)
Time Fixed-effects	Yes	Yes	Yes	Yes	Yes	Yes
Stock Fixed-effects	No	Yes	No	Yes	No	Yes
Manager × Client Fixed-effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	214,250	214,250	214,250	214,250	214,250	214,250
Adjusted R^2	0.23	0.31	0.27	0.31	0.21	0.24

Panel D: Robustness checks						
Dependent variable:	Number of Days		Number of Brokers		Broker HHI	
	(1)	(2)	(3)	(4)	(5)	(6)
$\mathbb{1}$(Informed Motivated)	0.60***	0.48***	0.18***	0.18***	−3.46***	−3.09***
	(3.91)	(4.19)	(3.79)	(3.70)	(−3.88)	(−4.31)
Trading volume, % in CRSP volume	2.97***	1.86***	−0.19***	0.39	−5.44**	−11.10**
	(5.41)	(3.52)	(−3.08)	(1.58)	(−2.11)	(−2.39)
Time Fixed-effects	Yes	Yes	Yes	Yes	Yes	Yes
Stock Fixed-effects	No	Yes	No	Yes	No	Yes
Manager × Client Fixed-effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	214,242	214,242	214,242	214,242	214,242	214,242
Adjusted R^2	0.13	0.17	0.13	0.16	0.12	0.15

Panel E: Robustness checks						
Dependent variable:	Number of Days		Number of Brokers		Broker HHI	
	(1)	(2)	(3)	(4)	(5)	(6)
$\mathbb{1}$(Informed Motivated)	0.60***	0.47***	0.18***	0.17***	−3.34***	−3.04***
	(3.81)	(4.23)	(3.97)	(3.75)	(−4.02)	(−4.39)
Trading volume, % in Outstanding Share	4.50***	2.83***	0.47***	0.90***	−21.96***	−21.28***
	(6.23)	(4.61)	(3.21)	(3.04)	(−4.30)	(−3.65)
Time Fixed-effects	Yes	Yes	Yes	Yes	Yes	Yes
Stock Fixed-effects	No	Yes	No	Yes	No	Yes
Manager × Client Fixed-effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	214,250	214,250	214,250	214,250	214,250	214,250
Adjusted R^2	0.13	0.17	0.13	0.16	0.12	0.15

Table 3: Less Implementation Shortfall with splitting order through Brokers

This table examines whether splitting order strategy can reduce trading costs or not. Specifically, I interact $\textit{Split Brokers}_{i,k,s,t}$ with $\mathbb{1}(\textit{Informed Motivated}_{i,k,s,t})$ in my specification as follows:

$$\begin{aligned}\textit{Implementation Shortfall}_{i,k,s,b,t} &= \delta \times \mathbb{1}(\textit{Informed Motivated}_{i,k,s,t}) \times \textit{Split Brokers}_{i,k,s,t} \\ &+ \beta \times \mathbb{1}(\textit{Informed Motivated}_{i,k,s,t}) + \lambda \times \textit{Split Brokers}_{i,k,s,t} \\ &+ \alpha_{i,k} + \gamma_d + \rho_s + \eta_b + \theta_t + \varepsilon_{i,k,s,b,d,t}\end{aligned}$$

where i indexes managers, k indexes clients, s index stocks, b indexes brokers, d indexes order-sequence and t indexes trade date. The dependent variable $\textit{Implementation Shortfall}_{i,k,s,b,t}$, is the percentage difference between the execution price and a benchmark price. The independent variable of interest, $\mathbb{1}(\textit{Informed Motivated}_{i,k,s,t})$ denotes whether the order is driven by informed reason or not and $\textit{Split Brokers}_{i,k,s,t}$ is one if an institutional investor hires more than one broker, zero otherwise. (Daily Basis) In Panel B, I further control for trading dollar volumes. In Panel C and D, I use sub-sample which includes multi-day order. Depending on the specification, the regression includes manager × client (fund) fixed effects ($\alpha_{i,k}$), stock fixed effects (ρ_s), broker fixed effects (η_b), and time fixed-effects (θ_t). All regressions include order-sequence fixed effects (γ_d). Standard errors are clustered by manager × client and the resulting t-statistics are reported in parentheses. Statistical significance at the 10%, 5%, and 1% level is indicated by *, **, and ***, respectively.

Panel A: Baseline

Dependent variable:	Implementation Shortfall (%)			
	(1)	(2)	(3)	(4)
$\mathbb{1}$(Split Brokers) × $\mathbb{1}$(Informed Motivation)	-0.16^{***}	-0.11^{***}	-0.08^{***}	-0.08^{**}
	(-3.61)	(-3.37)	(-2.65)	(-2.36)
$\mathbb{1}$(Informed Motivation)	0.04^{**}	0.05^{***}	0.03^{***}	0.03^{***}
	(2.56)	(3.66)	(2.80)	(2.77)
$\mathbb{1}$(Split Brokers)	-0.005	-0.004	0.03	0.04
	(-0.09)	(-0.06)	(0.61)	(0.79)
Time Fixed-effects	No	Yes	Yes	Yes
Order Sequence Fixed-effects	Yes	Yes	Yes	Yes
Stock Fixed-effects	No	No	Yes	Yes
Broker Fixed-effects	No	No	No	Yes
Manager × Client Fixed-effects	Yes	Yes	Yes	Yes
Observations	425,219	425,219	425,219	425,219
Adjusted R^2	0.02	0.11	0.13	0.13

