# Institutional Brokerage Networks: Facilitating Liquidity Provision*

Munhee Han    Sanghyun (Hugh) Kim    Vikram K. Nanda

April 29, 2019

*Han (munhee.han@utdallas.edu), Kim (sanghyun.kim1@utdallas.edu), and Nanda (vikram.nanda@utdallas.edu) are at the University of Texas at Dallas, 800 West Campbell Road, Richardson, TX 75080. We thank Steven Xiao, Kelsey Wei, Paul Irvine, Stacey Jacobsen (discussant), Junghoon Lee, Charles Trzcinka, Talis Putnins (discussant), Johan Sulaeman (discussant), Kumar Venkataraman, seminar participants at the University of Texas at Dallas and Texas Christian University, and conference participants at the 29th Annual Conference on Financial Economics and Accounting (CFEA), the 31st Australasian Finance and Banking Conference, Society for Financial Studies (SFS) Cavalcade Asia-Pacific 2018, and the 2019 Midwest Finance Association (MFA) Annual Meeting for helpful comments.

# Institutional Brokerage Networks: Facilitating Liquidity Provision


**Abstract**

We argue institutional brokerage networks facilitate liquidity provision and mitigate price impact of large non-information motivated trades. We use commission payments to map trading networks of mutual-funds and brokers. We find central-funds outperform peripheral-funds, especially in terms of return gap. Outperformance is more pronounced when trading is primarily liquidity driven to accommodate large redemptions. The fund–centrality premium is enhanced by brokers' incentives to generate greater commissions and by trading relationships between brokers and funds. Exploiting large brokerage mergers as exogenous shocks to network structure, we show that shocks to network centrality are accompanied by predicted changes in return gap.



*Keywords:* Institutional brokerage networks, mutual funds, return gap, trading costs, liquidity provision

# 1 Introduction

Brokers play a vital role in institutional trading in equity markets. When executing large client orders, brokers can mitigate price impact by actively searching for potential counterparties across various trading venues and, on occasion, by committing their own capital and acting more as dealers. Brokers often break up their clients' large orders and then strategically reveal to other clients who may be willing to fill the orders, while concealing from those who might front-run them (see Harris (2002) for an overview). Thus, trading between institutional investors tends to be broker-intermediated, with its efficacy closely tied to the trading networks of institutional investors and their brokers. In this paper, we argue that institutional brokerage networks facilitate liquidity provision and mitigate price impact for non-information driven trades.

Using brokerage commission payments, we map trading networks of mutual funds and their brokers as affiliation networks in which mutual funds are connected through their overlapping brokerage relationships. In these networks, mutual funds that trade through brokers that are also heavily used by other funds will tend to be more central. A key finding of the paper is that central funds in institutional brokerage networks outperform peripheral funds, especially as measured by their trading performance. In order to shed light on the specific mechanisms driving the positive relation between mutual funds' brokerage network centrality and their trading performance i.e., *fund–centrality premium*, we propose a liquidity provision hypothesis.

Our notion is that centrality in brokerage networks is especially valuable when mutual funds are forced to trade for liquidity reasons. As is well-recognized, open-end mutual funds incur substantial trading costs due to the adverse market impact of their trades when they liquidate holdings in response to investor redemptions (e.g., Edelen (1999)). In market microstructure models (such as Glosten and Milgrom (1985), Kyle (1985)), risk-neutral market makers are unable to identify trading motives. In these models, market makers set market prices and expect to lose to informed traders, while breaking even with gains from uninformed, liquidity traders. Thus, as in Admati and Pfleiderer (1991), liquidity traders who are transacting large quantities for non-informational reasons have an incentive to make their trading intentions known (i.e., engage in "sunshine trading") to distinguish themselves from informed traders and attract more traders to provide liquidity.[1] While large liquidity traders may be unable to signal their trading motives directly to market participants, our view is that they might achieve the desired outcome by relying on their brokerage network and upstairs block trading.

We contend that institutions trading for liquidity reasons may be able to credibly convey their trading motives to brokers with whom they have well-established relationships. The credibility of a mutual fund will be enhanced if misrepresentation of its trading motives is likely to be costly in terms of a loss of reputation capital and trust in the broker-institution relationship. Central funds, connected to a larger network of brokers and funds will have more at risk in terms of potential loss of reputation and, hence, are likely to have greater credibility. The fund's brokers, in turn, could certify their clients' liquidity motives and execute trades at better prices (Seppi (1990)).[2] In addition, upstairs brokers can expand the available liquidity pool using information about their clients' latent trading interests and reaching out to wider set of potential counterparties to lower trading costs (Grossman (1992)).[3] Thus, even though all funds may have similar access to the available pool of expressed liquidity, for instance, through an electronic limit order book in the downstairs market, central funds will be better positioned to tap into larger pools of unexpressed liquidity through their brokers, especially when submitting large blocks of liquidity-motivated orders.[4]

[1] A concern, however, is that strategic traders that become aware of, say, a large liquidation could engage in "predatory trading", an argument advanced in Brunnermeier and Pedersen (2005). The notion is that strategic traders would trade the asset in the same direction prior to or simultaneously with the liquidating trader, before subsequently reversing the trade, to profit from the price impact at the expense of the liquidating trader. Bessembinder et al. (2016), however, show that traders supply liquidity to rather than exploit predictable trades in resilient markets and provide empirical evidence that a larger number of individual trading accounts provide liquidity around the time of large and predictable futures "roll" trades undertaken by a large exchange-traded fund (ETF) designed to provide returns that track crude oil prices.

[2] An upstairs market is an off-exchange market where a block broker facilitates the trading process by locating counterparties to the trade, and it operates as a search-brokerage mechanism where the terms of trade are determined through negotiation. Madhavan and Cheng (1997), Smith, Turnbull, and White (2001), and Booth et al. (2002) present evidence consistent with the Seppi (1990) hypothesis that upstairs market makers effectively screen out information-motivated orders and execute large liquidity-motivated orders at a lower cost than the downstairs market in the New York Stock Exchange (NYSE), the Toronto Stock Exchange (TSE), and the Helsinki Stock Exchange (HSE), respectively.

[3] Bessembinder and Venkataraman (2004) present direct evidence in support of the Grossman (1992) prediction that upstairs brokers lower execution costs by tapping into unexpressed liquidity. The authors find that execution costs for upstairs trades on the Paris Bourse are much lower than would be expected if the trades were executed against the expressed (displayed and hidden) liquidity in the downstairs limit order book.

[4] In a related literature on inter-dealer networks in the over-the-counter (OTC) municipal bond market, Li and Schürhoff (Forthcoming) find that dealers that are more central in the networks have better access to clients and more

Liquidity traders may have their own concerns about revealing their trading intentions to brokers. In the context of outflow-driven fire sales, Barbon et al. (Forthcoming) document that institutional brokers appear to foster predatory trading by leaking their clients' order flow information about impending fire sales to other important clients. These clients then sell the stocks being liquidated – only to buy them back later at lower prices. Our view is that while brokers may occasionally disclose client trades, they are unlikely to do so against their important clients, if it puts their trading relationships in jeopardy. Institutional investors would share trading intentions only if brokerage firms had valuable reputation capital: capital that could be lost if brokers did not act in their clients' interests. A broker disclosing client information faces the risk of being readily detected due to, for instance, the visibility of the price impacts (see, e.g., Smith, Turnbull, and White (2001)). In a broader context, our contention is that brokers will tend to use information about large liquidity-motivated orders to mitigate trading costs associated with adverse selection and invite more traders to provide liquidity, especially when the brokers' reputation costs are sufficiently high. The brokers used by central funds are apt to have greater reputation capital as indicated, for instance, by their well-established relationships to many other funds (and greater costs to being seen as untrustworthy). Hence, central funds are likely to benefit from lower costs for their liquidity motivated trades.[5]

To test our liquidity provision hypothesis, we exploit a unique dataset on brokerage commissions for a comprehensive sample of mutual funds from Form N-SAR semi-annual reports filed with the Securities and Exchange Commission (SEC). Using techniques from graph theory, we map the connections between mutual funds and their brokers as affiliation networks represented by weighted bi-partite graphs.[6] The

information about which securities are available and who wants to buy or sell, which results in shorter "intermediation chains," i.e., that fewer dealers are involved before a bond is transferred to another customer.

[5] Our paper is complementary to Barbon et al. (Forthcoming) in the sense of Carlin, Lobo, and Viswanathan (2007), who present a multi-period model of trading based on liquidity needs. In their model, traders cooperate most of the time through repeated interaction, providing liquidity to one another. However, "episodically" this cooperation breaks down when the stakes are high enough, leading to predatory trading.

[6] In affiliation networks, members are connected with one another through the organizations to which they belong. One can imagine, for instance, how movie stars are connected to one another through the movies in which they have co-appeared. Affiliation networks can be represented by bi-partite graphs, which have two types of nodes with one node of one type only connected to another node of a different type. In our case, a mutual fund is directly connected to its brokers and any pair of mutual funds can be connected with each other only indirectly through their overlapping brokerage connections. The connection between two funds is stronger if the extent to which their

weight of the bi-partite graph represents the strength of connection between a given fund-broker pair and is calculated as a fraction of brokerage commissions paid to the given broker. Further, to measure mutual funds' brokerage network centrality, we reduce this bi-partite graph of funds and brokers into a mono-partite graph in which fund-to-fund links are operationalized through their overlapping broker ties. We then use degree centrality and eigenvector centrality to quantify the importance of a given fund's position in the network.

Mutual funds that trade through many brokers that many other funds also trade through tend to be central in the network. Goldstein et al. (2009) note that most institutions concentrate their order flows with a small number of brokers in order to become their important clients, whereas large institutions can easily obtain the premium status from most brokers. Consistent with this observation, we find that funds that are large or belong to large fund families tend to be more central in the network, as they can afford to trade through a large number of brokers that are themselves central in the network.[7] We also find that mutual funds' brokerage network centrality is highly persistent, reflecting the persistence in the underlying brokerage relationships.

We begin our empirical analysis by showing that mutual funds' brokerage network centrality positively predicts their trading performance. Since we do not directly observe trading activities of mutual funds, we use as our measure of trading performance the return gap, which is calculated as the difference between the reported fund return and the return on a hypothetical portfolio that invests in the previously disclosed fund holdings (Grinblatt and Titman (1989), Kacperczyk, Sialm, and Zheng (2008)). We find that mutual funds in the highest quintile of brokerage network centrality have average monthly return gaps that are about five basis points larger than mutual funds in the lowest quintile over the period from July 1994 to December 2016. The results are statistically significant, insensitive to the choice of centrality measures, and robust to risk adjustments.

The economic magnitude of the relation between brokerage network centrality and return gap is

brokerage connections overlap is larger.

[7] However, other fund characteristics do not explain much of variation in brokerage network centrality. In contrast, fixed-effects, especially fund fixed-effects, account for a large amount of variation in brokerage network centrality, suggesting that we can identify the network effects that are orthogonal to the size effects.

meaningful as well. To put the numbers in perspective, we find that the return gap differential between the highest and lowest quintile portfolios sorted on brokerage network centrality is nearly half as large as that sorted on past return gap (Kacperczyk, Sialm, and Zheng (2008)). Furthermore, in our sub-sample analysis, we find that the fund–centrality premium is economically large and statistically significant in both early (1994-2007) and later (2008-2016) periods. This suggests that even in today's fragmented market with dark pools and smart order-routing systems, upstairs trading and institutional brokerage networks remain highly relevant to large institutional investors, as reported in the *Wall Street Journal*.[8]

In order to understand the specific mechanisms driving the return–gap premium associated with mutual funds' brokerage network centrality, it is useful to recognize key factors affecting the return gap. The return gap was originally proposed by Grinblatt and Titman (1989) as a measure of total transactions costs for mutual funds. Thus, at first brush, the fund–centrality premium is pretty much in line with our hypothesis that institutional brokerage networks mitigate mutual fund trading costs. Grinblatt and Titman (1989), however, point out that the return gap may be affected by interim trades within a quarter (Puckett and Yan (2011)) and possibly window-dressing activities. Kacperczyk, Sialm, and Zheng (2008) further note that skilled fund managers can use their informational advantage to time the trades of individual stocks optimally and show that the past return gap helps predict fund performance.[9]

We also recognize that the network formation is likely endogenous.[10] In order to rule out potential

[8] "'Upstairs' Trading Draws More Big Investors," by Bradley Hope, the *Wall Street Journal*, December 8, 2013. The article quotes a trader as stating that "It's like trying to fill up your gas tank, but you have to go to 15 gas stations. By the time you get to the 15th one, they've increased the price because they've heard you were coming. Wouldn't someone rather go to two or three stations and fill up the tank in blocks?"

[9] It may seem plausible as an alternative hypothesis that central funds can acquire privileged information about company fundamentals through their strong brokerage connections and trade on it. Put it differently, under the information channel hypothesis, the positive relation between brokerage network centrality and return gap could be driven by interim trades within a quarter, rather than trading costs. As we will show in our subsequent analyses, however, the fund–centrality premium is more pronounced when funds' trading activities are largely driven by liquidity reasons, rather than information motivated.

[10] For instance, marginal benefits of brokerage networks are likely higher for better skilled ones, fund managers with superior trading skills might self-select into central positions in institutional brokerage networks. There might exist an unobservable (to the econometrician) factor that is correlated with both brokerage network centrality and return gap. For instance, Anand et al. (2012) show that institutional trading costs are closely linked to trading desks' execution skills over and above selecting better brokers. In Section 5, we provide evidence supportive of our causal interpretation that institutional brokerage networks *improve* institutional trading performance, by exploiting mergers of large brokerage houses as plausibly exogenous shocks to the network structure.

confounding factors, we use panel regressions with fund fixed-effects to control for fund characteristics and unobserved heterogeneity. Consistent with our time-series results, we continue to find robust evidence that brokerage network centrality positively predicts future return gap, even after controlling for fund characteristics, including past return gap, and fund fixed-effects.

Now we turn to testing key predictions of our liquidity provision hypothesis. The primary prediction that we can derive from our hypothesis is that the fund–centrality premium should be more pronounced when funds' trading activities are largely driven by liquidity motives and funds can credibly signal this to their brokers. We use large outflow events to identify such periods of liquidity-motivated trading. When a mutual fund is experiencing severe redemptions, the fund is forced to liquidate a large fraction of its holdings in several stocks and their selling is, to a large extent, uninformed (see, e.g., Coval and Stafford (2007), Alexander, Cici, and Gibson (2007)). In addition, such forced liquidations are likely to send a particularly strong signal to the brokers that its sell orders are driven by liquidity reasons, rather than information motivated, thus helping the brokers communicate more credibly with other institutional clients to take the other end of the trades. Consistent with this prediction, we find that the fund–centrality premium is more pronounced when funds are forced to unwind their positions to accommodate large outflows.[11]

Second, our liquidity provision hypothesis also requires an active role on the part of brokers, such as in discerning their clients' uninformed trading motives and communicating with other institutional clients. As made clear in Carlin, Lobo, and Viswanathan (2007), whether the brokers facilitate liquidity provision or foster predatory trading is likely to hinge on the incentives they face and the strength of repeated interaction with their clients. To the extent that brokers are incentivized to maximize the expected value of future commission revenue streams, central funds with greater commission revenue generating potential are most likely to benefit from liquidity provision facilitated by their brokers. Using aggregate

[11] One potential concern is that the above results could be also consistent with cross-subsidization within a fund family: when a fund is suffering severe redemptions, another fund in the same family could step in to provide liquidity. For instance, Bhattacharya, Lee, and Pool (2013) show that affiliated funds of mutual funds that invest only in other funds within the family provide an insurance pool against temporary liquidity shocks to other funds in the family. This alternative cross-subsidization hypothesis may seem plausible because we find that funds that belong to large families are more central and large fund families are likely better equipped to provide cross-subsidization. Nevertheless, we continue to find qualitatively similar results when we exclude funds that belong to large fund families.

brokerage commissions as a proxy for the broker's incentives, we find that the fund–centrality premium is more pronounced for the funds that are likely more valuable for the brokers. Furthermore, we find that the effect of brokers' incentives on the fund–centrality premium is further amplified when funds are experiencing severe investor redemptions.

Third, our hypothesis relies on the repeated nature of interaction between institutional clients and their brokers. Institutional investors must build reputation for being truthful in order to credibly signal liquidity motives for their uninformed orders to their brokers. The brokers, in turn, must develop their reputation capital for being discreet when handling their clients' orders. Thus, the signaling and certification of uninformed trading motives is likely most effective if funds have already built strong trading relationships their brokers. Consistent with this prediction, we find that the fund–centrality premium is larger for the clients that have stronger existing trading relationships with their brokers, especially when funds are forced to liquidate to accommodate large outflows.[12]

One could still argue that central funds can obtain the return–gap premium because central funds can more easily slice up large orders and spread across many brokers who can then further split their clients' orders across many counterparties. Although not mutually exclusive with this alternative hypothesis, our liquidity provision hypothesis has clear predictions about the relation between the fund–centrality premium and the information content of trading. We provide further evidence that the fund–centrality premium is mostly concentrated in the periods that can be characterized by uninformed trading activities, e.g., when funds are trading *with flows*, rather than *against flows*. In addition, the fund–centrality premium is further amplified when the orders are also likely larger, suggesting that central funds can obtain the return–gap premium when central funds submit large *uninformed* orders.

Before concluding, we provide evidence supportive of our causal interpretation that institutional brokerage networks *improve* institutional trading performance, by exploiting mergers of large brokerage houses as plausibly exogenous shocks to the network structure. Following Hong and Kacperczyk (2010),

[12] This result is also consistent with that found in a related literature on client-dealer networks. For instance, Di Maggio, Kermani, and Song (2017) show that prior trading relationships are valuable especially in turbulent times in the OTC corporate bond market

we are able to identify and match a total of 26 brokerage mergers with our N–SAR data during the period from 1995 to 2015. The shock strength, however, is a major concern for our natural experiment, given the complexity of our network structure (which typically consists of thousands of nodes connected by tens of thousands of edges). In other words, moderate-sized brokerage mergers, especially as stand-alone events (which amount to cutting a small number of edges connected to a single node), are unlikely to serve as meaningful shocks. Therefore, we focus on two waves of five largest mergers of institutional brokers that took place around 2000 and 2008.[13]

Another challenge for our natural experiment is that the treatment of a shock is *a priori* unclear. However, we can reason that funds that traded largely through the acquiring brokers but not heavily though the target brokers are most likely to benefit from exogenous shocks to the network, since the acquiring broker would retain at least some of the target broker's clients. Following this intuition, we first construct hypothetical post-merger brokerage networks as would be formed if every fund were to maintain its pre-merger brokerage relationships and the funds hiring target brokers were to simply redistribute commissions to their remaining brokers on a pro-rata basis.[14] We then estimate the expected change in brokerage network centrality for each fund by calculating the difference between its hypothetical post-merger network centrality and its actual pre-merger network centrality. We take top ten percent of funds with largest expected change as the treatment group. Using a difference-in-differences (DiD) with matching, we find that funds in the treatment group experience significant increases in both brokerage network centrality and return gap after the merger relative to a control group of funds. These findings provide plausible evidence that institutional brokerage networks have a causal impact on institutional trading performance.

The remainder of this paper is organized as follows. In the next section, we discuss our paper in the context of related literature. Section 3 introduces our data and describes how we construct networks. We report our main results in Section 4 and conduct a natural experiment in Section 5. Section 6 concludes.

[13] These five brokerage mergers include Credit Suisse First Boston (CFBS)' acquisition of Donaldson, Lufkin & Jenrette (DLJ) and UBS's acquisition of Paine Webber in 2000 and JP Morgan Chase's acquisition of Bear Stearns, Barclays' acquisition of Lehman Brothers, and Bank of America's acquisition of Merrill Lynch in 2008.