Table 3–*Continued*

Panel B: Robustness Checks				
Dependent variable:	Implementation Shortfall (%)			
	(1)	(2)	(3)	(4)
$\mathbb{1}$(Split Brokers) × $\mathbb{1}$(Informed Motivation)	−0.16***	−0.10***	−0.07**	−0.08**
	(−3.52)	(−3.26)	(−2.36)	(−2.25)
$\mathbb{1}$(Informed Motivation)	0.04***	0.05***	0.03**	0.03***
	(2.83)	(4.00)	(2.43)	(2.78)
$\mathbb{1}$(Split Brokers)	−0.01	−0.01	0.03	0.04
	(−0.13)	(−0.10)	(0.64)	(0.80)
log(Dollar Volume)	0.02*	0.02**	0.03***	0.03***
	(1.90)	(1.97)	(3.46)	(2.70)
Time Fixed-effects	No	Yes	Yes	Yes
Order Sequence Fixed-effects	Yes	Yes	Yes	Yes
Stock Fixed-effects	No	No	Yes	Yes
Broker Fixed-effects	No	No	No	Yes
Manager × Client Fixed-effects	Yes	Yes	Yes	Yes
Observations	425,219	425,219	425,219	425,219
Adjusted R^2	0.02	0.11	0.13	0.13

Panel C: Robustness Checks (Multi-Day order)				
Dependent variable:	Implementation Shortfall (%)			
	(1)	(2)	(3)	(4)
$\mathbb{1}$(Split Brokers) × $\mathbb{1}$(Informed Motivation)	−0.19***	−0.14**	−0.12*	−0.12*
	(−2.71)	(−2.24)	(−1.92)	(−1.83)
$\mathbb{1}$(Informed Motivation)	0.06***	0.09***	0.07***	0.06***
	(4.08)	(6.63)	(5.26)	(4.67)
$\mathbb{1}$(Split Brokers)	−0.02	−0.02	0.02	0.02
	(−0.37)	(−0.32)	(0.30)	(0.38)
Time Fixed-effects	No	Yes	Yes	Yes
Order Sequence Fixed-effects	Yes	Yes	Yes	Yes
Stock Fixed-effects	No	No	Yes	Yes
Broker Fixed-effects	No	No	No	Yes
Manager × Client Fixed-effects	Yes	Yes	Yes	Yes
Observations	262,895	262,895	262,895	262,895
Adjusted R^2	0.02	0.12	0.14	0.15

Table 3–*Continued*

Panel D: Robustness Checks (Multi-Day order)				
Dependent variable:	Implementation Shortfall (%)			
	(1)	(2)	(3)	(4)
𝟙(Split Brokers) × 𝟙(Informed Motivation)	−0.19***	−0.14**	−0.12*	−0.12*
	(−2.70)	(−2.22)	(−1.86)	(−1.81)
𝟙(Informed Motivation)	0.06***	0.09***	0.06***	0.06***
	(3.99)	(6.36)	(4.80)	(4.48)
𝟙(Split Brokers)	−0.02	−0.02	0.02	0.02
	(−0.36)	(−0.31)	(0.40)	(0.42)
log(Dollar Volume)	0.01	0.01	0.02***	0.02**
	(1.01)	(1.20)	(3.30)	(2.06)
Time Fixed-effects	No	Yes	Yes	Yes
Order Sequence Fixed-effects	Yes	Yes	Yes	Yes
Stock Fixed-effects	No	No	Yes	Yes
Broker Fixed-effects	No	No	No	Yes
Manager × Client Fixed-effects	Yes	Yes	Yes	Yes
Observations	262,895	262,895	262,895	262,895
Adjusted R^2	0.02	0.12	0.14	0.15

Table 4: Less Implementation Shortfall: Dynamics over Parent Order

This table examines dynamics of institutional investors' usage of brokers and shows sending orders to new brokers can help reduce trading costs. Specifically, I interact $\mathbb{1}(\textit{Informed Motivated}_{i,k,s,t})$ with $\mathbb{1}(\textit{Order Sent to Different Broker}_{i,k,s,t})$ in my specification as follows:

$$\begin{aligned}\textit{Implementation Shortfall}_{i,k,s,b,t} &= \delta \times \mathbb{1}(\textit{Informed Motivated}_{i,k,s,t}) \times \mathbb{1}(\textit{Order Sent to Different Broker}_{i,k,s,t}) \\ &+ \beta \times \mathbb{1}(\textit{Informed Motivated}_{i,k,s,t}) + \lambda \times \mathbb{1}(\textit{Order Sent to Different Broker}_{i,k,s,t}) \\ &+ \alpha_{i,k} + \gamma_d + \rho_s + \eta_b + \theta_t + \varepsilon_{i,k,s,b,d,t}\end{aligned}$$

where i indexes managers, k indexes clients, s index stocks, b indexes brokers, d indexes order-sequence, and t indexes trade date. The dependent variable $\textit{Implementation Shortfall}_{i,k,s,b,t}$, is the percentage difference between the execution price and a benchmark price. The independent variable of interest, $\mathbb{1}(\textit{Informed Motivated}_{i,k,s,t})$ denotes whether the order is driven by informed reason or not and $\mathbb{1}(\textit{Order Sent to Different Broker}_{i,k,s,t})$ is one if an order is sent to different brokers after the first day, zero otherwise. In Panel B, I further control for trading dollar volumes. Depending on the specification, the regression includes manager × client (fund) fixed effects ($\alpha_{i,k}$), stock fixed effects (ρ_s), broker fixed effects (η_b), and time fixed-effects (θ_t). All regressions include order-sequence fixed effects (γ_d). Standard errors are clustered by manager × client and the resulting t-statistics are reported in parentheses. Statistical significance at the 10%, 5%, and 1% level is indicated by *, **, and ***, respectively.

Panel A: Baseline				
Dependent variable:	Implementation Shortfall (%)			
	(1)	(2)	(3)	(4)
𝟙(Order Sent to Different Broker) × 𝟙(Informed Motivation)	−0.06***	−0.05**	−0.05***	−0.05***
	(−2.89)	(−2.46)	(−3.36)	(−3.31)
𝟙(Informed Motivation)	0.01	0.04*	0.04*	0.03*
	(0.48)	(1.92)	(1.85)	(1.91)
𝟙(Order Sent to Different Broker)	0.0003	−0.01	0.003	0.002
	(0.04)	(−1.01)	(0.46)	(0.37)
Time Fixed-effects	No	Yes	Yes	Yes
Order Sequence Fixed-effects	Yes	Yes	Yes	Yes
Stock Fixed-effects	No	No	Yes	Yes
Broker Fixed-effects	No	No	No	Yes
Manager × Client Fixed-effects	Yes	Yes	Yes	Yes
Observations	295,145	295,145	295,145	295,145
Adjusted R^2	0.02	0.13	0.16	0.16

Table 4–*Continued*

Panel B: Robustness Checks				
Dependent variable:	Implementation Shortfall (%)			
	(1)	(2)	(3)	(4)
𝟙(Order Sent to Different Broker) × 𝟙(Informed Motivation)	−0.06***	−0.05**	−0.05***	−0.05***
	(−2.89)	(−2.47)	(−3.34)	(−3.32)
𝟙(Informed Motivation)	0.01	0.05**	0.03*	0.03*
	(0.54)	(2.07)	(1.70)	(1.82)
𝟙(Order Sent to Different Broker)	−0.001	−0.01	0.002	0.003
	(−0.13)	(−1.27)	(0.38)	(0.46)
log(Dollar Volume)	0.01	0.01	0.02***	0.02**
	(1.06)	(1.25)	(3.88)	(2.16)
Time Fixed-effects	No	Yes	Yes	Yes
Order Sequence Fixed-effects	Yes	Yes	Yes	Yes
Stock Fixed-effects	No	No	Yes	Yes
Broker Fixed-effects	No	No	No	Yes
Manager × Client Fixed-effects	Yes	Yes	Yes	Yes
Observations	295,145	295,145	295,145	295,145
Adjusted R^2	0.02	0.13	0.16	0.16

Table 5: Camouflage strategies: Evidence from Confidential Holdings

This table examines whether institutional investors submit more order to their brokers who would execute their informed order. (submitting camouflage orders to brokers) Specifically, I regress $\textit{log(Number of Stocks)}_{i,k,b,t}$ on $\mathbb{1}(\textit{Informed Order to Broker}_{i,k,b,t})$ in my specification as follows:

$$\textit{log(Number of Stocks)}_{i,k,b,t} = \beta \times \mathbb{1}(\textit{Informed Order to Broker}_{i,k,b,t}) + \alpha_{i,k} + \eta_b + \theta_t + \varepsilon_{i,k,b,t}$$

where i indexes managers, k indexes clients, b indexes brokers, and t indexes time. The dependent variable $\textit{log(Number of Stocks)}_{i,k,b,t}$, is the number of stocks which are executed through broker at manager-client-tradedate-broker level. The independent variable of interest, $\mathbb{1}(\textit{Informed Order to Broker}_{i,k,b,t})$ is an indicator variable which is equal to one if a investor submit at least one informed order to a broker, zero otherwise in Panel A. In Panel B, I replace $\textit{log(Number of Stocks)}_{i,k,b,t}$ with $\textit{Stock HHI}_{i,k,b,t}$. $\textit{Stock HHI}_{i,k,b,t}$ is calculated by squaring the portion of the traded aggregate volumes of each stock traded through a broker for investors and then summing the resulting numbers. I control for the number of shares. Depending on the specification, the regression includes includes manager × client (fund) fixed effects ($\alpha_{i,k}$), broker fixed effects (η_b), and time fixed-effects (θ_t). Standard errors are clustered by manager × client and the resulting t-statistics are reported in parentheses. Statistical significance at the 10%, 5%, and 1% level is indicated by *, **, and ***, respectively.