[14] In our status-quo assumption, the funds that did not trade through the target broker (candidate treated funds) do not change their brokerage relationships, as they don't need to, but nonetheless experience exogenous increases in brokerage network centrality after the merger, because *other* funds need to reconfigure their brokerage relationships.

# 2 Related Literature

Our paper uncovers novel network effects in equity markets by documenting the return–gap premium associated with mutual funds' brokerage network centrality.[15] We contribute to a growing literature on broker-dealer networks in financial markets by shedding light on a unique role of institutional brokers in facilitating liquidity provision through the network. Whereas there is a large literature on dealer networks in over-the-counter (OTC) markets (see Section V. D of Bessembinder, Spatt, and Venkataraman (Forthcoming) for a comprehensive survey), studies on broker networks in the stock market have been relatively scant and our paper attempts to fill this gap.

In a recent paper, Di Maggio et al. (Forthcoming) shows that central brokers can extrapolate large informed trades from order flows and selectively leak this information to their more important clients, thereby facilitating "back-running" as described by Yang and Zhu (Forthcoming). Given such rent-extraction behavior, it is thus unclear whether central brokers can obtain "best execution" for their institutional clients. Our paper shows that central funds that trade through *many* central brokers can obtain the return–gap premium by effectively leveraging their strong brokerage connections to mitigate trading costs associated with adverse selection. Our paper is consistent with a related literature on client-dealer networks in the OTC corporate bond market. Hendershott et al. (2017) shows that many insurers use only one dealer, but execution costs decrease as a non-monotone function of the network size until it reaches 20 dealers, consistent with insurers trading off the benefits of relationship trading against dealer competition.

Our paper is related to, but differs from, recent studies that document evidence of information flows or leakages from some clients to the others through the brokers. Chung and Kang (2016) shows strong return comovement among hedge funds sharing the same prime broker and argue that the prime broker provides profitable information to its hedge fund clients. As potential sources of such profitable information, Kumar et al. (2018) points to privileged information on corporate borrowers from the affiliated

[15] There has been a growing interest in studying network effects in equity markets. For instance, Ahern (2013) shows that industries that are more central in intersectoral trade networks earn higher stock returns than industries that are less central. Ozsoylev et al. (2014) estimate empirical investor networks using account-level trading data from the Istanbul Stock Exchange and find that more central individual investors earn higher returns and trade earlier than peripheral investors with respect to information events.

banking division of an investment bank with prime brokerage business and Di Maggio et al. (Forthcoming) hints at client order flow information about large informed trades by hedge funds or activist investors right before 13D filings. Our paper, however, differs substantially from these papers in that our focus is on information flows regarding large liquidity-motivated trades, rather than private information about company fundamentals.[16]

Our paper is most closely related to Barbon et al. (Forthcoming) who document that institutional brokers can foster predatory trading by leaking their clients' order flow information about impending fire sales to other important clients, such as prime brokerage hedge fund clients. The clients then sell the stocks being liquidated along with the distressed funds only to buy them back later at much lower prices, thereby exacerbating price impacts. Brokerage firms, however, value their reputation capital and institutional clients can easily monitor whether a particular broker is acting in their interests thanks to the visibility of the price impacts and the ongoing broker–client relationships (see, e.g., Smith, Turnbull, and White (2001)).

In a broader context, we show that brokers tend to use information about large liquidity-motivated orders to mitigate trading costs associated with adverse selection and invite more traders to provide liquidity, especially when the brokers' reputation costs are sufficiently high. Our paper is complementary to Barbon et al. (Forthcoming) in the sense of Carlin, Lobo, and Viswanathan (2007), who present a multi-period model of trading based on liquidity needs. In their model, traders cooperate most of the time through repeated interaction, providing liquidity to one another. However, "episodically" this cooperation breaks down when the stakes are high enough, leading to predatory trading.[17]

[16] In a similar sense, our paper differs from the papers that shows how institutional investors can gain informational advantage through their brokerage connections. Examples of such information channels include early access to sell-side research or tipping (Irvine, Lipson, and Puckett (2007)) and invitations to broker-hosted investor conferences (Green et al. (2014)).

[17] Some investment banks generate a substantial amount of fee revenues from hedge funds that use their prime brokerage services, such as securities lending, margin financing, and risk management. Consistent with high-powered incentives of prime brokerage business, Kumar et al. (2018) find strong evidence that investment banks sometimes leak privileged information about their corporate borrowers to their prime brokerage hedge fund clients who subsequently trade on and profit from it, whereas Griffin, Shu, and Topaloglu (2012) find little evidence of such information-based trading by the average brokerage house client of investment banks.

# 3 Data and Variable Construction

Section 3.1 describes our primary data on brokerage commissions and explains how we construct other fund-level variables. Section 3.2 explains how we construct institutional brokerage networks and centrality measures, discusses the characteristics of the network, and examines the determinants of mutual funds' brokerage network centrality.

## 3.1 Brokerage Commissions and Other Fund-Level Variables

Our primary data comes from the SEC Form N–SAR filings, which we combine with other data sets. We obtain data on mutual fund monthly returns, total net assets (TNA), and fund expenses from the Center for Research in Security Prices (CRSP) Survivor-Bias-Free Mutual Fund Database. The returns are net of fees, expenses, and brokerage commissions, but before any front-end or back-end loads. The stock holdings of mutual funds are from Thomson-Reuter Ownership Database (Thomson s12). We use the MFLINKS files available through Wharton Research Data Services (WRDS) to merge CRSP and Thomson data sets. For funds with multiple share classes in CRSP, we aggregate share-class-level variables at the fund-level by computing the sum of total net assets and the value-weighted average of returns and expenses.

Under the Investment Company Act of 1940, all registered investment companies are required to file Form N–SAR with the SEC on a semi-annual basis. N–SAR reports are filed at the registrant level. A registrant typically consists of a single mutual fund and thus is simply referred to as a fund in our paper, except when the distinction is likely important.[18] N–SAR filings disclose information about fund operations and financials under 133 numbered items with alphabetized sub-items. We extract all N-SAR reports filed between 1994 and 2016 available through the SEC's Electronic Data Gathering, Analysis, and Retrieval (EDGAR) system.

[18] A registrant can consist of multiple funds or be part of a fund family, although it is just a single mutual fund in about 65% of the N–SAR filings. We emphasize that a registrant does not refer to a fund family, but rather is a filing unit under which a fund family reports its funds together in a single filing. For instance, according to our N–SAR data, Fidelity reported its 466 mutual funds with about $1.5 trillion assets under management using 82 separate N–SAR filings during the first half of 2016. Many items are reported at the fund level, but some of the items such as brokerage commissions are aggregated and reported at the registrant level.

Since our focus is on U.S. domestic equity funds, we exclude N-SAR funds that are not equity-oriented (Item 66.A), international funds (Item 68.B), and the funds with percentage of TNA invested in common stocks (Item 74.F divided by Item 74.T) below 80% or above 105%. We also exclude N-SAR reports where aggregate brokerage commissions paid (Item 21) are reported as zero or missing.[19] From the CRSP–Thomson merged data set, we eliminate international, municipal, bonds and preferred, and metals funds using the investment objective code from Thomson (ioc) and screen for U.S. domestic equity funds using the investment objective code from CRSP (crsp_obj_cd). We also exclude all observations where the fund's TNA does not exceed $5 million or the number of stock holdings does not exceed 10.

After the above data screens, we automatically match N–SAR fund names (Item 1.A and a colon followed by Item 7.C) with CRSP fund names after removing share-class identifiers using the generalized Levenshtein (1966) edit distance while exploiting the typical structure of CRSP fund name (*FUND FAMILY NAME: FUND NAME; SHARE CLASS*). In the automated name matching process, we require that the monthly average net assets (TNA) during the reporting period (Item 75.B) and the corresponding TNA value constructed from CRSP and MFLINKS be within the 5% range from each other. Finally, we manually check the accuracy of the matches and remove the ones that appear inaccurate. The total number and aggregate TNA of our CRSP–Thomson–NSAR matched sample funds are reported in Table A1 in the Appendix.

Of particular interest to our study are brokerage commissions paid to the ten brokers that received the largest amount from the fund during the reporting period and the names of those brokers (Item 20). Table 1 provides an example of brokerage commission payments along with some descriptive statistics.

[Insert Table 1]

We recognize that N–SAR filings do not report all brokerage firms to which the fund paid commissions and, as a result, we are likely to miss some of the less important brokerage connections. As an example, Panel A of Table 1 reports brokerage commissions that T. Rowe Price Blue Chip Growth Fund paid to

[19] Reuter (2006) reports that in his sample, approximately 82% of the N-SAR filings that report paying no brokerage commissions are from investment companies that consist solely of bond funds, which do not pay explicit brokerage commissions on their transactions.

its top ten brokers and the aggregate commissions paid to all brokers during the first half of 2016. As is typically the case, the sum of brokerage commissions do not add up to the aggregate commissions, suggesting that the fund employed more than ten brokers.[20] In general, as shown in Panel B of Table 1, brokerage commissions are highly concentrated with a few important brokers for each fund, but the top ten brokers reported in N–SAR filings on average account for only 72.45% (or 71.62% at the median) of the aggregate brokerage commissions that the fund paid to *all* brokers. Nevertheless, partial data issues are unlikely to cause any bias in our results, since centrality calculated in the reduced network is highly correlated with full-network centrality (Ozsoylev et al. (2014)).[21]

Panel C presents a transition probability matrix of annual changes in broker rankings for each fund and shows strong persistence in brokerage relationships between a fund and its key brokers. If a broker is ranked top this year by the commission payments, the probability of the same broker staying on top for the same fund next year is close to 50%. As we move down the rankings, the persistence becomes gradually weaker. The concentration of commissions with several important brokers and the persistence in business relationships funds maintain with those brokers are generally in line with the literature on institutional brokers (e.g., Goldstein et al. (2009)).

Next, we describe how we construct other fund-level variables. We take the fund TNA directly from N–SAR (Item 74.T) and use the fund family code reported by the fund (Item 19.C) to calculate the fund family TNA. The trading volume is calculated by the sum of purchases (Item 71.A) and sales (Item 71.B). Since brokerage commissions are reported at the registrant level, we calculate the commission rate as a ratio of the aggregate commission payments (Item 21) to the sum of aggregate trading volumes across all funds reported together, following Edelen, Evans, and Kadlec (2012). This pro-rata algorithm implicitly assumes

[20] Mutual funds and institutional investors typically employ a large number of brokers not only for daily trade executions but also for various services that brokers provide, such as early access to sell-side research or tipping (Irvine, Lipson, and Puckett (2007)), favorable allocations of hot IPO stocks (Reuter (2006)), invitations to broker-hosted investor conferences (Green et al. (2014)), and marketing and retail distribution support (Edelen, Evans, and Kadlec (2012)).

[21] For example, in simulations Ozsoylev et al. (2014) show that even when a reduced network represents only 10% of the links in the full network, the correlation between true centrality and centrality calculated in the reduced network is about 0.5. In our study, the reduced network typically represents more than 70% of the weighted links in the full network that be constructed from the complete information on commissions paid to all brokers.

the commission rates to be the same for all the funds of which a registrant consists. In a similar spirit, we estimate the fund's commission payments by taking the product of the commission rate and the fund trading volume. We take an index fund indicator from N–SAR (Item 69). For each fund-quarter, size, value, and momentum percentiles are calculated as percentiles of market capitalization, book-to-market ratio, and 12-month returns skipping the most recent month, respectively, averaged across all stock holdings. For each fund-halfyear, we take the most recent quarterly observation of average size-value-momentum percentiles.

Last, following the literature (e.g., Coval and Stafford (2007)), we calculate monthly net flows for each fund share class $i$ during month $t$ as follows:

$$FLOW_{i,t} = TNA_{i,t} - TNA_{i,t-1} \times (1 + R_{i,t}) \tag{1}$$

where $FLOW_{i,t}$ is the dollar value of fund flow (net new issues and redemptions), $TNA_{i,t}$ is the total net asset, and $R_{i,t}$ is the monthly return. To compute the monthly fund flow for the fund, we sum monthly fund flows for all share classes belonging to the same fund as identified by MFLINKS. Monthly fund flows are summed over the half-year to calculate the semi-annual fund flow. For the percentage figures, we divide the dollar value of fund flows by the beginning-of-period $TNA$. The summary statistics are reported in Table 2.

[Insert Table 2]

## 3.2 Institutional Brokerage Networks

Using brokerage commission payments, we map trading networks of mutual funds and their brokers as affiliation networks represented by weighted bi-partite graphs. In a graph, agents can be represented by nodes and connections (ties) between agents by edges. In a bi-partite graph, nodes can be partitioned into two types and nodes of one type can only be connected to the nodes of the other type, not with the ones of the same type. Such bi-partite graphs are typically used to model affiliation networks where members form networks through organizations to which they belong. We illustrate how we construct institutional brokerage networks and calculate brokerage network centrality step-by-step using a simple example in

Figure 1. Panel A of Figure 1 presents a graphical representation of the network consisting of ten funds and four brokers.

[Insert Figure 1]

Like any graph, a bi-partite graph can be represented by an adjacency matrix, denoted $G$, where rows index mutual funds and columns index brokers. Each element $g_{i,k}$ of $G$ represents the strength of connection between fund $i$ and broker $k$ and is defined as the brokerage commissions paid to broker $k$, scaled by the sum of brokerage commissions paid to the top ten brokers. If broker $k$ does not appear as one of the top ten brokers for fund $i$, then $g_{i,k}$ is assumed zero. Panel B of Figure 1 shows the transpose of the adjacency matrix $G$ representing our simple network in extended-form.

To measure a mutual fund's connections to all the other mutual funds through their overlapping brokerage connections, we reduce the bi-partite graph of mutual funds and brokers into a mono-partite graph of mutual funds only by defining its adjacency matrix $A$ as

$$a_{i,j} = \sum_{k} \min(g_{i,k}, g_{j,k}) \text{ if } i \neq j \tag{2}$$

where $i$ and $j$ index funds and $k$ indexes brokers. The strength of connection between any pair of funds is simply the percentage overlap (Jaccard distance) of brokerage connections between two funds. Panel C of Figure 1 shows the adjacency matrix $A$ representing our simple network in reduced-form. We emphasize that the connections between funds are *indirect* and made through overlapping brokerage connections. For instance, Fund 1 and Fund 2 are connected through Broker $A$ and Broker $B$, and the strength of connection between these two funds is 0.35 $(= \min(0.85, 0.20) + \min(0.15, 0.33) + \min(0, 0.22) + \min(0, 0.25))$.

We borrow techniques from graph theory and social network literature to quantify the importance of a mutual fund's position in the network. The importance of a node in a network is typically measured by its centrality and we use degree centrality (Freeman (1979)) and eigenvector centrality (Bonacich (1972, 1987)).[22] Degree centrality is defined as the sum of each row in the adjacency matrix, $A$, defining the

[22] Many different measures of centrality have been proposed and among the most commonly used measures

network, scaled by the number of rows minus one. Eigenvector centrality is defined as the principal eigenvector of the adjacency matrix defining the network. That is,

$$\lambda v = Av \tag{3}$$

where $A$ is the adjacency matrix of the graph, $\lambda$ is a constant (the eigenvalue), and $v$ is the eigenvector.

Panel D of Figure 1 reports brokerage network centrality calculated for all funds in our simple network. For instance, degree centrality for Fund 1 is 0.339 ($= (0.35 + 0.30 + 0.50 + 0.35 + 0.15 + 0.35 + 0.30 + 0.60 + 0.15)/9$). As can be seen in Panel A of Figure 1, funds that are positioned in the center of the network (e.g., 2, 3, and 5) are indeed more central than funds located in the periphery (e.g., 1, 4, and 10). In general, funds that trade through many brokers that many other funds also trade through tend to be central in the network.

In order to line up with the semi-annual N–SAR reporting frequency, we construct networks every half-year at the end of June (December) for N–SAR filings with reporting period ending in January to June (July to December) from the first half of 1994 to the first half of 2016. Since brokerage commission payments are only reported at the registrant level and are not broken down by fund, we construct networks at the registrant level and all funds within the same registrant inherit the same network structure. Figure 2 shows institutional brokerage networks constructed using our N–SAR data for the first half of 2016.

[Insert Figure 2]

Now we examine what types of mutual funds are more central in institutional brokerage networks

of centrality are degree, closeness, betweenness, and eigenvector centrality. When choosing the most appropriate measure, one must be careful about the implicit assumptions underlying these centrality measures. As laid out in Borgatti (2005), closeness centrality and betweenness centrality are built upon an implicit assumption that traffic flows along the shortest paths until it reaches a pre-determined destination like the package delivery process. In institutional brokerage networks, traffic is likely to freely flow from one fund (the fund submitting a trade order) to another (a potential fund that could absorb the submitted trade order) through the broker intermediating the trade. Since this type of traffic must flow through unrestricted walks, rather than via geodesics, closeness centrality and betweenness centrality can be safely ruled out. See also Ahern (2013) for a similar discussion.

by estimating the following linear regression model:

$$Centrality_{i,t} = \gamma \times Covariates_{i,t} + \alpha_i + \theta_t + \varepsilon_{i,t} \tag{4}$$

where $i$ indexes mutual funds and $t$ indexes time in half-years. The dependent variable is $Centrality_{i,t}$, fund $i$'s brokerage network centrality measured at the end of half-year $t$. $Covariates_{i,t}$ are a vector of fund-level characteristics that include log of fund TNA, log of family TNA, expense ratio, commission rate, trading volume, and average size-value-momentum percentiles of stock holdings, all measured at the end of half-year $t$. $\alpha_i$ denotes fund fixed-effects, $\theta_t$ denotes time fixed-effects, and standard errors are clustered at the fund level.

Table 3 presents the regression results. The dependent variable is degree centrality in columns (1) through (5) and eigenvector centrality in columns (6) through (10). Overall, we find that funds that are large or belong to large fund families tend to be more central in the network, as these funds can afford to trade through a large number of brokers that are themselves central in the network. This result suggests that brokerage relationships are costly to build and is consistent with Goldstein et al. (2009) who note that most institutions concentrate their order flows with a small number of brokers in order to become their important clients, whereas large institutions can easily obtain the premium status from most brokers.

[Insert Table 3]

As can be seen in columns (1) and (6), fund and family sizes alone can explain 19% and 28% of variation in degree centrality and eigenvector centrality, respectively. Adding other fund characteristics in columns (2) and (7) only marginally improves the explanatory power, raising adjusted $R^2$ to 26% and 31% for degree centrality and eigenvector centrality, respectively. In contrast, fixed-effects, especially fund fixed-effects, account for a large amount of variation in brokerage network centrality, suggesting that we can identify the network effects that are orthogonal to the size effects. Adding time and fund fixed-effects in columns (5) and (10) raises adjusted $R^2$ to 74% and 72% for degree centrality and eigenvector centrality, respectively. This result also implies that brokerage network centrality is highly persistent, reflecting the persistence in the underlying brokerage relationships.

# 4 Brokerage Network Centrality and Trading Performance

In Section 4.1, we begin our empirical analysis by showing that mutual funds' brokerage network centrality predicts their trading performance as measured by return gap. In Section 4.2, we turn to inspecting the specific mechanisms behind the return–gap premium associated with mutual funds' brokerage network centrality (simply the fund–centrality premium or the return–gap premium).

## 4.1 The Fund–Centrality Premium

### 4.1.1 The Time-Series Evidence

Despite extensive disclosure requirements, mutual funds are only required to disclose their holdings on a quarterly basis and their trading activities are generally unobservable (Kacperczyk, Sialm, and Zheng (2008)). In order to examine how institutional brokerage networks affect mutual fund trading performance, we use the return gap as our measure of trading performance. The return gap is calculated as the difference between the reported fund return and the return on a hypothetical portfolio that invests in the previously disclosed fund holdings (Grinblatt and Titman (1989), Kacperczyk, Sialm, and Zheng (2008)):

$$Return\ Gap_{i,t} = RET_{i,t} - (HRET_{i,t} - EXP_{i,t}) \tag{5}$$

where $RET_{i,t}$, is the fund $i$'s reported return net of expenses during month $t$, $EXP_{i,t}$, is the expense ratio for fund $i$ reported prior to month $t$, and $HRET_{i,t}$ is the fund $i$'s holdings return during month $t$, which is defined as:

$$HRET_{i,t} = \sum_k w_{i,k,t-1} R_{k,t} \tag{6}$$

where $w_{i,k,t-1}$ is the fund $i$'s portfolio weight on stock $k$ at the end of month $t-1$ and $R_{k,t}$ is the return on stock $k$ during month $t$.