Panel A: Baseline						
Dependent variable:	log(Number of Stocks)					
	(1)	(2)	(3)	(4)	(5)	(6)
$\mathbb{1}$(Informed Order to Broker)	0.68***	0.72***	0.48***	0.55***	0.57***	0.39***
	(3.76)	(3.68)	(9.80)	(6.03)	(5.90)	(19.78)
log(volume)				0.21***	0.21***	0.19***
				(3.88)	(3.85)	(6.37)
Time Fixed-effects	No	Yes	Yes	No	Yes	Yes
Broker Fixed-effects	No	No	Yes	No	No	Yes
Manager × Client Fixed-effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	142,383	142,383	142,383	142,383	142,383	142,383
Adjusted R^2	0.41	0.42	0.55	0.55	0.56	0.65
Panel B: Robustness checks						
Dependent variable:	Stock HHI					
	(1)	(2)	(3)	(4)	(5)	(6)
$\mathbb{1}$(Informed Order to Broker)	−18.78***	−19.80***	−13.82***	−14.56***	−15.16***	−10.88***
	(−4.29)	(−4.22)	(−12.20)	(−8.28)	(−8.15)	(−25.90)
log(volume)				−6.67***	−6.74***	−6.12***
				(−4.85)	(−4.76)	(−8.49)
Time Fixed-effects	No	Yes	Yes	No	Yes	Yes
Broker Fixed-effects	No	No	Yes	No	No	Yes
Manager × Client Fixed-effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	142,383	142,383	142,383	142,383	142,383	142,383
Adjusted R^2	0.36	0.37	0.46	0.47	0.48	0.54

Table 6: Randomizing Order Flows across Brokers: Evidence from Confidential Holdings

This table examines whether the information motivated order is more likely to shuffle the order flows across brokers. Specifically, I regress $Dissimilarity_{i,k,t}$ on $\mathbb{1}(Informed\ Shares > 0.5_{i,k,t})$ in my specification as follows:

$$Dissimilarity_{i,k,t} = \beta \times \mathbb{1}(Informed\ Shares > 0.5_{i,k,t}) + \alpha_{i,k} + \theta_t + \varepsilon_{i,k,t}$$

where i indexes managers, k indexes clients, and t indexes time. The dependent variable $Dissimilarity_{i,k,t}$, is 1−cosine similarity between the set of manager-client's aggregate trading dollar volume through its broker at today and that from 25 days before to 5 days before today in Panel A. As a robustness check in Panel B of Table 6, I construct a dissimilarity measure by focusing top broker for manager-client. To be specific, I take only one broker who executed the largest dollar volume for the manager-client from 25 days before to 5 days before for every day. Then, the proxy is taking how much the manager-client place order through *the* top broker today in terms of the fraction of the aggregate dollar trading volumes. The independent variable of interest, $\mathbb{1}(Informed\ Shares > 0.5_{i,k,t})$ is a dummy variable identifying the day when enough informed trades happen. I control for several measures of order size which are daily aggregate dollar volumes, the number of shares, and the number of stocks. Depending on the specification, the regression includes manager × client (fund) fixed effects ($\alpha_{i,k}$) and time fixed-effects (θ_t). Standard errors are clustered by manager × client and the resulting t-statistics are reported in parentheses. Statistical significance at the 10%, 5%, and 1% level is indicated by *, **, and ***, respectively.

Panel A: Baseline						
Dependent variable:	Dissimilarity × 100 (based on cosine similarity)					
	(1)	(2)	(3)	(4)	(5)	(6)
$\mathbb{1}$(Informed Shares > 0.5)	4.48***	4.88***	3.16***	2.24**	2.31**	2.29**
	(3.83)	(3.83)	(2.79)	(2.04)	(2.12)	(2.10)
log(Dollar Volume)			−2.69***		0.38	
			(−4.69)		(0.66)	
log(volume)						0.51
						(1.03)
log(NStocks)				−7.11***	−7.59***	−7.73***
				(−6.39)	(−7.54)	(−7.00)
Time Fixed-effects	No	Yes	Yes	Yes	Yes	Yes
Manager × Client Fixed-effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	12,625	12,625	12,625	12,625	12,625	12,625
Adjusted R^2	0.43	0.49	0.50	0.52	0.52	0.52

Panel B: Robustness checks						
Dependent variable:	Dissimilarity × 100 (based on overlapped dollar volumes)					
	(1)	(2)	(3)	(4)	(5)	(6)
$\mathbb{1}$(Informed Shares > 0.5)	0.97**	1.57***	1.53***	1.40**	1.45**	1.44**
	(2.09)	(2.62)	(2.65)	(2.53)	(2.57)	(2.57)
log(Dollar Volume)			−0.06		0.24	
			(−0.25)		(1.09)	
log(volume)						0.33
						(1.49)
log(NStocks)				−0.44	−0.74*	−0.84*
				(−1.08)	(−1.94)	(−1.91)
Time Fixed-effects	No	Yes	Yes	Yes	Yes	Yes
Manager × Client Fixed-effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	12,625	12,625	12,625	12,625	12,625	12,625
Adjusted R^2	0.59	0.63	0.63	0.63	0.63	0.63

Table 6–*Continued*

Panel C: Robustness checks						
Dependent variable:	Dissimilarity × 100 (based on Top Broker)					
	(1)	(2)	(3)	(4)	(5)	(6)
𝟙(Informed Shares > 0.5)	2.63**	2.86**	2.76**	2.52**	2.60**	2.58**
	(2.20)	(2.06)	(2.24)	(2.04)	(2.13)	(2.11)
log(Dollar Volume)			−0.15		0.42	
			(−0.21)		(0.62)	
log(volume)						0.55
						(1.05)
log(NStocks)				−0.90	−1.42**	−1.57*
				(−0.82)	(−1.98)	(−1.77)
Time Fixed-effects	No	Yes	Yes	Yes	Yes	Yes
Manager × Client Fixed-effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	12,625	12,625	12,625	12,625	12,625	12,625
Adjusted R^2	0.32	0.38	0.38	0.38	0.38	0.38

Table 7: Randomizing order submission strategies and institutional investors' performance

This table reports the monthly returns (raw-returns and Daniel et al. (1997) benchmark-adjusted returns) from January 1999 to November 2009. I sort funds into quantiles based on similarity based on cosine similarity based on the dollar volume of orders executed by each broker between today's set of brokers used by a fund and the core set of brokers used by the same fund (from day t-25 to day t-6). Then I measure the value-weighted buy-and-hold return on the stocks bought on each day by each fund, weighted by dollar volume, for the next month (typically from day t+1 to day t+21). For each quintile, I first average value-weighted returns across funds then use the time series average value-weighted returns for statistical inference in the spirit of Fama and MacBeth (1973) and adjust for serial correlation up to a lag of 20 trading days following Newey and West (1987). The heteroskedasticity robust t-statistics are reported in parentheses. Statistical significance at the 10%, 5%, and 1% level is indicated by *, **, and ***, respectively.

	cosine similarity					
	Similar(Q_1)	Q_2	Q_3	Q_4	Dis-similar(Q_5)	D - S
Raw-Return	0.76*	0.85*	0.93*	0.98**	1.01**	0.24***
	(1.69)	(1.80)	(1.93)	(2.06)	(2.14)	(3.20)
DGTW-Adjusted Return	0.29***	0.34***	0.38***	0.45***	0.46***	0.17***
	(3.81)	(4.78)	(5.15)	(6.07)	(6.22)	(2.77)

Table 8: Less Implementation Shortfall with more shuffling Order Flows across Brokers

This table examines whether randomization strategy across brokers leads to decrease in trading costs or not. Specifically, I interact $\textit{Dissimilarity}_{i,k,t}$ with $\mathbb{1}(\textit{Informed Motivated}_{i,k,s,t})$ in my specification as follows:

$$\begin{aligned}\textit{Implementation Shortfall}_{i,k,s,b,t} &= \delta \times \mathbb{1}(\textit{Informed Motivated}_{i,k,s,t}) \times \textit{Dissimilarity}_{i,k,t} \\ &+ \beta \times \mathbb{1}(\textit{Informed Motivated}_{i,k,s,t}) + \lambda \times \textit{Dissimilarity}_{i,k,t} \\ &+ \alpha_{i,k} + \rho_s + \eta_b + \theta_t + \varepsilon_{i,k,s,b,t}\end{aligned}$$

where i indexes managers, k indexes clients, s index stocks, b indexes brokers, and t indexes time in day. The dependent variable $\textit{Implementation Shortfall}_{i,k,s,b,t}$, is the percentage difference between the execution price and a benchmark price. The independent variable of interest, $\mathbb{1}(\textit{Informed Motivated}_{i,k,s,t})$ denotes whether the order is driven by informed reason or not and $\textit{Dissimilarity}_{i,k,t}$ is 1−cosine similarity between the set of manager-client (fund)'s aggregate trading dollar volume through its broker at today and the ones from 25 days before to 5 days before today. In columns (4)-(6), I examine block trades which include a trade's volume is greater than or equal to 10,000 or its dollar volume is greater than or equal to $200,000. Panel B presents results when controlling for log(Dollar Volume) or log (volume). In Panel C, I control for trading volume, % in CRSP volume or control for trading volume, % in outstanding share. Depending on the specification, the regression includes manager × client (fund) fixed effects ($\alpha_{i,k}$), stock fixed effects (ρ_s), and broker fixed effects (η_b). All regressions include time fixed-effects (θ_t). Standard errors are clustered by manager × client and the resulting t-statistics are reported in parentheses. Statistical significance at the 10%, 5%, and 1% level is indicated by *, **, and ***, respectively.