At the end of every June and December, we sort mutual funds into quintile portfolios, based on their brokerage network centrality. The average time-series monthly returns from July 1994 to December 2016 are reported in Table 4. The full-sample results reported in Panel A show that the average return gap increases monotonically from the portfolio of peripheral funds (the lowest quintile of brokerage network centrality) to the portfolio of central funds (the highest quintile). The difference in average return gaps between central funds and peripheral funds is about five basis points per month (t-statistic = 5.03 to 5.26). After adjusting for the Fama–French–Carhart four–factor loadings, the central–minus–peripheral portfolio delivers an average alpha of four basis points per month (t-statistic = 4.48 to 4.75).

[Insert Table 4]

The economic magnitude of the relation between brokerage network centrality and return gap is meaningful as well. To put the numbers in perspective, we find that the return gap differential between the highest and lowest quintile portfolios sorted on brokerage network centrality is nearly half as large as that sorted on past return gap (Kacperczyk, Sialm, and Zheng (2008)). Furthermore, in our sub-sample analysis, we find that the fund–centrality premium is economically large and statistically significant in both early (1994-2007) and later (2008-2016) periods reported in Panel B and Panel C, respectively. This suggests that even in today's fragmented market with dark pools and smart order-routing systems, upstairs trading and institutional brokerage networks remain highly relevant to large institutional investors, as reported in the *Wall Street Journal*.[23]

### 4.1.2 The Cross-Sectional Evidence

In order to understand the specific mechanisms driving the return–gap premium associated with brokerage network centrality, it is important to recognize key factors affecting the return gap. The return gap is originally proposed by Grinblatt and Titman (1989) as a measure of total transactions costs for mutual funds. Therefore, at first brush, the fund–centrality premium is very much in line with our hypothesis that institutional brokerage networks mitigate mutual fund trading costs. Grinblatt and Titman (1989),

[23] "'Upstairs' Trading Draws More Big Investors," by Bradley Hope, the *Wall Street Journal*, December 8, 2013.

however, point out that the return gap may be affected by interim trades within a quarter and possibly window-dressing activities. Kacperczyk, Sialm, and Zheng (2008) further note that skilled fund managers can use their informational advantage to time the trades of individual stocks optimally and show that the past return gap helps predict fund performance.

We also recognize that the network formation is likely endogenous. For instance, marginal benefits of institutional brokerage networks are likely higher for better skilled ones, fund managers with superior trading skills might self-select into central positions in the network. There might also exist an unobservable (to the econometrician) factor that is correlated with both mutual funds' brokerage network centrality and their trading performance. For example, Kacperczyk, Sialm, and Zheng (2008) document the persistence in return gap and propose the return gap as a measure of interim trading skills of fund managers (see also Puckett and Yan (2011)). Anand et al. (2012) show that trading costs are closely linked to trading desks' execution skills over and above selecting better brokers.

In order to mitigate these confounding factors, we use cross-sectional regressions with fund fixed-effects to control for unobserved heterogeneity along with observable fund characteristics. Specifically, we estimate the following linear regression model:

$$\textit{Return Gap}_{i,t} = \beta \times \textit{Centrality}_{i,t-1} + \gamma \times \textit{Covariates}_{i,t-1} + \alpha_i + \theta_t + \varepsilon_{i,t} \tag{7}$$

where $i$ indexes mutual funds and $t$ indexes time in half-years. The dependent variable is $\textit{Return Gap}_{i,t}$ which is fund $i$'s average return gap during half-year $t$. $\textit{Centrality}_{i,t-1}$ is fund $i$'s brokerage network centrality measured at the end of half-year $t-1$. $\textit{Covariates}_{i,t-1}$ are a vector of fund-level characteristics that include log of fund TNA, log of family TNA, expense ratio, commission rate, trading volume, and average size-value-momentum percentiles of stock holdings, all measured at the end of half-year $t-1$. Depending on the specification, the regression includes fund fixed-effects $((\alpha_i))$ and lagged return gap. All regressions include time fixed-effects $(\theta_t)$ and standard errors are clustered at the fund level.

We present the regression results in Table 5. Columns (1) and (4) report our baseline specification including fund characteristics and time-fixed effects. The coefficients on $\textit{Centrality}_{i,t-1}$ are all positive and statistically significant at 1% levels. Interestingly, the our main coefficients change little when we add lagged return gap in columns (2) and (5). In the remaining columns, our main coefficients remain positive and statistically significant even after the inclusion of fund fixed-effects, mitigating endogeneity concerns that the fund–centrality premium could be driven by some unobserved heterogeneity.

[Insert Table 5]

Later in Section 5, we further address endogeneity concerns that could arise, for instance, from reverse causality. By exploiting mergers of large brokerage houses as plausibly exogenous shocks to the network structure, we provide evidence supportive of our causal interpretation that institutional brokerage networks *improve* institutional trading performance. Next, we turn our attention to testing our our hypothesis that institutional brokerage networks facilitate liquidity provision and mitigate trading costs associated with adverse selection.

## 4.2 Inspecting the Mechanism

### 4.2.1 The Fund–Centrality Premium when Funds Experience Severe Redemptions

The primary prediction that we can derive from our hypothesis is that the fund–centrality premium should be more pronounced when funds' trading activities are largely driven by liquidity motives and funds can credibly signal this to their brokers. We use large outflow events to identify such periods of liquidity-motivated trading. When a mutual fund is experiencing severe redemptions, the fund is forced to liquidate a large fraction of its holdings in several stocks and their selling is, to a large extent, uninformed (see, e.g., Coval and Stafford (2007), Alexander, Cici, and Gibson (2007)). In addition, such forced liquidations are likely to send a particularly strong signal to the brokers that its sell orders are driven by liquidity reasons, rather than information motivated, thus helping the brokers communicate more credibly with other institutional clients to take the other end of the trades.

In order to test this prediction, we estimate the following linear regression model:

$$\begin{aligned} \textit{Return Gap}_{i,t} &= \delta \times \textit{Centrality}_{i,t-1} \times \mathbb{1}(\textit{Outflow}_{i,t} > 5\%) + \beta \times \textit{Centrality}_{i,t-1} \\ &\quad + \rho \times \mathbb{1}(\textit{Outflow}_{i,t} > 5\%) + \gamma \times \textit{Covariates}_{i,t-1} + \alpha_i + \theta_t + \varepsilon_{i,t} \end{aligned} \tag{8}$$

where $\mathbb{1}(\textit{Outflow}_{i,t} > 5\%)$ is an indicator variable that is equal to 1 if fund $i$'s outflow during half-year $t$ exceeds five percent and the rest of the model is the same as in Equation (7). In some specifications, we include fund fixed-effects ($\alpha_i$). All regressions include time fixed-effects ($\theta_t$) and standard errors are clustered at the fund level.

We present the regression results in Table 6. The dependent variable is degree centrality in columns (1) and (2) and eigenvector centrality in columns (3) and (4). In the baseline specification without fund fixed-effects in columns (1) and (3), the coefficients on $\textit{Centrality}_{i,t-1}$ and $\textit{Centrality}_{i,t-1} \times \mathbb{1}(\textit{Outflow}_{i,t} > 5\%)$ are all positive and statistically significant at 1% levels. These results suggest that central funds tend to outperform peripheral funds in terms of return gap during normal times, but the fund–centrality premium is more pronounced when funds are faced with large outflows.

[Insert Table 6]

Next, in columns (2) and (4), we add fund fixed-effects to our baseline specification to control for unobserved heterogeneity such as trading skills of fund managers and execution skills of trading desks. By exploiting within-fund variation in investor flows, we continue to find that the fund–centrality premium is more pronounced when funds are forced to liquidate due to large outflows. As a robustness check, we re-define a large outflow event as a half-year during which the fund's outflow exceeds ten percent, instead of five percent, and still obtain qualitatively similar results, reported in Panel A of Table A2 in the Appendix.

One potential concern is that the above results could be also consistent with cross-subsidization within a fund family: when a fund is suffering severe redemptions, another fund in the same family could step in to provide liquidity. For instance, Bhattacharya, Lee, and Pool (2013) show that affiliated funds of mutual funds that invest only in other funds within the family provide an insurance pool against temporary liquidity shocks to other funds in the family. This alternative cross-subsidization hypothesis may seem plausible because we find that funds that belong to large families are more central and large fund families are likely better equipped to provide cross-subsidization. Nevertheless, we continue to find qualitatively similar results when we exclude funds that belong to large fund families, reported in Panel B of Table A2 in the Appendix.

Before we move on, we can further rule out another important alternative hypothesis. Many studies on brokerage connections have focused on various information channels.[24] Thus, it may seem plausible that central funds can acquire privileged information through their strong brokerage connections and trade on it. Our evidence, however, is at odds with this alternative information channel hypothesis: the fund–centrality premium is more pronounced when funds' trading activities are largely driven by liquidity reasons, rather than information motivated. We provide further evidence along this line in Section 4.2.4.

#### 4.2.2 The Fund–Centrality Premium for Valuable Clients

Second, our liquidity provision hypothesis requires an active role on the part of brokers, such as in discerning their clients' uninformed trading motives and communicating with other institutional clients. As made clear in Carlin, Lobo, and Viswanathan (2007), whether the brokers facilitate liquidity provision or foster predatory trading is likely to hinge on the incentives they face and the strength of repeated interaction with their clients. To the extent that brokers are incentivized to maximize the expected value of future commission revenues, central funds with greater revenue generating potential for brokers are most likely to benefit from liquidity provision facilitated by their brokers. In addition, combined with our primary prediction, the effect of brokers' incentives on the fund–centrality premium should be further amplified when funds are forced to liquidate in order to accommodate severe redemptions.

In order to test these predictions, we first interact a proxy for brokers' incentives with brokerage

[24] Such information channels include, but not limited to, early access to sell-side research or tipping (Irvine, Lipson, and Puckett (2007)), invitations to broker-hosted investor conferences (Green et al. (2014)), and information leakages on company fundamentals, especially in the context of hedge funds and their prime brokers (Chung and Kang (2016), Kumar et al. (2018), Di Maggio et al. (Forthcoming))

network centrality and estimate the following linear regression model:

$$\begin{aligned} \textit{Return Gap}_{i,t} = \delta \times \textit{Centrality}_{i,t-1} \times \textit{Broker Incentive}_{i,t-1} + \beta \times \textit{Centrality}_{i,t-1} \\ + \rho \times \textit{Broker Incentive}_{i,t-1} + \gamma \times \textit{Covariates}_{i,t-1} + \alpha_i + \theta_t + \varepsilon_{i,t} \end{aligned} \tag{9}$$

where $\textit{Broker Incentive}_{i,t-1}$ is our proxy for brokers' incentives as measured by fund $i$'s aggregate dollar commissions during half-year $t-1$ and the rest of the model is the same as in Equation (7).

We present the regression results in Panel A of Table 7. In columns (1) and (2), $\textit{Broker Incentive}_{i,t-1}$ is an indicator variable that is equal to one if fund $i$'s aggregate dollar commissions during half-year $t-1$ is greater than its top quartile value. Consistent with our prediction that brokers' incentives drive up the fund–centrality premium, we find a positive and statistically significant coefficient on $\textit{Centrality}_{i,t-1} \times \textit{Broker Incentive}_{i,t-1}$. In contrast, the coefficients on $\textit{Centrality}_{i,t-1}$ are small and statistically insignificant, suggesting that the fund–centrality premium is mostly accrued to central funds that are also likely valuable for brokers. As a robustness check in columns (3) and (4), we replace an indicator variable with its continuous counterpart, log of aggregate dollar commissions, for $\textit{Broker Incentive}_{i,t-1}$. We continue to obtain qualitatively similar, albeit somewhat weaker, results that essentially brokers' incentives drive up the fund–centrality premium.

[Insert Table 7]

Next, we add an indicator variable for contemporaneous large outflows as an additional interaction term in Equation (9) and run triple interaction regressions. We present the results in Panel B of Table 7. In all specifications, the coefficients on the triple interaction term, $\textit{Centrality}_{i,t-1} \times \textit{Broker Incentive}_{i,t-1} \times \mathbb{1}(\textit{Outflow}_{i,t} > 5\%)$, are positive and statistically significant at conventional levels. Overall, these results suggest that the effect of brokers' incentives on the fund–centrality premium is further amplified when funds' trading activities are largely driven by liquidity motives and funds can credibly signal this to their brokers.

### 4.2.3 The Fund–Centrality Premium for Relationship Clients

Third, our hypothesis relies on the repeated nature of interaction between institutional clients and their brokers. Institutional investors must build reputation for being truthful in order to credibly signal liquidity motives for their uninformed orders to their brokers. The brokers, in turn, must develop their reputation capital for being discreet when handling their clients' orders. Thus, the signaling and certification of uninformed trading motives is likely most effective if funds have already built strong trading relationships with their brokers.

In order to test this prediction, we interact a measure of existing trading relationships with brokerage network centrality and estimate the following linear regression model:

$$\begin{aligned} \textit{Return Gap}_{i,t} &= \delta \times \textit{Centrality}_{i,t-1} \times \textit{Trading Relationship}_{i,t-1} + \beta \times \textit{Centrality}_{i,t-1} \\ &+ \rho \times \textit{Trading Relationship}_{i,t-1} + \gamma \times \textit{Covariates}_{i,t-1} + \alpha_i + \theta_t + \varepsilon_{i,t} \end{aligned} \tag{10}$$

where $\textit{Trading Relationship}_{i,t-1}$, or simply, $\textit{Relationship}_{i,t-1}$ is our proxy for fund $i$'s strength of trading relationships with its current set of brokers, as measured by taking the minimum of a fraction of fund $i$'s commissions paid to its broker $k$ during half-year $t-1$ (current) and that during $t-3$ (a year before) and then summing it over all brokers currently employed by the fund. Intuitively, $\textit{Relationship}_{i,t-1}$ measures the extent (Jaccard distance) to which fund $i$'s current set of brokers overlap with the set of brokers the fund traded through a year before. The rest of the model is the same as in Equation (7).

We present the regression results in Panel A of Table 8. We find some evidence that trading relationships drive up the fund–centrality premium. In all specifications, the coefficients on $\textit{Centrality}_{i,t-1} \times \textit{Trading Relationship}_{i,t-1}$ are positive, but statistically significant only in columns (3) and (4) when we use eigenvector centrality. These somewhat weaker results, however, are not inconsistent with our liquidity provision hypothesis, which predicts that the fund–centrality premium is primarily driven by liquidity-motivated trades.

[Insert Table 8]

To test whether trading relationships drive up the fund–centrality premium especially in periods of heavy liquidity-motivated trades, we add an indicator variable for contemporaneous large outflows as an additional interaction term in Equation (10) and run triple interaction regressions. We present the results in Panel B of Table 8. Consistent with our prediction, the coefficients on the triple interaction term, $\mathit{Centrality}_{i,t-1} \times \mathit{Broker\ Relationship}_{i,t-1} \times \mathbb{1}(\mathit{Outflow}_{i,t} > 5\%)$, are positive and statistically significant at conventional levels in all specifications. Our results are also consistent with those found in a related literature on client-dealer networks. For instance, Di Maggio, Kermani, and Song (2017) show that prior trading relationships are valuable especially in turbulent times in the OTC corporate bond market.

#### 4.2.4 The Fund–Centrality Premium When Funds Submit Uninformed Large Orders

Our results thus far suggest that the return–gap premium associated with brokerage network centrality is more pronounced when funds' trading activities are largely driven by liquidity reasons and funds can credibly signal this to their brokers. In addition, we find that brokers' incentives and trading relationships further drive up the fund–centrality premium, corroborating our liquidity provision hypothesis. One could still argue that central funds can obtain the return–gap premium because central funds can more easily slice up large orders and spread across many brokers who can then further spread their clients' orders across many counterparties. Although not mutually exclusive with this alternative hypothesis, our liquidity provision hypothesis has clear predictions about the relation between the fund–centrality premium and the information content of trading. We provide further evidence that the fund–centrality premium is mostly concentrated in the periods that can be characterized by uninformed trading activities. We do so using an alternative measure, which puts emphasis on funds' trading volume in relation to fund flows, to identify such periods. In addition, we show that the fund–centrality premium is further amplified when the orders are also likely larger.

We identify periods of heavy information-motivated buying and selling activities following Alexander, Cici, and Gibson (2007). We calculate $BF$ and $SF$ metrics as follows:

$$BF_{i,t} = \frac{BUYS_{i,t} - FLOW_{i,t}}{TNA_{i,t-1}} \quad \& \quad SF_{i,t} = \frac{SELLS_{i,t} + FLOW_{i,t}}{TNA_{i,t-1}}$$

where $BUYS_{i,t}$ is fund $i$'s dollar volume of stock purchases during half-year $t$, $SELLS_{i,t}$ is fund $i$'s dollar volume of stock sales during half-year $t$, $FLOW_{i,t}$ is fund $i$'s net investor flow (inflow minus outflow) during half-year $t$, and $TNA_{i,t-1}$ is fund $i$'s total net assets at the end of half-year $t-1$. Exploiting within-fund variation in $BF$ and $SF$ metrics, Alexander, Cici, and Gibson (2007) show that buy (sell) portfolios with high $BF$ ($SF$) tend to outperform buy (sell) portfolios with low $BF$ ($SF$). Intuitively, trading against investor flows is likely motivated by superior private information, whereas trading with flows is likely driven by liquidity reasons, i.e., scaling up to accommodate inflows and scaling down to accommodate outflows (see also Coval and Stafford (2007)).

Since heavy informed buying activities do not necessarily coincide with heavy informed selling activities, we assign half-years in which both $BF$ and $SF$ fall below its respective top quartile value as periods of uninformed trading (or at least less informed trading). We interact an indicator variable for period of uninformed trading with brokerage network centrality and estimate the following linear regression model:

$$\begin{aligned} \mathit{Return\ Gap}_{i,t} &= \delta \times \mathit{Centrality}_{i,t-1} \times \mathbb{1}(BF_{i,t} < Q_3 \ \&\ SF_{i,t} < Q_3) + \beta \times \mathit{Centrality}_{i,t-1} \\ &\quad + \rho \times \mathbb{1}(BF_{i,t} < Q_3 \ \&\ SF_{i,t} < Q_3) + \gamma \times \mathit{Covariates}_{i,t-1} + \alpha_i + \theta_t + \varepsilon_{i,t} \end{aligned} \tag{11}$$

where $\mathbb{1}(BF_{i,t} < Q_3 \ \&\ SF_{i,t} < Q_3)$ is an indicator variable that is equal to 1 if both $BF_{i,t}$ and $SF_{i,t}$ fall below its respective top quartile value during half-year $t$ and the rest of the model is the same as in Equation (7).

We present the regression results in Panel A of Table 9. We find that the coefficients on $\mathit{Centrality}_{i,t-1} \times \mathbb{1}(BF_{i,t} < Q_3 \ \&\ SF_{i,t} < Q_3)$ are positive and statistically significant at 1% and 5% levels, whereas the coefficients on $\mathit{Centrality}_{i,t-1}$ are small and statistically insignificant. These results are consistent with our main results based on large outflow events and further suggest that the fund–centrality premium is associated with trading motives and mostly concentrated in periods of uninformed trading, i.e., when funds are trading *with flows*, rather than *against flows*.