Panel A: Baseline						
Dependent variable:	Implementation Shortfall (%)					
	(1)	(2)	(3)	(4)	(5)	(6)
Dissimilarity × 𝟙(Informed Motivation)	−0.13***	−0.12***	−0.12***	−0.41***	−0.36***	−0.37**
	(−2.86)	(−2.58)	(−2.58)	(−2.86)	(−2.64)	(−2.55)
Dissimilarity	0.04	0.02	0.01	0.08	0.03	0.03
	(0.93)	(0.39)	(0.28)	(0.88)	(0.33)	(0.31)
𝟙(Informed Motivation)	0.04***	0.02***	0.02**	0.06***	0.03*	0.03*
	(4.04)	(3.11)	(2.34)	(3.55)	(1.92)	(1.77)
Time Fixed-effects	Yes	Yes	Yes	Yes	Yes	Yes
Stock Fixed-effects	No	Yes	Yes	No	Yes	Yes
Broker Fixed-effects	No	No	Yes	No	No	Yes
Manager × Client Fixed-effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	279,758	279,758	279,758	148,864	148,864	148,864
Adjusted R^2	0.09	0.11	0.11	0.09	0.11	0.11

Table 8–*Continued*

Panel B: Robustness

Dependent variable:	Implementation Shortfall (%)					
	(1)	(2)	(3)	(4)	(5)	(6)
Dissimilarity × 𝟙(Informed Motivation)	-0.13^{***}	-0.11^{**}	-0.11^{**}	-0.12^{***}	-0.11^{**}	-0.11^{**}
	(−2.78)	(−2.35)	(−2.48)	(−2.74)	(−2.36)	(−2.49)
Dissimilarity	0.04	0.01	0.01	0.04	0.01	0.01
	(0.89)	(0.27)	(0.26)	(0.88)	(0.28)	(0.27)
𝟙(Informed Motivation)	0.04^{***}	0.02^{**}	0.02^{**}	0.04^{***}	0.02^{**}	0.02^{**}
	(3.97)	(2.22)	(2.02)	(3.73)	(2.20)	(2.00)
log(Dollar Volume)	0.01^{**}	0.02^{***}	0.02^{**}			
	(1.99)	(6.00)	(2.47)			
log(volume)				0.01^{***}	0.02^{***}	0.01^{**}
				(2.69)	(5.90)	(2.29)
Time Fixed-effects	Yes	Yes	Yes	Yes	Yes	Yes
Stock Fixed-effects	No	Yes	Yes	No	Yes	Yes
Broker Fixed-effects	No	No	Yes	No	No	Yes
Manager × Client Fixed-effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	279,758	279,758	279,758	279,758	279,758	279,758
Adjusted R^2	0.09	0.11	0.11	0.09	0.11	0.11

Panel C: Robustness

Dependent variable:	Implementation Shortfall (%)					
	(1)	(2)	(3)	(4)	(5)	(6)
Dissimilarity × 𝟙(Informed Motivation)	-0.13^{***}	-0.14^{***}	-0.14^{***}	-0.13^{***}	-0.12^{***}	-0.12^{***}
	(−2.86)	(−3.01)	(−3.00)	(−2.73)	(−2.58)	(−2.63)
Dissimilarity	0.04	0.02	0.01	0.04	0.02	0.01
	(0.95)	(0.46)	(0.22)	(0.91)	(0.39)	(0.28)
𝟙(Informed Motivation)	0.04^{***}	0.04^{***}	0.03^{***}	0.04^{***}	0.02^{***}	0.02^{**}
	(3.39)	(4.55)	(3.66)	(3.27)	(3.07)	(2.52)
% in Trading Volume	−0.07	-0.63^{***}	-0.85^{***}			
	(−0.93)	(−4.78)	(−4.75)			
% in Outstanding Share				0.13^{*}	−0.02	−0.18
				(1.67)	(−0.20)	(−1.33)
Time Fixed-effects	Yes	Yes	Yes	Yes	Yes	Yes
Stock Fixed-effects	No	Yes	Yes	No	Yes	Yes
Broker Fixed-effects	No	No	Yes	No	No	Yes
Manager × Client Fixed-effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	279,758	279,758	279,758	279,758	279,758	279,758
Adjusted R^2	0.09	0.11	0.11	0.09	0.11	0.11

Appendix

Table A1: Split or Concentrate Large Orders? Evidence from Order in earlier period of Quarters