[Insert Table 9]

Next, we proxy for average order sizes using average trade sizes inferred from consecutive portfolio disclosures, adjusting for trading volume in the market as follows:

$$\overline{\mathit{Trade\ Size}}_{i,t} = \frac{1}{N_{i,t}} \sum_{k} \frac{\mid Shares_{i,k,t} - Shares_{i,k,t-1} \mid}{\overline{VOL}_{k,t}^{CRSP}} \tag{12}$$

where $Shares_{i,k,t}$ is the split-adjusted number of shares held in stock $k$ by fund $i$ at the end of half-year (or quarter) $t$, $\overline{VOL}_{k,t}^{CRSP}$ is the average CRSP monthly volume between portfolio disclosures, and the averages are taken over stocks for which $Shares_{i,k,t} \neq Shares_{i,k,t-1}$. To arrive at the semi-annual figure, we take the average of quarterly numbers, if two quarterly observations are available.

In order to test whether the fund–centrality premium is more pronounced when funds submit *large* uninformed orders, we add in Equation (11) an additional interaction term, $\mathbb{1}(\overline{\text{Trade Size}} > Q_3)$, which is an indicator variable that is equal to 1 if $\overline{\mathit{Trade\ Size}}_{i,t}$ is above its top quartile value. We present the triple interaction results in Panel B of of Table 9. In columns (1) and (3), the coefficients on the triple interaction term, $\mathit{Centrality}_{i,t-1} \times \mathbb{1}(BF_{i,t} < Q_3 \ \&\ SF_{i,t} < Q_3) \times \mathbb{1}(\overline{\mathit{Trade\ Size}}_{i,t} > Q_3)$, are positive and significant at 5% and 10% levels. In contrast, the coefficients on $\mathit{Centrality}_{i,t-1} \times \mathbb{1}(\overline{\mathit{Trade\ Size}}_{i,t} > Q_3)$ are small and statistically insignificant. These results suggest that central funds can obtain the return–gap premium when central funds submit large *uninformed* orders. As a robustness check in columns (3) and (4), we replace an indicator variable with its continuous counterpart, $\overline{\mathit{Trade\ Size}}_{i,t}$ and continue to obtain qualitatively similar results. Overall, our results suggest that the fund–centrality premium is more pronounced when the trading orders are larger, but only when the trades are likely motivated by liquidity reasons. These results are largely consistent with our hypothesis that institutional brokerage networks facilitate liquidity provision and mitigate trading costs associated with adverse selection.

# 5 A Natural Experiment

We recognize that our results are not completely free from endogeneity concerns that could be derived from, for instance, reverse causality. Hence, we conduct a natural experiment to provide evidence supportive of our causal interpretation that institutional brokerage networks *improve* institutional trading performance. To accomplish this, we exploit mergers of large brokerage houses as plausibly exogenous shocks to the network structure.

## 5.1 Backgrounds on Brokerage Mergers and Identification

Following Hong and Kacperczyk (2010), we identify mergers among brokerage houses by relying on information from the SDC Mergers and Acquisition database. We choose all the mergers that the acquiring broker belongs to the four-digit SIC code 6211 ("investment Commodity Firms, Dealers, and Exchanges"). Next, we manually match brokerage mergers identified in the SDC data using broker names and narrow down to the mergers in which broker names show up in at least 100 N-SAR filings.[25] This process gives rise to twenty six brokerage mergers during the period from 1995 to 2015. Table 10 lists all twenty six brokerage mergers. The table also reports average broker shares before (from18 months to 6 months) and after (from 6 months to 18 months) the merger and changes in average broker shares around the merger.

[Insert Table 10]

The shock strength, however, is a major concern for our natural experiment, given the complexity of the network structure (which typically consists of thousands of nodes connected by tens of thousands edge). Moderate-sized brokerage mergers, especially as stand-alone events (which amounts to cutting a small number of edges connected to a single node) are unlikely to have an economically meaningful impact on the entire structure of institutional brokerage networks. Therefore, we focus on two waves of five largest mergers of institutional brokerage houses that took place around 2000 and 2008, in which more than ten percent of edges were served.[26]

Figure 3 plots the changes in average broker shares around each of these mergers. A visual inspection suggests that these five mergers were likely to have a meaningful impact on institutional brokerage networks. Specifically, the average brokerage shares of the acquired brokers dramatically decreased following the

[25] Our N–SAR sample period runs from 1994 to 2016. But we exclude the first and last years to facilitate a difference-in-differences (DiD) analysis around the merger.

[26] These five brokerage mergers include CSFB's acquisition of DLJ and UBS'acquistion of PaineWebber in 2000 and JP Morgan Chase's acquisition of Bear Stearns, Barclay's acquisition of Lehman Brothers, and Bank of America's acquisition of Merrill Lynch in 2008.

merger in all cases, whereas those of the acquiring brokers increased notably after the merger except for the case of Bank of America. For instance, mutual funds on average paid about 4.02 % of its brokerage commissions to CSFB as one of the top 10 brokers, while the figure for DLJ was 4.40%. After the merger, CSFB's average broker shares increased to 6.40%. One notable exception is Bank of America's acquisition of Merrill Lynch. After the merger, the merged firm's brokerage services were carried out under the name of Merrill Lynch for a while and thus reported as such in N–SAR reports.

[Insert Figure 3]

## 5.2 Empirical Design and Results

Our analysis of the causal effect of mutual funds' brokerage network centrality on their trading performance exploits large brokerage mergers in a quasi-natural experiment setting to overcome potential concerns about endogeneous network formation. As stated earlier, we exploit two waves of five largest mergers of brokerage houses and the empirical methodology of our analysis is a difference-in-differences (DiD). In a standard DiD approach, the sample needs to be divided into treatment and control groups. Here comes another challenge for our natural experiment: the treatment of shock is *a priori* unclear. Nevertheless, we can reason that mutual funds that traded largely through the acquiring brokers but not heavily through the target (acquired) brokers are most likely to benefit from exogenous shocks to the network, since the acquiring broker would retain at least some of the target broker's clients.

Building on this intuition, we construct hypothetical post-merger brokerage network centrality under a fairly conservative assumption. Specifically, we assume that funds who had relationships with a target broker before the merger were to simply redistribute commissions to their existing brokers on a pro-rata basis following the merger.[27] Then, we proceed by calculating the expected change in brokerage network

[27] In particular, we re-scale each mutual fund's normalized commission payment vector ($g_{i,\cdot}$) a half-year prior to the merger event window, denoted $\tilde{g}_{i,\cdot}$, as follows:

$$\tilde{g}_{i,k} = \begin{cases} 0 & \text{if } k \in S; \\ \frac{g_{i,k}}{\sum_{k \notin S} g_{i,k}} & \text{if } k \notin S. \end{cases} \tag{13}$$

where $i$ indexs funds, $k$ indexs brokers, and $S$ denotes set of acquired brokers. For instance, if a mutual fund $i$ hired centrality as the difference between the hypothetical post-merger centrality and the actual pre-merger centrality around the merger event. Under this assumption, the funds that did not trade through the target broker (candidate treated funds) do not change their brokerage relationships, as they don't need to, but nonetheless experience exogenous increases in brokerage network centrality after the merger, because *other* funds need to reconfigure their brokerage relationships. We form the treatment group by choosing the top ten percent of mutual funds sorted based on the expected change in brokerage network centrality.

Our empirical methodology also requires that we specify the event window around the mergers. In general, most event studies focus on a very narrow window because choosing a window that is too long may include irrelevant information with the focused events (Hong and Kacperczyk (2010)). However, a window that is too short would result in the loss of many observations containing relevant information and we thus choose a relatively longer time window than other event studies. Specifically, we examine one year before and one year after the event window of brokerage mergers. Figure 4 illustrates the event timelines for our natural experiment.

[Insert Figure 4]

If we denote the average outcome variables in the treatment (T) and control (C) groups in the pre- and post-event periods by $O_{T,1}$, $O_{T,2}$, $O_{C,1}$, and $O_{C,2}$, respectively, the partial effect of change due to the merger can be estimated as

$$DiD = (O_{T,2} - O_{T,1}) - (O_{C,2} - O_{C,1}). \tag{14}$$

A potential concern with the above estimation is that the results could be affected by fund characteristics. In other words, if the funds in the treatment and control groups have different fund characteristics, then those characteristics could potentially bias our results. To resolve this concern, we use a matching technique. As mentioned earlier, we assign top ten percent of funds with the largest expected change in brokerage network centrality as the treatment group. Among the remaining 90% of the sample, we construct the control group by matching on pre-treatment (pre-event) outcome variables and all fund characteristics

broker A, B, C, and D and its $g_i = \begin{bmatrix}0.1 & 0.3 & 0.4 & 0.2\end{bmatrix}$ and C is an acquired broker, then $\tilde{g}_i = \begin{bmatrix}\frac{0.1}{0.6} & \frac{0.3}{0.6} & 0 & \frac{0.2}{0.6}\end{bmatrix}$.

used in our previous analyses except for *log(Family TNA)*[28] following *Genetic Matching* algorithm proposed by Diamond and Sekhon (2013). Matching on observable pre-event fund characteristics and pre-treatment outcome variables can remove (at least to a degree) common influences of fund characteristics that could affect return gap other than changes in brokerage network centrality.

Table 11 reports the results of our matching using *Degree Centrality* as our measure of brokerage network centrality. As seen in the table, some of the variables are remarkably different before matching, but those differences largely disappear after matching. Panel A presents the matching balance results for the brokerage mergers in 2000. Before matching, *Degree Centrality* is significantly different at the 1% level, i.e., the treated funds were more central to begin with. In addition, among covariates, *Expense Ratio*, *Size Percentile*, and *Value Percentile* are significantly different at the conventional levels. Our matching appears successful and all p-values for post-matching differences in means are above 10%, with the smallest p-value of 0.18. Panel B presents the matching balance results for the brokerage mergers in 2008. Before matching, *Degree Centrality* is also significantly different at the 1% level and several covariates including *log(Fund TNA)*, *Expense Ratio*, *Commission Rate*, and *Trade Volume* are also significantly different at the conventional levels. Again, the matching appears similarly successful.

[Insert Table 11]

To be consistent with our causal interpretation, if the brokerage mergers had indeed served as positive exogenous shocks to the brokerage network centrality of mutual funds in the treatment group, then the return gap of treated funds would have experienced significant increases relative to that of the control group of funds following the mergers.

Table 12 presents our DiD results. Panel A shows the results of our DiD analysis of the brokerage mergers in 2000. The average *Degree Centrality* of the treatment group increased from 0.206 to 0.235, while the average *Degree Centrality* of the matched control group only increased from 0.205 to 0.222. Thus, we observe a discernible increase in *Degree Centrality* of 0.013, using a DiD estimator. This effect

[28] It turns out that it is very difficult to match on fund family size and *all* the other fund characteristics including pre-event outcome variables.

is statistically significant at the 5% level. Moreover, the average *Return Gap* also substantially increased around the mergers in 2000 by 9.3 basis points per month relative to a control group of funds, significant at the 10% level. Similarly, Panel B presents the results of our DiD analysis of the brokerage mergers in 2008. We similarly observe a discernible increase in *Degree Centrality* by 0.034, using a DiD estimator, significant at the 1% level. At the same time, the average *Return Gap* of the treated funds also substantially increased by 6.8 basis points per month relative to a control group of funds, significant at the 10% level. In sum, the DiD results indicate that exogenous changes in brokerage network centrality due to large brokerage mergers are accompanied by predicted changes in return gap performance.

[Insert Table 12]

As a robustness check, we re-do our DiD analysis with *Eigenvector Centrality* instead of *Degree Centrality*. We obtain qualitatively similar results, as reported in Table A3 and Table A4. To sum up, positive changes in brokerage network centrality as a result of exogenous shocks to the brokerage network are accompanied by positive changes in return gap. These results are consistent with our causal interpretation that institutional brokerage networks *improve* institutional trading performance.

# 6 Conclusion

Using a unique dataset on brokerage commission payments for a comprehensive sample of mutual funds, we map trading networks of mutual funds and their brokers as affiliation networks in which mutual funds are connected through their overlapping brokerage relationships. Mutual funds that trade through many brokers that many other funds also trade through are central in the network. We find that central funds outperform peripheral ones, especially in terms of return gap. In order to shed light on the specific mechanisms behind the return–gap premium associated with brokerage network centrality (simply the fund–centrality premium), we propose a liquidity provision hypothesis.

Suppose, for instance, that a mutual fund faced with an extreme fund outflow is forced to sell large blocks of its holdings in several stocks at the same time. The sell orders would tend to be submitted to brokers with which the fund has strong relationships and that could infer the underlying liquidity reasons for the orders. The brokers, in turn, may be likely to turn to other institutional clients with whom they have strong relationships to absorb the orders while communicating the likely liquidity motives for the trades to ease their concerns about trading against better informed traders. Thus, central funds are better positioned to tap into larger pools of unexpressed liquidity, especially when submitting large blocks of liquidity-motivated orders.

Consistent with our liquidity provision hypothesis, we find that the fund–centrality premium is more pronounced when funds' trading activities are largely driven by liquidity motives, such as to accommodate large fund outflows. We also find that the fund–centrality premium is further driven up by brokers' incentives to generate greater commission revenues and by trading relationships that funds have established with their brokers. Exploiting large brokerage mergers as plausibly exogenous shocks to the network structure, we provide evidence supportive of our causal interpretation that institutional brokerage networks improve institutional trading performance.

# References

Admati, A. R., and P. Pfleiderer. 1991. Sunshine Trading and Financial Market Equilibrium. *Review of Financial Studies* 4:443–481.

Ahern, K. R. 2013. Network Centrality and the Cross Section of Stock Returns. *University of Southern California Working Paper* pp. 1–51.

Alexander, G. J., G. Cici, and S. Gibson. 2007. Does Motivation Matter when Assessing Trade Performance? An Analysis of Mutual Funds. *Review of Financial Studies* 20:125–150.

Anand, A., P. Irvine, A. Puckett, and K. Venkataraman. 2012. Performance of Institutional Trading Desks: An Analysis of Persistence in Trading Costs. *Review of Financial Studies* 25:557–598.

Barbon, A., M. Di Maggio, F. A. Franzoni, and A. Landier. Forthcoming. Brokers and Order Flow Leakage: Evidence from Fire Sales. *Journal of Finance* .

Bessembinder, H., A. Carrion, L. Tuttle, and K. Venkataraman. 2016. Liquidity, resiliency and market quality around predictable trades: Theory and evidence. *Journal of Financial Economics* 121:142–166.

Bessembinder, H., C. S. Spatt, and K. Venkataraman. Forthcoming. A Survey of the Microstructure of Fixed-Income Markets. *Journal of Financial and Quantitative Analysis* .

Bessembinder, H., and K. Venkataraman. 2004. Does an Electronic Stock Exchange Need an Upstairs Market? *Journal of Financial Economics* 73:3–36.

Bhattacharya, U., J. H. Lee, and V. K. Pool. 2013. Conflicting Family Values in Mutual Fund Families. *Journal of Finance* 68:173–200.

Bonacich, P. 1972. Factoring and Weighting Approaches to Status Scores and Clique Identification. *The Journal of Mathematical Sociology* 2:113–120.

Bonacich, P. 1987. Power and Centrality: A Family of Measures. *American Journal of Sociology* 92:1170–1182.

Booth, G. G., J.-C. Lin, T. Martikainen, and Y. Tse. 2002. Trading and Pricing in Upstairs and Downstairs Stock Markets. *Review of Financial Studies* 15:1111–1135.

Borgatti, S. P. 2005. Centrality and Network Flow. *Social Networks* 27:55–71.

Brunnermeier, M. K., and L. H. Pedersen. 2005. Predatory Trading. *Journal of Finance* 60:1825–1863.

Carlin, B. I., M. S. Lobo, and S. Viswanathan. 2007. Episodic Liquidity Crises: Cooperative and Predatory Trading. *Journal of Finance* 62:2235–2274.

Chung, J.-W., and B. U. Kang. 2016. Prime Broker-level Comovement in Hedge Fund Returns: Information or Contagion? *Review of Financial Studies* 29:3321–3353.

Coval, J., and E. Stafford. 2007. Asset Fire Sales (and Purchases) in Equity Markets. *Journal of Financial Economics* 86:479–512.

Di Maggio, M., F. A. Franzoni, A. Kermani, and C. Sommavilla. Forthcoming. The Relevance of Broker Networks for Information Diffusion in the Stock Market. *Journal of Financial Economics* .

Di Maggio, M., A. Kermani, and Z. Song. 2017. The value of trading relations in turbulent times. *Journal of Financial Economics* 124:266–284.

Diamond, A., and J. S. Sekhon. 2013. Genetic Matching for Estimating Causal Effects. *The Review of Economics and Statistics* 95:932–945.

Edelen, R. M. 1999. Investor Flows and the Assessed Performance of Open-end Mutual Funds. *Journal of Financial Economics* 53:439–466.

Edelen, R. M., R. B. Evans, and G. B. Kadlec. 2012. Disclosure and agency conflict: Evidence from mutual fund commission bundling. *Journal of Financial Economics* 103:308–326.

Freeman, L. C. 1979. Centrality in Social Networks Conceptual Clarification. *Social Networks* 1:215–239.

Glosten, L. R., and P. R. Milgrom. 1985. Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. *Journal of financial economics* 14:71–100.

Goldstein, M. A., P. Irvine, E. Kandel, and Z. Wiener. 2009. Brokerage Commissions and Institutional Trading Patterns. *Review of Financial Studies* 22:5175–5212.

Green, T. C., R. Jame, S. Markov, and M. Subasi. 2014. Broker-hosted Investor Conferences. *Journal of Accounting and Economics* 58:142–166.

Griffin, J. M., T. Shu, and S. Topaloglu. 2012. Examining the Dark Side of Financial Markets: Do Institutions Trade on Information from Investment Bank Connections? *Review of Financial Studies* 25:2155–2188.

Grinblatt, M., and S. Titman. 1989. Mutual Fund Performance: An Analysis of Quarterly Portfolio Holding. *Journal of Business* 62:393–416.

Grossman, S. J. 1992. The Informational Role of Upstairs and Downstairs Trading. *Journal of Business* 65:509–528.

Harris, L. 2002. *Trading and Exchanges: Market Microstructure for Practitioners*. Oxford University Press.

Hendershott, T., D. Li, D. Livdan, and N. Schürhoff. 2017. Relationship Trading in OTC Markets. *Swiss Finance Institute Research Paper No. 17-30* .

Hong, H., and M. Kacperczyk. 2010. Competition and Bias. *Quarterly Journal of Economics* 125:1683–1725.

Irvine, P., M. Lipson, and A. Puckett. 2007. Tipping. *Review of Financial Studies* 20:741–768.

Kacperczyk, M., C. Sialm, and L. Zheng. 2008. Unobserved Actions of Mutual Funds. *Review of Financial Studies* 21:2379–2416.

Kumar, N., K. Mullally, S. Ray, and Y. Tang. 2018. Prime (Information) Brokerage. *Working Paper, University of Florida* .

Kyle, A. S. 1985. Continuous Auctions and Insider Trading. *Econometrica* 53:1315–1335.

Levenshtein, V. I. 1966. Binary codes capable of correcting deletions, insertions, and reversals. *Soviet Physics Doklady* 10:707–701.

Li, D., and N. Schürhoff. Forthcoming. Dealer Networks. *Journal of Finance* .

Madhavan, A., and M. Cheng. 1997. In Search of Liquidity: Block Trades in the Upstairs and and Downstairs. *Review of Financial Studies* 10:175–203.

Ozsoylev, H. N., J. Walden, M. D. Yavuz, and R. Bildik. 2014. Investor Networks in the Stock Market. *Review of Financial Studies* 27:1323–1366.

Puckett, A., and X. S. Yan. 2011. The Interim Trading Skills of Institutional Investors. *Journal of Finance* 66:601–633.

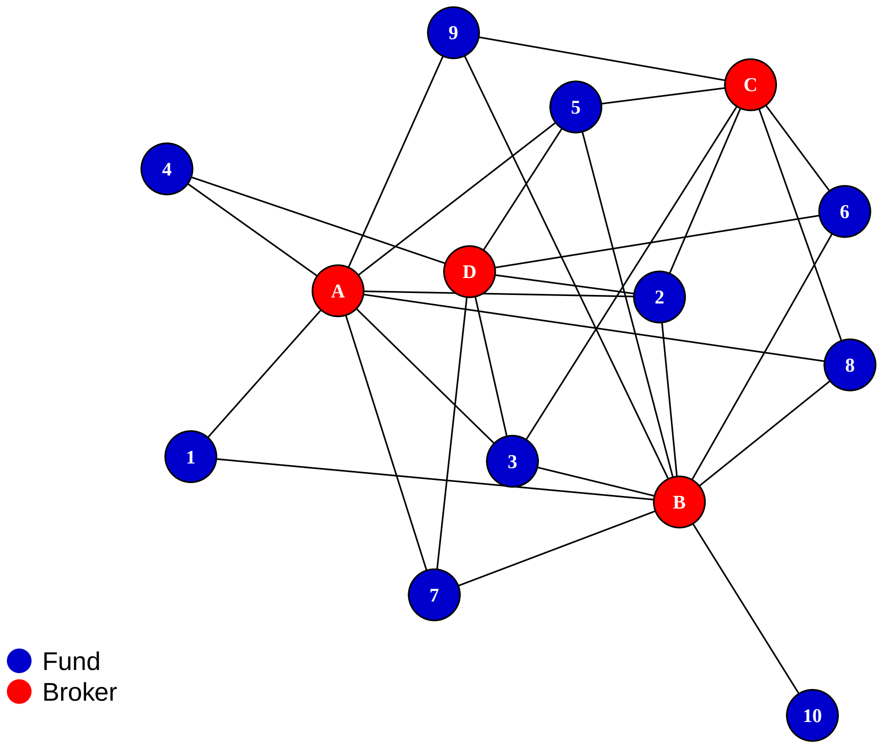
Reuter, J. 2006. Are IPO Allocations for Sale? Evidence from Mutual Funds. *Journal of Finance* 61:2289–2324.

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

Smith, B. F., D. A. S. Turnbull, and R. W. White. 2001. Upstairs Market for Principal and Agency Trades: Analysis of Adverse Information and Price Effects. *Journal of Finance* 56:1723 –1746.

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

*Panel A: Graphical representation*




*Panel B: Extended-form representation*

$$G^T = \text{Broker} \begin{array}{c} \\ A \\ B \\ C \\ D \end{array} \begin{array}{c} \text{Fund} \\ \begin{array}{cccccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \end{array} \\ \begin{bmatrix} 0.85 & 0.20 & 0.15 & 0.50 & 0.20 & 0.00 & 0.20 & 0.15 & 0.50 & 0.00 \\ 0.15 & 0.33 & 0.40 & 0.00 & 0.17 & 0.25 & 0.50 & 0.50 & 0.10 & 1.00 \\ 0.00 & 0.22 & 0.10 & 0.00 & 0.28 & 0.40 & 0.00 & 0.35 & 0.40 & 0.00 \\ 0.00 & 0.25 & 0.35 & 0.50 & 0.35 & 0.35 & 0.30 & 0.00 & 0.00 & 0.00 \end{bmatrix} \end{array}$$

*Panel C: Reduced-form representation*

$$A = \begin{array}{c} \\ \text{Fund} \end{array} \begin{array}{c|cccccccccc} & \multicolumn{10}{c}{\text{Fund}} \\ & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ \hline 1 & 0.00 & 0.35 & 0.30 & 0.50 & 0.35 & 0.15 & 0.35 & 0.30 & 0.60 & 0.15 \\ 2 & 0.35 & 0.00 & 0.83 & 0.45 & 0.84 & 0.72 & 0.78 & 0.70 & 0.52 & 0.33 \\ 3 & 0.30 & 0.83 & 0.00 & 0.50 & 0.77 & 0.70 & 0.85 & 0.65 & 0.35 & 0.40 \\ 4 & 0.50 & 0.45 & 0.50 & 0.00 & 0.55 & 0.35 & 0.50 & 0.15 & 0.50 & 0.00 \\ 5 & 0.35 & 0.84 & 0.77 & 0.55 & 0.00 & 0.80 & 0.67 & 0.60 & 0.58 & 0.17 \\ 6 & 0.15 & 0.72 & 0.70 & 0.35 & 0.80 & 0.00 & 0.55 & 0.60 & 0.50 & 0.25 \\ 7 & 0.35 & 0.78 & 0.85 & 0.50 & 0.67 & 0.55 & 0.00 & 0.65 & 0.30 & 0.50 \\ 8 & 0.30 & 0.70 & 0.65 & 0.15 & 0.60 & 0.60 & 0.65 & 0.00 & 0.60 & 0.50 \\ 9 & 0.60 & 0.52 & 0.35 & 0.50 & 0.58 & 0.50 & 0.30 & 0.60 & 0.00 & 0.10 \\ 10 & 0.15 & 0.33 & 0.40 & 0.00 & 0.17 & 0.25 & 0.50 & 0.50 & 0.10 & 0.00 \end{array}$$

*Panel D: Brokerage network centrality*

| Fund | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| Degree Centrality | 0.339 | 0.613 | 0.594 | 0.389 | 0.592 | 0.513 | 0.572 | 0.528 | 0.450 | 0.267 |
| Eigenvector Centrality | 0.550 | 1.000 | 0.972 | 0.654 | 0.974 | 0.870 | 0.928 | 0.864 | 0.730 | 0.468 |

**Figure 1:** Institutional Brokerage Networks: A Toy Example

This figure illustrates how we construct institutional brokerage networks and calculate brokerage network centrality using a simple example network consisting of ten funds and four brokers.

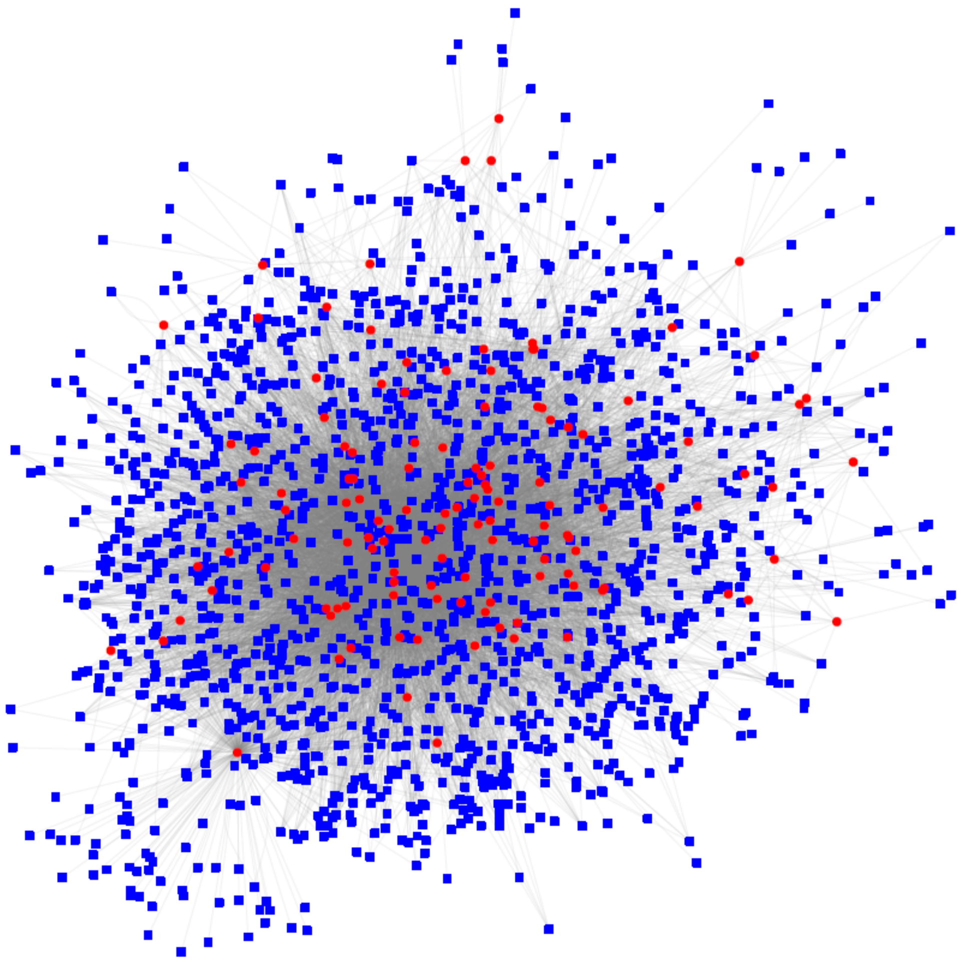
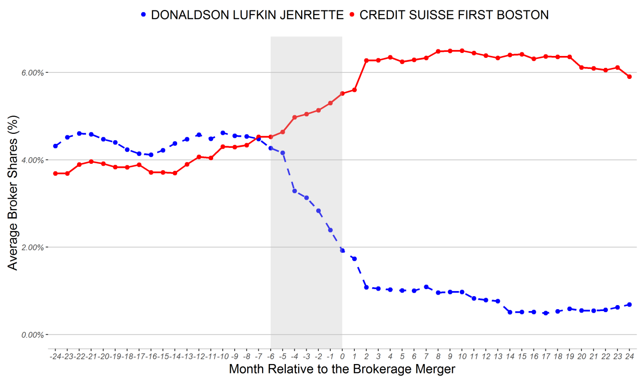
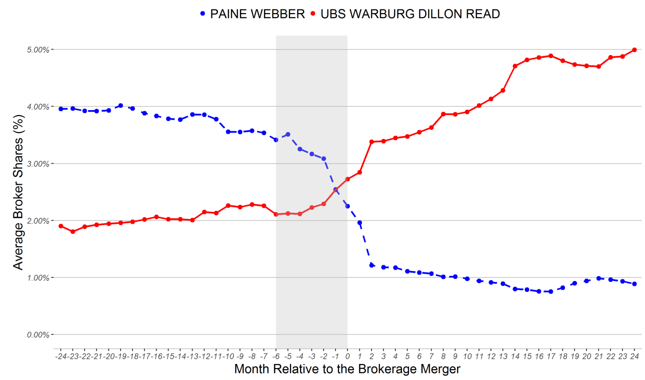


**Figure 2:** Institutional Brokerage Networks

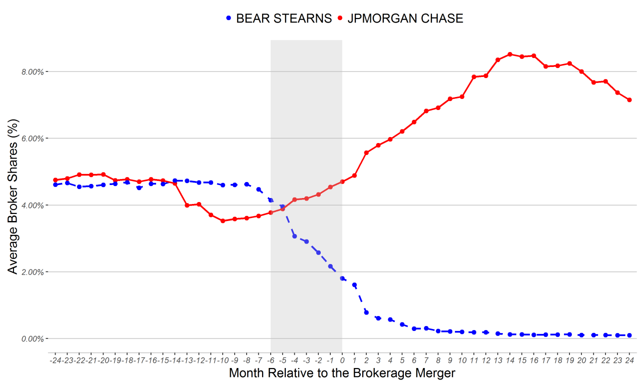
This figure shows a snapshot of institutional brokerage networks at the end of June 2016. Blue nodes represent mutual funds, red nodes represent institutional brokers, and lines represent connections between mutual funds and their brokers.



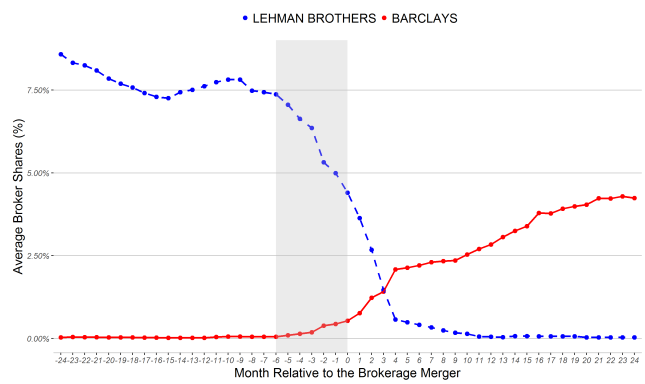

**(a)** CSFB's acquisition of DLJ



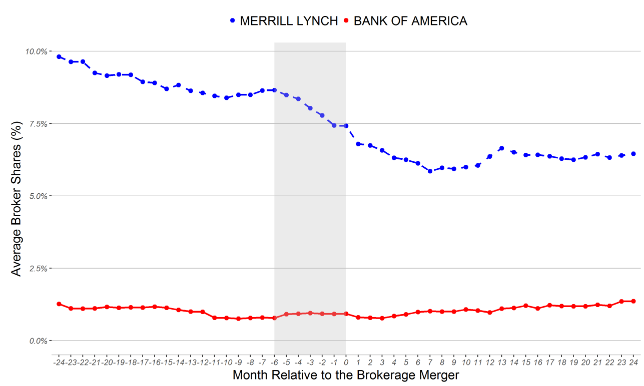

**(b)** UBS Warburg Dillon Read's acquisition of Paine Webber




**(c)** JP Morgan Chase's acquisition of Bear Stearns



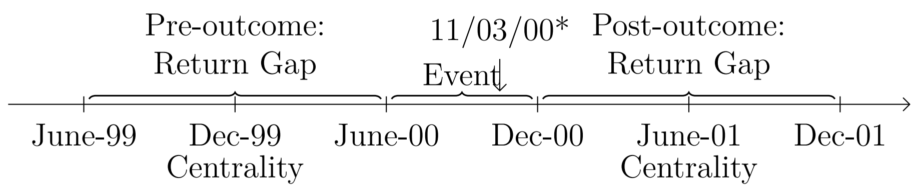

**(d)** Barclay's acquisition of Lehman Brothers



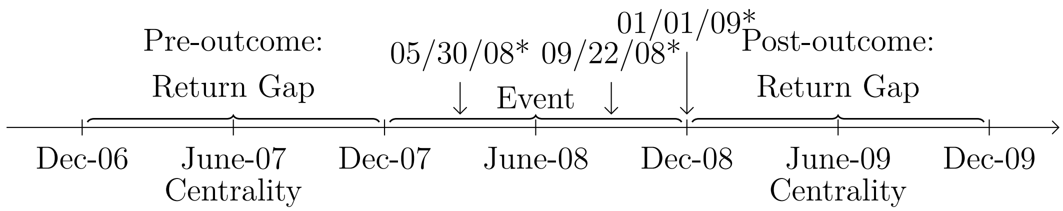

**(e)** Bank of America's acquisition of Merrill Lynch

**Figure 3:** Average Brokerage Share around Brokerage Merger

This figure shows changes in average broker shares for the acquiring brokers and target brokers around the mergers. A broker share is defined as a fraction of the commission payments to the given broker by the fund and broker shares are averaged across funds each month on a rolling basis around each of the following mergers: Credit Suisse First Boston (CSFB)'s acquisition of Donaldson Lufkin Jenrette (DLJ) in 2000 (a); UBS Warburg Dillon Read's acquisition of Paine Webber in 2000 (b); JP Morgan Chase's acquisition of Bear Stearns in 2008 (c); Barclays's acquisition of Lehman Brothers in 2008 (d); and Bank of America's acquisition of Merrill Lynch in 2008 (e).




**(a)** Event Timeline of Brokerage Mergers in 2000




**(b)** Event Timeline of Brokerage Mergers in 2008

**Figure 4:** Event Timeline of Brokerage Mergers

Figure 4a depicts the event timeline of the 2000 mergers: Credit Suisse First Boston (CSFB)'s acquisition of Donaldson Lufkin Jenrette (DLJ) and UBS Warburg Dillon Read's acquisition of Paine Webber in 2000. The effective date of both mergers is November 3rd, 2000. We set the second half of 2000 as the event window.
Figure 4b depicts the event timeline of the 2008 mergers: JP Morgan Chase's acquisition of Bear Stearns, Barclays's acquisition of Lehman Brothers, and Bank of America's acquisition of Merrill Lynch in 2008. The effective dates are May 30th, 2008, September 22nd, 2008, and January 1st, 2009, respectively. We set the entire year of 2008 as the event window.
* Effective date is as reported by SDC Platinum Financial Securities Data.

**Table 1:** Brokerage Commission Payments: Example and Descriptions

This table provides an example of and some descriptive statistics on brokerage commission payments. N-SAR filings report brokerage commissions paid to the 10 brokers that received the largest amount (Item 20) from the fund and the aggregate brokerage commission payments (Item 21). Panel A provides an example for T. Rowe Price Blue Chip Growth Fund for the period ending in June 30, 2016. Panel B reports the concentration level of brokerage commissions for the top 1, 3, 5, 7, and 10 brokers to which the fund paid the largest amount. Panel C reports the transition matrix of year-to-year changes in the broker rankings for the fund by the amount of commission payments.