This table examines whether the orders in earlier period of quarters tend to split the order across more days and more brokers. Specifically, I regress $\textit{Number of Days}_{i,k,s,t}$ on $\textit{Days to Quarter-end}_{i,k,s,t}$ in my specification as follows:

$$\textit{Number of Days}_{i,k,s,t} = \beta \times \textit{Days to Quarter-end}_{i,k,s,t} + \alpha_{i,k} + \rho_s + \theta_t + \varepsilon_{i,k,s,t}$$

where i indexes managers, k indexes clients, and t indexes the first day in parent order. The dependent variable, $\textit{Number of Days}_{i,k,s,t}$, is the number of days in each parent order. In column (3) through (4), $\textit{Number of Days}_{i,k,s,t}$ is replaced by $\textit{Number of Days}_{i,k,s,t}$, which is the number of brokers in each parent order. In column (5) through (6), the dependent variable is replaced by $\textit{Broker HHI}_{i,k,s,t}$ which is the measure of institutional investors' concentration on usage of brokers. The independent variable of interest, Days to Quarter-end$_{i,k,s,t}$ is the number of days from the first day of order to quarter-end. Panel A presents the baseline results. In Panel B, I examine block trades which include a trade's volume is greater than or equal to 10,000 or its dollar volume is greater than or equal to 200,000. Panel C presents results when controlling for log(Dollar Volume). In Panel D, I control for trading volume, % in CRSP volume and I also control for trading volume, % in outstanding share in Panel E. Depending on the specification, the regression includes manager×client fixed effects ($\alpha_{i,k}$), stock fixed effects (ρ_s) and quarter fixed-effects (θ_t). Standard errors are clustered by manager × client and the resulting t-statistics are reported in parentheses. Statistical significance at the 10%, 5%, and 1% level is indicated by *, **, and ***, respectively.

Panel A: Baseline						
Dependent variable:	Number of Days		Number of Brokers		Broker HHI	
	(1)	(2)	(3)	(4)	(5)	(6)
Days to Quarter-end	0.001***	0.001***	0.001***	0.001***	−0.01***	−0.01***
	(5.08)	(5.15)	(3.07)	(3.11)	(−2.95)	(−2.96)
Quarter Fixed-effects	Yes	Yes	Yes	Yes	Yes	Yes
Stock Fixed-effects	No	Yes	No	Yes	No	Yes
Manager × Client Fixed-effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	10,105,494	10,105,494	10,105,494	10,105,494	10,105,494	10,105,494
Adjusted R^2	0.09	0.10	0.11	0.13	0.10	0.12

Panel B: Block Trade (Volume ≥ 10K or Dollar volume ≥ 200K)						
Dependent variable:	Number of Days		Number of Brokers		Broker HHI	
	(1)	(2)	(3)	(4)	(5)	(6)
Days to Quarter-end	0.003***	0.003***	0.001***	0.001***	−0.02***	−0.02***
	(4.27)	(4.23)	(2.85)	(3.08)	(−3.32)	(−3.65)
Quarter Fixed-effects	Yes	Yes	Yes	Yes	Yes	Yes
Stock Fixed-effects	No	Yes	No	Yes	No	Yes
Manager × Client Fixed-effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3,344,741	3,344,741	3,344,741	3,344,741	3,344,741	3,344,741
Adjusted R^2	0.16	0.18	0.22	0.24	0.17	0.18

Table A1–*Continued*

Panel C: Robustness checks						
Dependent variable:	Number of Days		Number of Brokers		Broker HHI	
	(1)	(2)	(3)	(4)	(5)	(6)
Days to Quarter-end	0.001***	0.001***	0.0004**	0.0004**	−0.01**	−0.01**
	(4.32)	(4.32)	(2.31)	(2.30)	(−2.22)	(−2.20)
log(Dollar volume)	0.20***	0.22***	0.16***	0.16***	−2.92***	−2.88***
	(8.91)	(9.55)	(5.61)	(6.71)	(−6.61)	(−8.67)
Quarter Fixed-effects	Yes	Yes	Yes	Yes	Yes	Yes
Stock Fixed-effects	No	Yes	No	Yes	No	Yes
Manager × Client Fixed-effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	10,105,494	10,105,494	10,105,494	10,105,494	10,105,494	10,105,494
Adjusted R^2	0.20	0.22	0.25	0.25	0.18	0.19

Panel D: Robustness checks						
Dependent variable:	Number of Days		Number of Brokers		Broker HHI	
	(1)	(2)	(3)	(4)	(5)	(6)
Days to Quarter-end	0.001***	0.001***	0.001***	0.001***	−0.01***	−0.01***
	(5.03)	(5.11)	(3.09)	(3.13)	(−2.96)	(−2.97)
Trading volume, % in CRSP volume	0.77	0.90***	−1.27**	−0.18*	29.12**	9.05***
	(1.43)	(3.57)	(−2.57)	(−1.74)	(2.12)	(2.75)
Time Fixed-effects	Yes	Yes	Yes	Yes	Yes	Yes
Stock Fixed-effects	No	Yes	No	Yes	No	Yes
Manager × Client Fixed-effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	10,099,641	10,099,641	10,099,641	10,099,641	10,099,641	10,099,641
Adjusted R^2	0.09	0.10	0.11	0.13	0.10	0.12