Panel A: Example: T ROWE PRICE BLUE CHIP GROWTH FUND (CIK = 902259), June 30, 2016

| Item 20 | Name of Broker | IRS Number | Commisions ($000) |
|---|---|---|---|
| 1 | BANK OF AMERICA MERRILL LYNCH | 13-5674085 | 415 |
| 2 | JPMORGAN CHASE | 13-4994650 | 292 |
| 3 | MORGAN STANLEY CO INC | 13-2655998 | 252 |
| 4 | DEUTSCHE BANK SECURITIES | 13-2730828 | 207 |
| 5 | RBC CAPITAL MARKETS | 41-1416330 | 159 |
| 6 | CITIGROUP GLOBAL MARKETS INC | 11-2418191 | 157 |
| 7 | CS FIRST BOSTON | 13-5659485 | 153 |
| 8 | BAIRD ROBERT W | 39-6037917 | 148 |
| 9 | GOLDMAN SACHS | 13-5108880 | 144 |
| 10 | SANFORD C BERNSTEIN | 13-2625874 | 115 |
| Item 21 | Aggregate Brokerage Commissions ($000) | | 3107 |

Panel B: Concentration of Brokerage Commissions

| Broker Share (%) | Mean | St. Dev. | Pctl(1) | Pctl(25) | Median | Pctl(75) | Pctl(99) |
|---|---|---|---|---|---|---|---|
| Top 1 Broker | 25.65 | 22.21 | 5.53 | 11.54 | 16.88 | 30.00 | 100.00 |
| Top 1–3 Brokers | 45.24 | 23.95 | 13.02 | 27.59 | 37.44 | 56.76 | 100.00 |
| Top 1–5 Brokers | 56.60 | 22.53 | 19.00 | 39.26 | 51.17 | 71.13 | 100.00 |
| Top 1–7 Brokers | 64.47 | 20.91 | 23.80 | 48.25 | 61.32 | 80.63 | 100.00 |
| Top 1–10 Brokers | 72.45 | 18.91 | 29.30 | 58.08 | 71.62 | 88.89 | 100.00 |

Panel C: Persistence in Brokerage Relationship (Transition Matrix)

| Probability (%) | Next Year | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Current Year | Top 1 | Top 2 | Top 3 | Top 4 | Top 5 | Top 6 | Top 7 | Top 8 | Top 9 | Top 10 |
| Top 1 | 46.74 | 20.55 | 13.35 | 10.06 | 7.57 | 6.44 | 5.41 | 4.99 | 4.41 | 4.00 |
| Top 2 | 17.37 | 23.69 | 17.29 | 13.03 | 10.96 | 8.89 | 7.58 | 6.63 | 6.30 | 5.65 |
| Top 3 | 10.71 | 15.83 | 17.64 | 14.52 | 12.34 | 10.93 | 9.61 | 8.53 | 7.14 | 7.62 |
| Top 4 | 7.17 | 11.24 | 13.33 | 15.12 | 13.31 | 11.82 | 10.10 | 9.59 | 9.42 | 8.47 |
| Top 5 | 5.31 | 8.09 | 10.53 | 12.73 | 13.83 | 12.84 | 12.16 | 10.54 | 10.10 | 9.67 |
| Top 6 | 4.00 | 6.47 | 8.87 | 10.49 | 11.84 | 13.00 | 12.91 | 12.12 | 10.78 | 10.67 |
| Top 7 | 3.12 | 5.30 | 6.65 | 8.18 | 10.38 | 11.67 | 12.77 | 13.11 | 12.71 | 11.21 |
| Top 8 | 2.41 | 3.60 | 5.00 | 6.56 | 8.02 | 9.93 | 11.72 | 13.73 | 13.70 | 13.53 |
| Top 9 | 1.85 | 2.86 | 4.06 | 5.12 | 6.15 | 8.22 | 10.22 | 11.28 | 14.08 | 13.93 |
| Top 10 | 1.32 | 2.36 | 3.28 | 4.19 | 5.60 | 6.25 | 7.52 | 9.48 | 11.35 | 15.26 |

**Table 2:** Summary Statistics

This table reports the summary statistics on degree centrality (Freeman (1979)), eigenvector centrality (Bonacich (1972, 1987)), and other fund-level characteristics over the period from the first half of 1994 through the first half of 2016. The fund TNA (Item 74.T) and an indicator for an index fund (Item 69) are directly taken from N-SAR filings and we use the family code reported by the fund (Item 19.C) to calculate the family TNA. The fund trading volume is calculated as the sum of purchases (Item 71.A) and sales (Item 71.B). Since brokerage commissions are reported at the registrant level, we calculate the commission rate as a ratio of the aggregate commission payments (Item 21) to the sum of all trading volumes across equity-oriented funds within the same registrant. We estimate the fund's commission payments as the product of the commission rate and the fund trading volume. The expense ratio is from CRSP and we calculate monthly net flows for each fund share class $i$ during month $t$ as follows: $FLOW_{i,t} = TNA_{i,t} - TNA_{i,t-1} \times (1 + R_{i,t})$ where $FLOW_{i,t}$ is the dollar value of fund flow (net new issues and redemptions), $TNA_{i,t}$ is the total net asset, and $R_{i,t}$ is the monthly return. To compute the monthly fund flow for the fund, we sum monthly fund flows for all share classes belonging to the same fund as identified by MFLINKS. Monthly fund flows are summed over the half-year to calculate the semi-annual fund flow. We scale the semi-annual fund flows by the beginning-of-period $TNA$. For each fund-quarter, size, value, and momentum percentiles are calculated as percentiles of market capitalization, book-to-market ratio, and 12-month returns skipping the most recent month, respectively, averaged across all stock holdings. For each fund-halfyear, we take the most recent quarterly observation of average size-value-momentum percentiles.

| Variable | Obs. | Mean | St. Dev. | $Q_1$ | Median | $Q_3$ |
|---|---|---|---|---|---|---|
| Degree Centrality | 54,331 | 0.16 | 0.08 | 0.10 | 0.16 | 0.21 |
| Eigenvector Centrality | 54,331 | 0.53 | 0.25 | 0.33 | 0.57 | 0.73 |
| Return Gap (%) | 54,331 | −0.03 | 0.39 | −0.20 | −0.02 | 0.14 |
| Fund TNA ($billion) | 54,331 | 1.42 | 3.43 | 0.08 | 0.29 | 1.09 |
| Family TNA ($billion) | 54,331 | 122.09 | 277.91 | 2.97 | 20.50 | 79.80 |
| Expense Ratio (%) | 54,331 | 1.13 | 0.42 | 0.92 | 1.11 | 1.35 |
| Commission Rate (%) | 54,331 | 0.12 | 0.13 | 0.06 | 0.09 | 0.14 |
| Trade Volume, as % of TNA | 54,331 | 86.32 | 80.10 | 34.81 | 63.84 | 108.64 |
| $\mathbb{1}$(Index Fund) | 54,331 | 0.10 | 0.30 | 0 | 0 | 0 |
| Size Percentile | 54,331 | 84.97 | 12.60 | 76.87 | 89.52 | 95.03 |
| Value Percentile | 54,331 | 37.46 | 11.92 | 28.02 | 37.21 | 46.15 |
| Momentum Percentile | 54,331 | 57.69 | 9.56 | 51.55 | 57.07 | 63.46 |
| Fund Flow, as % of TNA | 54,331 | 1.98 | 22.44 | −8.01 | −2.25 | 5.71 |

## **Table 3:** Determinants of Mutual Funds' Brokerage Network Centrality

This table presents the results of regressing degree centrality and eigenvector centrality on contemporaneous fund-level characteristics including log(fund TNA), log(family TNA), expense ratio, commission ratio, trading volume, and size-value-momentum percentiles. The details on the fund-level variables are reported in Table 2. The standard errors are clustered at the fund level and the resulting t-statistics are reported in parentheses. Statistical significance at the 10%, 5%, and 1% level is indicated by *, **, and ***, respectively.

| *Dependent variable:* | Degree Centrality ×100 | | | | | Eigenvector Centrality ×100 | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
| Constant | 2.34*** | −2.54** | | | | −1.32 | −21.41*** | | | |
| | (6.18) | (−2.33) | | | | (−1.08) | (−5.86) | | | |
| log(Fund TNA) | 0.09 | 0.25*** | 0.14** | 0.20*** | 0.21*** | 0.64*** | 0.90*** | 0.58*** | 0.85*** | 0.83*** |
| | (1.53) | (4.12) | (2.53) | (3.11) | (3.13) | (3.19) | (4.50) | (2.93) | (3.72) | (3.59) |
| log(Family TNA) | 1.36*** | 1.44*** | 1.62*** | 0.60*** | 0.59*** | 5.21*** | 5.33*** | 5.78*** | 1.92*** | 1.91*** |
| | (29.57) | (32.21) | (36.92) | (8.02) | (7.95) | (34.98) | (36.49) | (38.31) | (7.53) | (7.44) |
| Expense Ratio (%) | | 1.35*** | | | −0.22 | | 1.52* | | | −1.76** |
| | | (5.54) | | | (−0.89) | | (1.90) | | | (−2.07) |
| Commission Rate (%) | | 3.94*** | | | −0.82*** | | 6.21*** | | | −1.72* |
| | | (8.76) | | | (−2.86) | | (4.58) | | | (−1.78) |
| Trading Volume, as % of TNA | | 0.01*** | | | 0.002*** | | 0.03*** | | | 0.01*** |
| | | (12.34) | | | (4.11) | | (9.73) | | | (3.49) |
| Size Percentile | | 0.09*** | | | 0.02 | | 0.27*** | | | 0.03 |
| | | (11.16) | | | (1.59) | | (9.43) | | | (0.72) |
| Value Percentile | | −0.05*** | | | −0.01 | | −0.10*** | | | −0.003 |
| | | (−6.75) | | | (−0.77) | | (−3.68) | | | (−0.12) |
| Momentum Percentile | | −0.10*** | | | −0.01*** | | −0.12*** | | | −0.03* |
| | | (−15.80) | | | (−2.85) | | (−6.37) | | | (−1.94) |
| Time Fixed Effects | No | No | Yes | Yes | Yes | No | No | Yes | Yes | Yes |
| Fund Fixed Effects | No | No | No | Yes | Yes | No | No | No | Yes | Yes |
| Observations | 54,331 | 54,331 | 54,331 | 54,331 | 54,331 | 54,331 | 54,331 | 54,331 | 54,331 | 54,331 |
| Adjusted $R^2$ | 0.19 | 0.26 | 0.42 | 0.74 | 0.74 | 0.28 | 0.31 | 0.33 | 0.72 | 0.72 |

**Table 4:** Brokerage Network Centrality and Mutual Fund Performance: Portfolio Sorts

This table reports the average time-series monthly returns from July 1994 to December 2016. Funds are sorted into quintile portfolios based on degree centrality (in columns (1) to column (6)) and eigenvector centrality (in columns (7) to (12)). The investor return is decomposed into the holdings return (net of expenses) and the return gap following Equation (5). Raw returns as well as four-factor adjusted returns are reported for average return gap, average holdings return (net of expenses), and average investor return. Panel A reports the full sample results, whereas Panel B and Panel C report the split-sample results. The heteroskedasticity robust t-statistics are reported in parentheses. Statistical significance at the 10%, 5%, and 1% level is indicated by *, **, and ***, respectively.

Panel A: Full Sample: July 1994 to December 2016

| | Raw Return (% per month) | | | | | | 4-Factor Alpha (% per month) | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Peripheral | Q2 | Q3 | Q4 | Central | C − P | Peripheral | Q2 | Q3 | Q4 | Central | C − P |
| *Sorted on Degree Centrality* | | | | | | | | | | | | |
| Return Gap | −0.06*** | −0.05** | −0.04** | −0.03 | −0.01 | 0.05*** | −0.02 | −0.02 | −0.01 | 0.003 | 0.02 | 0.04*** |
| | (−3.12) | (−2.58) | (−2.32) | (−1.54) | (−0.74) | (5.03) | (−1.41) | (−1.08) | (−0.61) | (0.20) | (1.19) | (4.48) |
| Holdings Return | 0.88*** | 0.87*** | 0.88*** | 0.86*** | 0.87*** | −0.01 | 0.13*** | 0.13*** | 0.13*** | 0.13*** | 0.12*** | −0.005 |
| | (3.03) | (3.04) | (3.05) | (2.99) | (2.96) | (−0.19) | (3.20) | (3.28) | (2.95) | (2.95) | (3.18) | (−0.19) |
| Investor Return | 0.81*** | 0.83*** | 0.83*** | 0.84*** | 0.86*** | 0.05 | 0.11*** | 0.12*** | 0.12*** | 0.13*** | 0.14*** | 0.03 |
| | (2.98) | (3.00) | (3.04) | (3.02) | (3.04) | (1.55) | (2.67) | (2.85) | (2.78) | (2.98) | (3.58) | (1.53) |
| *Sorted on Eigenvector Centrality* | | | | | | | | | | | | |
| Return Gap | −0.07*** | −0.04** | −0.04** | −0.03 | −0.01 | 0.05*** | −0.02* | −0.01 | −0.01 | 0.002 | 0.02 | 0.04*** |
| | (−3.27) | (−2.43) | (−2.25) | (−1.61) | (−0.71) | (5.26) | (−1.66) | (−0.85) | (−0.56) | (0.11) | (1.22) | (4.75) |
| Holdings Return | 0.88*** | 0.87*** | 0.88*** | 0.87*** | 0.87*** | −0.02 | 0.13*** | 0.12*** | 0.13*** | 0.13*** | 0.12*** | −0.01 |
| | (3.06) | (3.01) | (3.04) | (3.01) | (2.96) | (−0.52) | (3.39) | (3.09) | (2.98) | (2.87) | (3.26) | (−0.46) |
| Investor Return | 0.82*** | 0.82*** | 0.83*** | 0.84*** | 0.85*** | 0.04 | 0.11*** | 0.11*** | 0.12*** | 0.13*** | 0.14*** | 0.03 |
| | (2.99) | (2.99) | (3.03) | (3.03) | (3.04) | (1.34) | (2.80) | (2.74) | (2.79) | (2.93) | (3.64) | (1.44) |
| *Sorted on Past Return Gap* | | | | | | | | | | | | |
| Return Gap | −0.09*** | −0.05*** | −0.05*** | −0.02 | 0.02 | 0.11*** | −0.03 | −0.02 | −0.03** | −0.002 | 0.05*** | 0.08*** |
| | (−3.50) | (−3.24) | (−3.33) | (−1.47) | (0.77) | (4.57) | (−1.41) | (−1.48) | (−2.02) | (−0.16) | (2.67) | (4.16) |
| Holdings Return | 0.88*** | 0.86*** | 0.88*** | 0.87*** | 0.88*** | −0.003 | 0.08 | 0.12*** | 0.17*** | 0.16*** | 0.11* | 0.03 |
| | (2.88) | (3.07) | (3.21) | (3.07) | (2.79) | (−0.05) | (1.54) | (2.72) | (4.11) | (3.38) | (1.93) | (0.53) |
| Investor Return | 0.79*** | 0.81*** | 0.83*** | 0.84*** | 0.90*** | 0.11* | 0.05 | 0.10** | 0.15*** | 0.16*** | 0.16*** | 0.11* |
| | (2.74) | (2.99) | (3.12) | (3.09) | (3.03) | (1.95) | (1.01) | (2.28) | (3.60) | (3.35) | (2.90) | (1.96) |

## Table 4–*Continued*

Panel B: Sub-sample: July 1994 to December 2007

| | Raw Return (% per month) | | | | | | 4-Factor Alpha (% per month) | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Peripheral | Q2 | Q3 | Q4 | Central | C − P | Peripheral | Q2 | Q3 | Q4 | Central | C − P |
| *Sorted on Degree Centrality* | | | | | | | | | | | | |
| Return Gap | −0.04 | −0.03 | −0.03 | −0.01 | 0.01 | 0.05*** | 0.01 | 0.02 | 0.02 | 0.03* | 0.05*** | 0.03*** |
| | (−1.53) | (−1.18) | (−1.11) | (−0.33) | (0.23) | (3.36) | (0.80) | (1.05) | (1.09) | (1.93) | (2.71) | (2.68) |
| Holdings Return | 0.99*** | 0.98*** | 0.99*** | 0.97*** | 0.97*** | −0.02 | 0.21*** | 0.23*** | 0.22*** | 0.22*** | 0.19*** | −0.02 |
| | (2.82) | (2.80) | (2.82) | (2.73) | (2.65) | (−0.51) | (3.95) | (3.81) | (3.31) | (3.36) | (3.41) | (−0.66) |
| Investor Return | 0.94*** | 0.95*** | 0.96*** | 0.96*** | 0.97*** | 0.03 | 0.23*** | 0.24*** | 0.24*** | 0.25*** | 0.24*** | 0.01 |
| | (2.89) | (2.88) | (2.90) | (2.85) | (2.82) | (0.64) | (4.31) | (4.22) | (3.76) | (4.00) | (4.45) | (0.47) |
| *Sorted on Eigenvector Centrality* | | | | | | | | | | | | |
| Return Gap | −0.05* | −0.03 | −0.03 | −0.01 | 0.01 | 0.06*** | 0.01 | 0.02 | 0.02 | 0.03** | 0.04** | 0.04*** |
| | (−1.75) | (−1.02) | (−1.01) | (−0.33) | (0.20) | (3.60) | (0.36) | (1.41) | (1.23) | (1.99) | (2.57) | (2.88) |
| Holdings Return | 1.00*** | 0.97*** | 0.98*** | 0.97*** | 0.96*** | −0.04 | 0.23*** | 0.21*** | 0.22*** | 0.21*** | 0.20*** | −0.03 |
| | (2.85) | (2.77) | (2.80) | (2.75) | (2.65) | (−0.89) | (4.20) | (3.57) | (3.37) | (3.22) | (3.56) | (−0.86) |
| Investor Return | 0.95*** | 0.95*** | 0.96*** | 0.96*** | 0.97*** | 0.02 | 0.23*** | 0.23*** | 0.24*** | 0.25*** | 0.24*** | 0.01 |
| | (2.91) | (2.86) | (2.89) | (2.88) | (2.81) | (0.38) | (4.50) | (4.04) | (3.81) | (3.94) | (4.52) | (0.40) |
| *Sorted on Past Return Gap* | | | | | | | | | | | | |
| Return Gap | −0.08** | −0.05** | −0.04** | −0.003 | 0.08** | 0.17*** | 0.01 | −0.01 | −0.01 | 0.03* | 0.11*** | 0.11*** |
| | (−2.33) | (−2.42) | (−2.13) | (−0.15) | (2.18) | (4.94) | (0.25) | (−0.44) | (−0.74) | (1.70) | (4.82) | (4.05) |
| Holdings Return | 1.02*** | 0.99*** | 0.97*** | 0.94*** | 0.97** | −0.05 | 0.14** | 0.22*** | 0.26*** | 0.25*** | 0.20** | 0.06 |
| | (2.66) | (2.97) | (2.96) | (2.76) | (2.39) | (−0.53) | (2.22) | (3.99) | (4.37) | (3.43) | (2.10) | (0.64) |
| Investor Return | 0.94*** | 0.94*** | 0.93*** | 0.94*** | 1.05*** | 0.11 | 0.14** | 0.22*** | 0.25*** | 0.28*** | 0.31*** | 0.16** |
| | (2.64) | (2.95) | (2.95) | (2.89) | (2.78) | (1.35) | (2.46) | (4.19) | (4.54) | (3.93) | (3.42) | (2.03) |

Panel C: Sub-sample: January 2008 to December 2016

| | Raw Return (% per month) | | | | | | 4-Factor Alpha (% per month) | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Peripheral | Q2 | Q3 | Q4 | Central | C − P | Peripheral | Q2 | Q3 | Q4 | Central | C − P |
| *Sorted on Degree Centrality* | | | | | | | | | | | | |
| Return Gap | −0.10*** | −0.07*** | −0.07** | −0.06** | −0.05* | 0.05*** | −0.06*** | −0.05*** | −0.04** | −0.04* | −0.02 | 0.04*** |
| | (−3.23) | (−3.04) | (−2.47) | (−2.27) | (−1.72) | (4.37) | (−3.42) | (−2.75) | (−2.06) | (−1.84) | (−1.15) | (4.36) |
| Holdings Return | 0.71 | 0.71 | 0.71 | 0.71 | 0.73 | 0.02 | −0.07* | −0.06** | −0.06* | −0.06* | −0.04 | 0.03 |
| | (1.43) | (1.45) | (1.45) | (1.46) | (1.48) | (0.55) | (−1.91) | (−2.03) | (−1.84) | (−1.73) | (−1.29) | (1.01) |
| Investor Return | 0.61 | 0.64 | 0.65 | 0.65 | 0.68 | 0.07** | −0.13*** | −0.11*** | −0.10*** | −0.10** | −0.07* | 0.06*** |
| | (1.29) | (1.34) | (1.36) | (1.37) | (1.43) | (2.15) | (−3.39) | (−3.31) | (−2.84) | (−2.56) | (−1.76) | (2.73) |
| *Sorted on Eigenvector Centrality* | | | | | | | | | | | | |
| Return Gap | −0.10*** | −0.07*** | −0.07** | −0.06** | −0.04 | 0.05*** | −0.06*** | −0.05*** | −0.04** | −0.04** | −0.02 | 0.04*** |
| | (−3.19) | (−3.01) | (−2.51) | (−2.43) | (−1.61) | (4.53) | (−3.34) | (−2.75) | (−2.11) | (−2.03) | (−1.01) | (4.42) |
| Holdings Return | 0.71 | 0.71 | 0.72 | 0.71 | 0.73 | 0.02 | −0.07* | −0.06** | −0.06* | −0.06* | −0.04 | 0.03 |
| | (1.43) | (1.45) | (1.46) | (1.45) | (1.48) | (0.56) | (−1.95) | (−2.06) | (−1.73) | (−1.75) | (−1.31) | (1.02) |
| Investor Return | 0.61 | 0.64 | 0.65 | 0.65 | 0.68 | 0.07** | −0.13*** | −0.11*** | −0.10*** | −0.10*** | −0.06* | 0.07*** |
| | (1.30) | (1.34) | (1.36) | (1.36) | (1.44) | (2.14) | (−3.39) | (−3.35) | (−2.74) | (−2.72) | (−1.70) | (2.74) |
| *Sorted on Past Return Gap* | | | | | | | | | | | | |
| Return Gap | −0.10*** | −0.05** | −0.06*** | −0.05** | −0.07* | 0.03 | −0.07*** | −0.03* | −0.04** | −0.04** | −0.03 | 0.03 |
| | (−2.77) | (−2.22) | (−2.83) | (−2.45) | (−1.82) | (0.96) | (−2.73) | (−1.70) | (−2.49) | (−2.09) | (−1.27) | (1.31) |
| Holdings Return | 0.67 | 0.66 | 0.74 | 0.75 | 0.74 | 0.07 | −0.12** | −0.11*** | −0.005 | −0.01 | −0.04 | 0.08 |
| | (1.33) | (1.36) | (1.56) | (1.54) | (1.48) | (1.29) | (−2.27) | (−3.18) | (−0.18) | (−0.31) | (−0.96) | (1.33) |
| Investor Return | 0.57 | 0.61 | 0.69 | 0.69 | 0.67 | 0.11 | −0.19*** | −0.14*** | −0.05* | −0.05 | −0.08 | 0.11* |
| | (1.18) | (1.28) | (1.46) | (1.47) | (1.41) | (1.64) | (−3.21) | (−3.70) | (−1.72) | (−1.35) | (−1.50) | (1.77) |