Panel E: Robustness checks						
Dependent variable:	Number of Days		Number of Brokers		Broker HHI	
	(1)	(2)	(3)	(4)	(5)	(6)
Days to Quarter-end	0.001***	0.001***	0.001***	0.001***	−0.01***	−0.01***
	(5.01)	(5.09)	(3.09)	(3.10)	(−2.97)	(−2.96)
Trading volume, % in Outstanding Share	16.21***	15.53***	−8.50**	6.11*	229.12	−29.07
	(3.55)	(3.57)	(−2.06)	(1.68)	(1.56)	(−0.75)
Time Fixed-effects	Yes	Yes	Yes	Yes	Yes	Yes
Stock Fixed-effects	No	Yes	No	Yes	No	Yes
Manager × Client Fixed-effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	10,105,494	10,105,494	10,105,494	10,105,494	10,105,494	10,105,494
Adjusted R^2	0.09	0.10	0.11	0.13	0.10	0.12

Table A2: Randomizing Order Flows across Brokers: Evidence from Order in earlier period of Quarters

This table provides further evidence whether the information motivated order is more likely to shuffle the order flows across brokers by looking at orders in earlier periods of quarter. Specifically, I regress $Dissimilarity_{i,k,t}$ on $Days\ to\ Quarter\text{-}end_{i,k,t}$ in my specification as follows:

$$Dissimilarity_{i,k,t} = \beta \times Days\ to\ Quarter\text{-}end_{i,k,t} + \alpha_{i,k} + \theta_t + \varepsilon_{i,k,t}$$

where i indexes managers, k indexes clients, and t indexes time. The dependent variable $Dissimilarity_{i,k,t}$, is 1−cosine similarity between the set of manager-client's aggregate trading dollar volume through its broker at today and that from 25 days before to 5 days before today in Panel A. As a robustness check in Panel B of Table A2, I construct a dissimilarity measure by focusing top broker for manager-client. To be specific, I take only one broker who executed the largest dollar volume for the manager-client from 25 days before to 5 days before for every day. Then, the proxy is taking how much the manager-client place order through *the* top broker today in terms of the fraction of the aggregate dollar trading volumes. The independent variable of interest, $Days\ to\ Quarter\text{-}end_{i,k,t}$ is the number of days from the first day of order to quarter-end. I control for several measures of order size which are daily aggregate dollar volumes, the number of shares, and the number of stocks. Depending on the specification, the regression includes manager × client (fund) fixed effects ($\alpha_{i,k}$), and time fixed-effects (θ_t). Standard errors are clustered by manager × client and the resulting t-statistics are reported in parentheses. Statistical significance at the 10%, 5%, and 1% level is indicated by *, **, and ***, respectively.

Panel A: Baseline

Dependent variable:	Dissimilarity × 100 (based on cosine similarity)					
	(1)	(2)	(3)	(4)	(5)	(6)
Days to Quarter-end	0.01***	0.01***	0.01***	0.01***	0.01***	0.01***
	(7.71)	(8.44)	(11.23)	(9.39)	(9.88)	(9.79)
log(Dollar Volume)			−2.91***		−0.61***	
			(−43.86)		(−7.18)	
log(volume)						−0.64***
						(−7.93)
log(NStocks)				−6.62***	−5.98***	−5.95***
				(−47.59)	(−32.36)	(−33.67)
Time Fixed-effects	No	Yes	Yes	Yes	Yes	Yes
Manager × Client Fixed-effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,607,092	1,607,092	1,607,092	1,607,092	1,607,092	1,607,092
Adjusted R^2	0.25	0.25	0.26	0.28	0.28	0.28

Panel B: Robustness checks

Dependent variable:	Dissimilarity × 100 (based on overlapped dollar volumes)					
	(1)	(2)	(3)	(4)	(5)	(6)
Days to Quarter-end	0.004***	0.004***	0.004***	0.004***	0.004***	0.004***
	(5.63)	(6.35)	(6.57)	(6.38)	(6.55)	(6.45)
log(Dollar Volume)			−0.14***		−0.13***	
			(−4.33)		(−2.78)	
log(volume)						−0.07
						(−1.64)
log(NStocks)				−0.17***	−0.03	−0.09
				(−3.23)	(−0.39)	(−1.25)
Time Fixed-effects	No	Yes	Yes	Yes	Yes	Yes
Manager × Client Fixed-effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,607,092	1,607,092	1,607,092	1,607,092	1,607,092	1,607,092
Adjusted R^2	0.36	0.36	0.36	0.36	0.36	0.36

Table A2–*Continued*

Panel C: Robustness checks						
Dependent variable:	Dissimilarity × 100 (based on Top Broker)					
	(1)	(2)	(3)	(4)	(5)	(6)
Days to Quarter-end	0.01***	0.01***	0.01***	0.01***	0.01***	0.01***
	(7.41)	(7.99)	(8.39)	(8.07)	(8.31)	(8.20)
log(Dollar Volume)			−0.46***		−0.31***	
			(−8.44)		(−4.02)	
log(volume)						−0.23***
						(−3.04)
log(NStocks)				−0.72***	−0.39***	−0.48***
				(−8.05)	(−3.12)	(−3.96)
Time Fixed-effects	No	Yes	Yes	Yes	Yes	Yes
Manager × Client Fixed-effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,607,092	1,607,092	1,607,092	1,607,092	1,607,092	1,607,092
Adjusted R^2	0.18	0.18	0.18	0.18	0.18	0.18