## **Table 5:** Brokerage Network Centrality and Return Gap: Panel Regressions

This table examines whether our previous results documenting the fund–centrality premium based on portfolio sorts continue to hold after controlling for fund characteristics, including lagged return gap, and fund fixed-effects. Specifically, this table presents the results of our baseline linear regression model:

$$Return\ Gap_{i,t} = \beta \times Centrality_{i,t-1} + \gamma \times Covariates_{i,t-1} + \alpha_i + \theta_t + \varepsilon_{i,t}$$

where $i$ indexes mutual funds and $t$ indexes time in half-years. The dependent variable is $Return\ Gap_{i,t}$ which is fund $i$'s average return gap during half-year $t$. The independent variable of interest is $Centrality_{i,t-1}$, which is fund $i$'s brokerage network centrality (degree centrality or eigenvector centrality) measured at the end of half-year $t-1$. $Covariates_{i,t-1}$ are a vector of fund-level variables that are measured at the end of time $t-1$ and include log(fund TNA), log(family TNA), expense ratio, commission rate, trading volume, and average size-value-momentum percentiles of the stocks in the fund's portfolio. More details on fund-level variables are provided in Table 2. In some specifications, the regression includes lagged return gap and fund fixed-effects ($\alpha_i$) and all regressions include time fixed-effects ($\theta_t$). Standard errors are clustered at the fund level and the resulting t-statistics are reported in parentheses. Statistical significance at the 10%, 5%, and 1% level is indicated by *, **, and ***, respectively.

| *Dependent variable:* | Return Gap (%) | | | | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Degree Centrality | 0.15*** | 0.13*** | 0.09** | | | |
| | (4.62) | (4.44) | (1.97) | | | |
| Eigenvector Centrality | | | | 0.04*** | 0.04*** | 0.03** |
| | | | | (4.70) | (4.52) | (2.01) |
| Past Return Gap (%) | | 0.08*** | 0.01 | | 0.08*** | 0.01 |
| | | (11.07) | (0.89) | | (11.08) | (0.89) |
| log(Fund TNA) | −0.01*** | −0.01*** | −0.03*** | −0.01*** | −0.01*** | −0.03*** |
| | (−6.91) | (−6.87) | (−9.72) | (−6.92) | (−6.89) | (−9.72) |
| log(Family TNA) | 0.01*** | 0.01*** | 0.01*** | 0.01*** | 0.01*** | 0.01*** |
| | (6.25) | (6.24) | (2.61) | (6.16) | (6.14) | (2.61) |
| Expense Ratio (%) | −0.01 | −0.01 | 0.02 | −0.01 | −0.01 | 0.02 |
| | (−1.09) | (−1.10) | (1.20) | (−1.05) | (−1.07) | (1.22) |
| Commission Rate (%) | −0.04*** | −0.04*** | −0.07*** | −0.04*** | −0.04*** | −0.07*** |
| | (−2.97) | (−3.00) | (−3.98) | (−2.99) | (−3.02) | (−3.99) |
| Trading Volume, as % of TNA | 0.0000 | 0.0000 | 0.0001* | 0.0000 | 0.0000 | 0.0001* |
| | (0.75) | (0.78) | (1.92) | (0.76) | (0.79) | (1.93) |
| Size Percentile | −0.001*** | −0.001*** | 0.001 | −0.001*** | −0.001*** | 0.001 |
| | (−4.11) | (−3.99) | (1.35) | (−4.13) | (−4.01) | (1.36) |
| Value Percentile | −0.002*** | −0.002*** | −0.001*** | −0.002*** | −0.002*** | −0.001*** |
| | (−10.66) | (−10.65) | (−2.73) | (−10.67) | (−10.66) | (−2.74) |
| Momentum Percentile | −0.001** | −0.001** | −0.001** | −0.001** | −0.001** | −0.001** |
| | (−2.32) | (−2.08) | (−2.03) | (−2.34) | (−2.10) | (−2.04) |
| Time Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes |
| Fund Fixed Effects | No | No | Yes | No | No | Yes |
| Observations | 54,331 | 54,331 | 54,331 | 54,331 | 54,331 | 54,331 |
| Adjusted $R^2$ | 0.07 | 0.08 | 0.10 | 0.07 | 0.08 | 0.10 |

**Table 6:** The Fund–Centrality Premium when Funds Experience Severe Redemptions

This table examines whether the fund–centrality premium is more pronounced when funds' trading activities are primarily driven by liquidity reasons, such as to accommodate large investor redemptions. Specifically, we interact an indicator variable for contemporaneous large outflows with lagged brokerage network centrality in our baseline specification as follows:

$$\begin{aligned} \textit{Return Gap}_{i,t} &= \delta \times \textit{Centrality}_{i,t-1} \times \mathbb{1}(\textit{Outflow}_{i,t} > 5\%) + \beta \times \textit{Centrality}_{i,t-1} \\ &+ \rho \times \mathbb{1}(\textit{Outflow}_{i,t} > 5\%) + \gamma \times \textit{Covariates}_{i,t-1} + \alpha_i + \theta_t + \varepsilon_{i,t} \end{aligned}$$

where $\mathbb{1}(\textit{Outflow}_{i,t} > 5\%)$ is an indicator variable that is equal to 1 if fund $i$'s outflow during half-year $t$ exceeds five percent and the rest of the model is the same as in Table 5. Standard errors are clustered at the fund level and the resulting t-statistics are reported in parentheses. Statistical significance at the 10%, 5%, and 1% level is indicated by *, **, and ***, respectively.

| *Dependent variable:* | Return Gap (%) | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Degree Centrality × $\mathbb{1}$(Outflow > 5%) | 0.14*** | 0.17*** | | |
| | (2.94) | (3.19) | | |
| Eigenvector Centrality × $\mathbb{1}$(Outflow > 5%) | | | 0.04*** | 0.04*** |
| | | | (2.89) | (2.70) |
| Degree Centrality | 0.10*** | 0.03 | | |
| | (2.89) | (0.67) | | |
| Eigenvector Centrality | | | 0.03*** | 0.01 |
| | | | (2.75) | (0.79) |
| $\mathbb{1}$(Outflow > 5%) | −0.03*** | −0.03*** | −0.03*** | −0.03*** |
| | (−3.32) | (−3.25) | (−3.23) | (−2.78) |
| log(Fund TNA) | −0.01*** | −0.03*** | −0.01*** | −0.03*** |
| | (−6.87) | (−9.71) | (−6.87) | (−9.65) |
| log(Family TNA) | 0.01*** | 0.01*** | 0.01*** | 0.01*** |
| | (6.20) | (2.59) | (6.13) | (2.60) |
| Expense Ratio (%) | −0.005 | 0.02 | −0.005 | 0.02 |
| | (−0.91) | (1.26) | (−0.86) | (1.30) |
| Commission Rate (%) | −0.04*** | −0.07*** | −0.04*** | −0.07*** |
| | (−2.94) | (−3.96) | (−2.98) | (−4.00) |
| Trading Volume, as % of TNA | 0.0000 | 0.0001** | 0.0000 | 0.0001** |
| | (0.90) | (1.97) | (0.93) | (2.00) |
| Size Percentile | −0.001*** | 0.001 | −0.001*** | 0.001 |
| | (−4.11) | (1.35) | (−4.13) | (1.36) |
| Value Percentile | −0.002*** | −0.001*** | −0.002*** | −0.001*** |
| | (−10.67) | (−2.71) | (−10.70) | (−2.73) |
| Momentum Percentile | −0.001** | −0.001** | −0.001** | −0.001** |
| | (−2.38) | (−2.12) | (−2.40) | (−2.14) |
| Time Fixed Effects | Yes | Yes | Yes | Yes |
| Fund Fixed Effects | No | Yes | No | Yes |
| Observations | 54,331 | 54,331 | 54,331 | 54,331 |
| Adjusted $R^2$ | 0.07 | 0.11 | 0.07 | 0.11 |

## Table 7: The Fund–Centrality Premium For Valuable Clients

This table examines whether the fund–centrality premium is larger for more valuable clients, especially when the client funds are forced to trade to accommodate large investor redemptions. In unconditional tests presented in Panel A, we interact a measure of brokerage revenue generating potential with brokerage network centrality in our baseline specification as follows:

$$\begin{aligned} \textit{Return Gap}_{i,t} &= \delta \times \textit{Centrality}_{i,t-1} \times \textit{Broker Incentive}_{i,t-1} + \beta \times \textit{Centrality}_{i,t-1} \\ &+ \rho \times \textit{Broker Incentive}_{i,t-1} + \gamma \times \textit{Covariates}_{i,t-1} + \alpha_i + \theta_t + \varepsilon_{i,t} \end{aligned}$$

where $\textit{Broker Incentive}_{i,t-1}$ is our proxy for fund $i$'s brokerage revenue generating potential as measured by an indicator variable that is equal to one if fund $i$'s aggregate dollar commissions during half-year $t-1$ is greater than its top quartile value. As a robustness check, we replace an indicator variable with its continuous counterpart, log of aggregate dollar commissions in columns (3) and (4). The rest of the model is the same as in Table 5. The independent variable of interest is $\textit{Centrality}_{i,t-1} \times \textit{Broker Incentive}_{i,t-1}$ to tease out the effect of brokers' incentives on the fund–centrality premium. In conditional tests presented in Panel B, we add an indicator variable for contemporaneous large outflows as an additional interaction term. Standard errors are clustered at the fund level 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:* | Return Gap (%) | | | |
|---|---|---|---|---|
| *Broker Incentive:* | $\mathbb{1}$(Dollar Commission > $Q_3$) | | log(Dollar Commission) | |
| | (1) | (2) | (3) | (4) |
| Degree Centrality × Broker Incentive | 0.26*** | | 0.04* | |
| | (3.23) | | (1.96) | |
| Eigenvector Centrality × Broker Incentive | | 0.05* | | 0.003 |
| | | (1.66) | | (0.52) |
| Degree Centrality | 0.03 | | 0.16*** | |
| | (0.53) | | (2.74) | |
| Eigenvector Centrality | | 0.02 | | 0.04* |
| | | (1.15) | | (1.94) |
| Broker Incentive | −0.06*** | −0.04** | −0.01*** | −0.01* |
| | (−3.70) | (−2.26) | (−2.74) | (−1.86) |
| log(Fund TNA) | −0.03*** | −0.03*** | −0.03*** | −0.03*** |
| | (−8.60) | (−8.58) | (−6.09) | (−5.98) |
| log(Family TNA) | 0.01*** | 0.01*** | 0.01*** | 0.01** |
| | (2.66) | (2.65) | (2.62) | (2.56) |
| Expense Ratio (%) | 0.02 | 0.02 | 0.01 | 0.02 |
| | (1.17) | (1.21) | (1.14) | (1.19) |
| Commission Rate (%) | −0.05*** | −0.06*** | −0.04* | −0.04* |
| | (−3.00) | (−3.08) | (−1.96) | (−1.91) |
| Trading Volume, as % of TNA | 0.0001** | 0.0001** | 0.0001** | 0.0001** |
| | (2.28) | (2.29) | (2.30) | (2.37) |
| Size Percentile | 0.001 | 0.001 | 0.001 | 0.001 |
| | (1.32) | (1.34) | (1.30) | (1.32) |
| Value Percentile | −0.001*** | −0.001*** | −0.001*** | −0.001*** |
| | (−2.70) | (−2.71) | (−2.72) | (−2.74) |
| Momentum Percentile | −0.001** | −0.001** | −0.001** | −0.001** |
| | (−2.00) | (−2.05) | (−1.97) | (−2.03) |
| Time Fixed Effects | Yes | Yes | Yes | Yes |
| Fund Fixed Effects | Yes | Yes | Yes | Yes |
| Observations | 54,331 | 54,331 | 54,331 | 54,331 |
| Adjusted $R^2$ | 0.11 | 0.11 | 0.11 | 0.10 |

**Table 7**–*Continued*

Panel B: Triple Interaction

| *Dependent variable:* | Return Gap (%) | | | |
|---|---|---|---|---|
| *Broker Incentive:* | $\mathbb{1}$(Dollar Commission $> Q_3$) | | log(Dollar Commission) | |
| | (1) | (2) | (3) | (4) |
| Degree Centrality × Broker Incentive × $\mathbb{1}$(Outflow > 5%) | 0.28** | | 0.05* | |
| | (2.08) | | (1.86) | |
| Eigenvector Centrality × Broker Incentive × $\mathbb{1}$(Outflow > 5%) | | 0.12*** | | 0.02** |
| | | (2.74) | | (2.00) |
| Degree Centrality × Broker Incentive | 0.14 | | 0.01 | |
| | (1.52) | | (0.62) | |
| Degree Centrality × $\mathbb{1}$(Outflow > 5%) | 0.11* | | 0.30*** | |
| | (1.90) | | (3.54) | |
| Eigenvector Centrality × Broker Incentive | | −0.004 | | −0.004 |
| | | (−0.14) | | (−0.64) |
| Eigenvector Centrality × $\mathbb{1}$(Outflow > 5%) | | 0.02 | | 0.09*** |
| | | (1.27) | | (3.30) |
| Broker Incentive × $\mathbb{1}$(Outflow > 5%) | −0.06** | −0.08*** | −0.01*** | −0.01*** |
| | (−2.24) | (−2.82) | (−2.63) | (−2.65) |
| Degree Centrality | −0.01 | | 0.05 | |
| | (−0.15) | | (0.69) | |
| Eigenvector Centrality | | 0.01 | | −0.0001 |
| | | (0.61) | | (−0.01) |
| Broker Incentive | −0.04* | −0.01 | −0.01 | −0.002 |
| | (−1.95) | (−0.39) | (−1.29) | (−0.48) |
| $\mathbb{1}$(Outflow > 5%) | −0.02** | −0.01 | −0.06*** | −0.06*** |
| | (−1.98) | (−1.40) | (−3.70) | (−3.46) |
| log(Fund TNA) | −0.03*** | −0.03*** | −0.03*** | −0.03*** |
| | (−8.63) | (−8.54) | (−6.19) | (−6.02) |
| log(Family TNA) | 0.01*** | 0.01*** | 0.01*** | 0.01** |
| | (2.62) | (2.61) | (2.58) | (2.51) |
| Expense Ratio (%) | 0.02 | 0.02 | 0.02 | 0.02 |
| | (1.24) | (1.33) | (1.18) | (1.26) |
| Commission Rate (%) | −0.05*** | −0.06*** | −0.04** | −0.04* |
| | (−3.00) | (−3.12) | (−1.98) | (−1.96) |
| Trading Volume, as % of TNA | 0.0001** | 0.0001** | 0.0001** | 0.0001** |
| | (2.37) | (2.41) | (2.39) | (2.48) |
| Size Percentile | 0.001 | 0.001 | 0.001 | 0.001 |
| | (1.31) | (1.32) | (1.31) | (1.33) |
| Value Percentile | −0.001*** | −0.001*** | −0.001*** | −0.001*** |
| | (−2.69) | (−2.73) | (−2.71) | (−2.75) |
| Momentum Percentile | −0.001** | −0.001** | −0.001** | −0.001** |
| | (−2.06) | (−2.12) | (−2.08) | (−2.14) |
| Time Fixed Effects | Yes | Yes | Yes | Yes |
| Fund Fixed Effects | Yes | Yes | Yes | Yes |
| Observations | 54,331 | 54,331 | 54,331 | 54,331 |
| Adjusted $R^2$ | 0.11 | 0.11 | 0.11 | 0.11 |

## **Table 8:** The Fund–Centrality Premium For Relationship Clients

This table examines whether the fund–centrality premium is larger for the clients that have established trading relationships with their brokers, especially when the client funds are forced to trade to accommodate large investor redemptions. In unconditional tests presented in Panel A, we interact a measure of existing trading relationships with brokerage network centrality in our baseline specification as follows:

$$\begin{aligned} \textit{Return Gap}_{i,t} &= \delta \times \textit{Centrality}_{i,t-1} \times \textit{Trading Relationship}_{i,t-1} + \beta \times \textit{Centrality}_{i,t-1} \\ &+ \rho \times \textit{Trading Relationship}_{i,t-1} + \gamma \times \textit{Covariates}_{i,t-1} + \alpha_i + \theta_t + \varepsilon_{i,t} \end{aligned}$$

where $\textit{Trading Relationship}_{i,t-1}$, or simply, $\textit{Relationship}_{i,t-1}$ is our proxy for fund $i$'s strength of trading relationships with its current set of brokers, as measured by taking the minimum of a fraction of fund $i$' commissions paid to its broker $k$ during half-year $t-1$ (current) and that during $t-3$ (a year before) and then summing it over all brokers currently employed by the fund. Intuitively, $\textit{Relationship}_{i,t-1}$ measures the extent to which fund $i$'s current set of brokers overlap with the set of brokers the fund traded through a year before. The rest of the model is the same as in Table 5. The independent variable of interest is $\textit{Centrality}_{i,t-1} \times \textit{Relationship}_{i,t-1}$ to tease out the effect of prior trading relationships on the fund–centrality premium. In conditional tests presented in Panel B, we add an indicator variable for contemporaneous large outflows as an additional interaction term. Standard errors are clustered at the fund level 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:* | Return Gap (%) | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Degree Centrality × Relationship | 0.13 | 0.12 | | |
| | (1.42) | (1.15) | | |
| Eigenvector Centrality × Relationship | | | 0.05* | 0.05* |
| | | | (1.82) | (1.69) |
| Degree Centrality | 0.07 | 0.03 | | |
| | (1.15) | (0.50) | | |
| Eigenvector Centrality | | | 0.01 | 0.001 |
| | | | (0.66) | (0.06) |
| Relationship | −0.01 | −0.02 | −0.01 | −0.03 |
| | (−0.55) | (−1.07) | (−0.86) | (−1.51) |
| log(Fund TNA) | −0.01*** | −0.03*** | −0.01*** | −0.03*** |
| | (−6.98) | (−9.71) | (−6.99) | (−9.72) |
| log(Family TNA) | 0.01*** | 0.01*** | 0.01*** | 0.01*** |
| | (6.18) | (2.59) | (6.07) | (2.61) |
| Expense Ratio (%) | −0.01 | 0.02 | −0.01 | 0.02 |
| | (−1.08) | (1.21) | (−1.04) | (1.25) |
| Commission Rate (%) | −0.04*** | −0.07*** | −0.04*** | −0.07*** |
| | (−2.99) | (−4.00) | (−3.03) | (−4.04) |
| Trading Volume, as % of TNA | 0.0000 | 0.0001* | 0.0000 | 0.0001* |
| | (0.73) | (1.92) | (0.74) | (1.93) |
| Size Percentile | −0.001*** | 0.001 | −0.001*** | 0.001 |
| | (−4.09) | (1.35) | (−4.09) | (1.38) |
| Value Percentile | −0.002*** | −0.001*** | −0.002*** | −0.001*** |
| | (−10.64) | (−2.72) | (−10.63) | (−2.73) |
| Momentum Percentile | −0.001** | −0.001** | −0.001** | −0.001** |
| | (−2.29) | (−2.05) | (−2.30) | (−2.06) |
| Time Fixed Effects | Yes | Yes | Yes | Yes |
| Fund Fixed Effects | No | Yes | No | Yes |
| Observations | 54,331 | 54,331 | 54,331 | 54,331 |
| Adjusted $R^2$ | 0.07 | 0.10 | 0.07 | 0.10 |

**Table 8**–*Continued*

Panel B: Triple Interaction

| *Dependent variable:* | Return Gap (%) | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Degree Centrality × Relationship × $\mathbb{1}$(Outflow > 5%) | 0.43** | 0.44** | | |
| | (2.12) | (2.06) | | |
| Eigenvector Centrality × Relationship × $\mathbb{1}$(Outflow > 5%) | | | 0.12* | 0.11* |
| | | | (1.87) | (1.66) |
| Degree Centrality × Relationship | −0.01 | −0.04 | | |
| | (−0.08) | (−0.33) | | |
| Degree Centrality × $\mathbb{1}$(Outflow > 5%) | −0.09 | −0.06 | | |
| | (−0.77) | (−0.45) | | |
| Eigenvector Centrality × Relationship | | | 0.01 | 0.01 |
| | | | (0.40) | (0.35) |
| Eigenvector Centrality × $\mathbb{1}$(Outflow > 5%) | | | −0.02 | −0.01 |
| | | | (−0.50) | (−0.24) |
| Relationship × $\mathbb{1}$(Outflow > 5%) | −0.08** | −0.10*** | −0.08** | −0.09** |
| | (−2.32) | (−2.68) | (−2.17) | (−2.37) |
| Degree Centrality | 0.10 | 0.05 | | |
| | (1.40) | (0.68) | | |
| Eigenvector Centrality | | | 0.02 | 0.004 |
| | | | (0.82) | (0.17) |
| Relationship | 0.02 | 0.02 | 0.01 | 0.01 |
| | (0.88) | (0.75) | (0.58) | (0.26) |
| $\mathbb{1}$(Outflow > 5%) | 0.01 | 0.02 | 0.01 | 0.02 |
| | (0.72) | (0.99) | (0.57) | (0.86) |
| log(Fund TNA) | −0.01*** | −0.03*** | −0.01*** | −0.03*** |
| | (−6.94) | (−9.77) | (−6.94) | (−9.71) |
| log(Family TNA) | 0.01*** | 0.01*** | 0.01*** | 0.01*** |
| | (6.12) | (2.59) | (6.02) | (2.60) |
| Expense Ratio (%) | −0.005 | 0.02 | −0.005 | 0.02 |
| | (−0.92) | (1.27) | (−0.86) | (1.33) |
| Commission Rate (%) | −0.04*** | −0.07*** | −0.04*** | −0.07*** |
| | (−2.96) | (−3.98) | (−3.01) | (−4.03) |
| Trading Volume, as % of TNA | 0.0000 | 0.0001* | 0.0000 | 0.0001** |
| | (0.88) | (1.96) | (0.91) | (1.99) |
| Size Percentile | −0.001*** | 0.001 | −0.001*** | 0.001 |
| | (−4.09) | (1.37) | (−4.09) | (1.39) |
| Value Percentile | −0.002*** | −0.001*** | −0.002*** | −0.001*** |
| | (−10.65) | (−2.69) | (−10.68) | (−2.71) |
| Momentum Percentile | −0.001** | −0.001** | −0.001** | −0.001** |
| | (−2.36) | (−2.12) | (−2.38) | (−2.14) |
| Time Fixed Effects | Yes | Yes | Yes | Yes |
| Fund Fixed Effects | No | Yes | No | Yes |
| Observations | 54,331 | 54,331 | 54,331 | 54,331 |
| Adjusted $R^2$ | 0.07 | 0.11 | 0.07 | 0.11 |

**Table 9:** The Fund–Centrality Premium When Funds Submit Uninformed Large Orders

This table attempts to generalize our main results in Table 6 by examining whether the fund–centrality premium is larger when funds' trading activities are primarily driven by liquidity reasons, for instance, when funds submit large uninformed orders. First, we identify periods of heavy information-motivated buying and selling activities following Alexander, Cici, and Gibson (2007). We calculate $BF$ and $SF$ metrics as follows:

$$BF_{i,t} = \frac{BUYS_{i,t} - FLOW_{i,t}}{TNA_{i,t-1}} \quad \& \quad SF_{i,t} = \frac{SELLS_{i,t} + FLOW_{i,t}}{TNA_{i,t-1}}$$

where $BUYS_{i,t}$ is fund $i$'s dollar volume of stock purchases during half-year $t$, $SELLS_{i,t}$ is fund $i$'s dollar volume of stock sales during half-year $t$, $FLOW_{i,t}$ is fund $i$'s net investor flow (inflow minus outflow) during half-year $t$, and $TNA_{i,t-1}$ is fund $i$'s total net assets at the end of half-year $t-1$. Exploiting within-fund variation in $BF$ and $SF$ metrics, Alexander, Cici, and Gibson (2007) show that buy (sell) portfolios with high $BF$ ($SF$) tend to outperform buy (sell) portfolios with low $BF$ ($SF$). Since we cannot separately evaluate trading performance associated with buys and sells, we assign half-years where both $BF$ and $SF$ fall below its respective top quartile value as periods of uninformed trading. In Panel A, we interact an indicator variable for period of uninformed trading with brokerage network centrality as follows:

$$\begin{aligned} \mathit{Return\ Gap}_{i,t} &= \delta \times \mathit{Centrality}_{i,t-1} \times \mathbb{1}(BF_{i,t} < Q_3 \ \&\ SF_{i,t} < Q_3) + \beta \times \mathit{Centrality}_{i,t-1} \\ &+ \rho \times \mathbb{1}(BF_{i,t} < Q_3 \ \&\ SF_{i,t} < Q_3) + \gamma \times \mathit{Covariates}_{i,t-1} + \alpha_i + \theta_t + \varepsilon_{i,t} \end{aligned}$$

where $\mathbb{1}(BF_{i,t} < Q_3 \ \&\ SF_{i,t} < Q_3)$ is an indicator variable that is equal to 1 if both $BF_{i,t}$ and $SF_{i,t}$ fall below its respective top quartile value during half-year $t$ and the rest of the model is the same as in Table 5. Next, we proxy for average order sizes using average trade sizes inferred from consecutive portfolio disclosures, adjusting for trading volume in the market as follows:

$$\overline{\mathit{Trade\ Size}}_{i,t} = \frac{1}{N_{i,t}} \sum_k \frac{|\ Shares_{i,k,t} - Shares_{i,k,t-1}\ |}{\overline{VOL}^{CRSP}_{k,t}}$$

where $Shares_{i,k,t}$ is the split-adjusted number of shares held in stock $k$ by fund $i$ at the end of half-year (or quarter) $t$, $\overline{VOL}^{CRSP}_{k,t}$ is the average CRSP monthly volume between portfolio disclosures, and the averages are taken over stocks for which $Shares_{i,k,t} \neq Shares_{i,k,t-1}$. To arrive at the semi-annual figure, we take the average of quarterly numbers, if two quarterly observations are available. In Panel B, we add as an additional interaction term $\overline{\mathit{Trade\ Size}}_{i,t}$ as an indicator variable that is equal to 1 if $\overline{\mathit{Trade\ Size}}_{i,t}$ is above its quartile value or as a continuous variable to examine whether the fund–centrality premium is larger when funds submit uninformed large orders. Standard errors are clustered at the fund level and the resulting t-statistics are reported in parentheses. Statistical significance at the 10%, 5%, and 1% level is indicated by *, **, and ***, respectively.

**Table 9**–*Continued*

| *Dependent variable:* | Return Gap (%) | |
| --- | --- | --- |
| | (1) | (2) |
| Degree Centrality × $\mathbb{1}(\text{BF} < Q_3 \ \& \ \text{SF} < Q_3)$ | 0.17*** | |
| | (2.85) | |
| Eigenvector Centrality × $\mathbb{1}(\text{BF} < Q_3 \ \& \ \text{SF} < Q_3)$ | | 0.04** |
| | | (2.04) |
| Degree Centrality | −0.01 | |
| | (−0.12) | |
| Eigenvector Centrality | | 0.004 |
| | | (0.19) |
| $\mathbb{1}(\text{BF} < Q_3 \ \& \ \text{SF} < Q_3)$ | −0.02 | −0.01 |
| | (−1.44) | (−0.79) |
| log(Fund TNA) | −0.03*** | −0.03*** |
| | (−9.90) | (−9.89) |
| log(Family TNA) | 0.01*** | 0.01*** |
| | (2.66) | (2.68) |
| Expense Ratio (%) | 0.02 | 0.02 |
| | (1.20) | (1.22) |
| Commission Rate (%) | −0.07*** | −0.07*** |
| | (−3.98) | (−4.00) |
| Trading Volume, as % of TNA | 0.0001** | 0.0001** |
| | (2.26) | (2.25) |
| Size Percentile | 0.001 | 0.001 |
| | (1.36) | (1.35) |
| Value Percentile | −0.001*** | −0.001*** |
| | (−2.63) | (−2.68) |
| Momentum Percentile | −0.001* | −0.001* |
| | (−1.89) | (−1.93) |
| Time Fixed Effects | Yes | Yes |
| Fund Fixed Effects | Yes | Yes |
| Observations | 54,331 | 54,331 |
| Adjusted $R^2$ | 0.11 | 0.11 |

**Table 9**–*Continued*

| *Dependent variable:* | Return Gap (%) | | | |
|---|---|---|---|---|
| *Brokerage Network Centrality:* | Degree Centrality | | Eigenvector Centrality | |
| | (1) | (2) | (3) | (4) |
| Centrality × $\mathbb{1}(\text{BF} < Q_3\ \&\ \text{SF} < Q_3) \times \mathbb{1}(\overline{\text{Trade Size}} > Q_3)$ | 0.31** | | 0.08* | |
| | (2.11) | | (1.67) | |
| Centrality × $\mathbb{1}(\text{BF} < Q_3\ \&\ \text{SF} < Q_3) \times \overline{\text{Trade Size}}$ | | 0.13* | | 0.04 |
| | | (1.68) | | (1.44) |
| Centrality × $\mathbb{1}(\text{BF} < Q_3\ \&\ \text{SF} < Q_3)$ | 0.09 | 0.09 | 0.02 | 0.02 |
| | (1.36) | (1.28) | (0.91) | (0.75) |
| $\mathbb{1}(\text{BF} < Q_3\ \&\ \text{SF} < Q_3) \times \mathbb{1}(\overline{\text{Trade Size}} > Q_3)$ | −0.05* | | −0.05 | |
| | (−1.91) | | (−1.56) | |
| $\mathbb{1}(\text{BF} < Q_3\ \&\ \text{SF} < Q_3) \times \overline{\text{Trade Size}}$ | | −0.02 | | −0.02 |
| | | (−1.16) | | (−1.01) |
| Centrality × $\mathbb{1}(\overline{\text{Trade Size}} > Q_3)$ | −0.13 | | −0.07 | |
| | (−0.93) | | (−1.51) | |
| Centrality × $\overline{\text{Trade Size}}$ | | −0.02 | | −0.02 |
| | | (−0.26) | | (−0.93) |
| Centrality | 0.02 | 0.002 | 0.02 | 0.02 |
| | (0.36) | (0.03) | (0.97) | (0.77) |
| $\mathbb{1}(\text{BF} < Q_3\ \&\ \text{SF} < Q_3)$ | −0.002 | −0.01 | 0.002 | 0.0001 |
| | (−0.18) | (−0.47) | (0.19) | (0.01) |
| $\mathbb{1}(\overline{\text{Trade Size}} > Q_3)$ | 0.02 | | 0.04 | |
| | (0.74) | | (1.27) | |
| $\overline{\text{Trade Size}}$ | | −0.01 | | 0.0002 |
| | | (−0.71) | | (0.01) |
| log(Fund TNA) | −0.03*** | −0.03*** | −0.03*** | −0.03*** |
| | (−9.83) | (−8.68) | (−9.86) | (−8.73) |
| log(Family TNA) | 0.01*** | 0.01*** | 0.01*** | 0.01*** |
| | (2.67) | (2.75) | (2.68) | (2.75) |
| Expense Ratio (%) | 0.02 | 0.02 | 0.02 | 0.02 |
| | (1.18) | (1.23) | (1.20) | (1.26) |
| Commission Rate (%) | −0.07*** | −0.07*** | −0.07*** | −0.07*** |
| | (−3.98) | (−3.92) | (−4.01) | (−3.93) |
| Trading Volume, as % of TNA | 0.0001** | 0.0001** | 0.0001** | 0.0001** |
| | (2.28) | (2.42) | (2.26) | (2.40) |
| Size Percentile | 0.001 | 0.001 | 0.001 | 0.001 |
| | (1.30) | (0.99) | (1.37) | (1.09) |
| Value Percentile | −0.001*** | −0.001** | −0.001*** | −0.001*** |
| | (−2.64) | (−2.57) | (−2.67) | (−2.58) |
| Momentum Percentile | −0.001* | −0.001** | −0.001* | −0.001** |
| | (−1.92) | (−2.00) | (−1.94) | (−2.00) |
| Time Fixed Effects | Yes | Yes | Yes | Yes |
| Fund Fixed Effects | Yes | Yes | Yes | Yes |
| Observations | 54,331 | 54,331 | 54,331 | 54,331 |
| Adjusted $R^2$ | 0.11 | 0.11 | 0.11 | 0.11 |

**Table 10:** A List of Brokerage Mergers (1995-2015)

This table reports a list of twenty six brokerage mergers, including the names of brokers involved in the merger, the merger effective date, the average brokerage shares pre- and post-merger, and changes in average broker shares around the merger. A broker share is defined as a fraction of the commission payments to the given broker by the fund. Broker shares are first averaged across funds each month on a rolling basis and then averaged over months $t-18$ to $t-7$ for the pre-merger and over months $t+7$ and $t+18$ for the post-merger. We highlight five largest mergers that will be used in our natural experiment.

| | Acquiring Broker | | | | Acquired Broker | | | |
|---|---|---|---|---|---|---|---|---|
| | | Average Broker Shares (%) | | | | Average Broker Shares (%) | | |
| Effective Date | Broker Name | Before | After | Change | Broker Name | Before | After | Change |
| 1997-05-31 | MORGAN STANLEY | 4.76 | 5.65 | 0.89 | DEAN WITTER REYNOLDS | 1.47 | 0.57 | −0.90 |
| 1997-09-02 | BT NEW YORK (SUCCESSOR: DEUTSCHE) | 0.28 | 0.44 | 0.16 | ALEX BROWN | 1.04 | 1.16 | 0.12 |
| 1997-11-28 | SMITH BARNEY (TRAVELERS) | 4.83 | 5.69 | 0.86 | SALOMON BROTHERS | 3.94 | 0.78 | −3.16 |
| 1998-06-30 | SOCIETE GENERALE SECURITIES | 0.18 | 0.18 | −0.004 | COWEN | 0.54 | 0.66 | 0.12 |
| 2000-02-24 | INSTINET | 3.28 | 2.67 | −0.61 | LYNCH JONES RYAN | 0.42 | 0.35 | −0.07 |
| 2000-11-02 | GOLDMAN SACHS GROUP | 5.72 | 7.23 | 1.52 | SPEAR LEEDS KELLOGG | 0.22 | 0.35 | 0.12 |
| **2000-11-03** | **CREDIT SUISSE FIRST BOSTON** | 4.02 | 6.40 | 2.38 | **DONALDSON LUFKIN JENRETTE** | 4.40 | 0.75 | −3.65 |
| **2000-11-03** | **UBS WARBURG DILLON READ** | 2.12 | 4.31 | 2.20 | **PAINE WEBBER** | 3.75 | 0.89 | −2.85 |
| 2001-04-30 | ABN-AMRO | 1.15 | 0.77 | −0.39 | ING BARING-US | 7.27 | 8.83 | 1.56 |
| 2001-09-04 | WACHOVIA | 0.41 | 0.50 | 0.09 | FIRST UNION CAPITAL MARKETS | 0.18 | 0.15 | −0.03 |
| 2002-02-04 | BANK OF NEW YORK | 0.08 | 0.27 | 0.19 | AUTRANET | 1.02 | 0.44 | −0.58 |
| 2003-07-01 | WACHOVIA | 0.47 | 0.86 | 0.40 | PRUDENTIAL | 1.30 | 1.05 | −0.25 |
| 2003-10-31 | LEHMAN BROTHERS | 5.99 | 7.33 | 1.34 | NEUBERGER BERMAN | 0.14 | 0.02 | −0.12 |
| 2003-12-08 | UBS AG | 5.69 | 5.11 | −0.58 | ABN-AMRO | 0.90 | 1.14 | 0.24 |
| 2005-03-31 | INSTINET | 1.72 | 1.49 | −0.24 | BRIDGE TRADING | 0.61 | 0.22 | −0.39 |
| 2007-02-02 | NOMURA HOLDINGS | 0.23 | 0.20 | −0.03 | INSTINET | 1.39 | 2.26 | 0.87 |
| 2007-10-01 | WACHOVIA | 0.26 | 0.13 | −0.13 | A.G. EDWARDS SONS | 0.28 | 0.004 | −0.28 |
| **2008-05-30** | **JPMORGAN CHASE** | 4.14 | 7.83 | 3.69 | **BEAR STEARNS** | 4.63 | 0.17 | −4.46 |
| **2008-09-22** | **BARCLAYS** | 0.04 | 3.02 | 2.98 | **LEHMAN BROTHERS** | 7.53 | 0.12 | −7.41 |
| **2009-01-01** | **BANK OF AMERICA** | 0.96 | 1.09 | 0.13 | **MERRILL LYNCH** | 8.69 | 6.23 | −2.45 |
| 2009-10-02 | MACQUARIE GROUP | 0.42 | 0.69 | 0.27 | FOX PITT KELTON | 0.09 | 0.002 | −0.09 |
| 2009-12-31 | WELLS FARGO SECURITIES | 0.04 | 0.16 | 0.12 | WACHOVIA | 0.12 | 0.11 | −0.01 |
| 2010-07-01 | STIFEL | 0.52 | 0.60 | 0.08 | THOMAS WEISEL PARTNERS | 0.20 | 0.02 | −0.18 |
| 2012-04-02 | RAYMOND JAMES FINANCIAL | 0.37 | 0.44 | 0.07 | MORGAN KEEGAN | 0.27 | 0.15 | −0.12 |
| 2013-02-15 | STIFEL | 0.63 | 0.73 | 0.10 | KEEFE BRUYETTE WOODS | 0.22 | 0.15 | −0.07 |
| 2014-09-03 | KEYBANK | 0.09 | 0.11 | 0.03 | PACIFIC CREST SECURITIES | 0.04 | 0.01 | −0.03 |

## **Table 11:** Testing for Matching Balance

This table reports the cross-sectional means and differences in means of the pre-treatment outcome variables and other pre-event fund-characteristics for the treated mutual funds and (matched) controls before and after the matching. We take top ten percent of funds with the largest expected changes in *Degree Centrality* as the treatment group. Among the remaining 90% of the sample, we construct the control group by matching on pre-treatment (pre-event) outcome variables and fund characteristics, using *Genetic Matching* algorithm proposed by Diamond and Sekhon (2013). The pre-treatment outcome variables include *Degree Centrality* and *Return Gap* and pre-event fund characteristics include *Log(Fund TNA)*, *Expense Ratio*, *Commission Rate*, *Trading Volume, as % of TNA*, *Index Fund (Yes=1)*, *Size Percentile*, *Value Percentile*, and *Momentum Percentile*. We choose one year just prior to the event window as the pre-event period. Return gaps are averaged over the twelve months in the pre-event period and mid-point values are taken for other variables. The event timelines are as depicted in Figure 4. The t-statistics are reported in parentheses. Statistical significance at the 10%, 5%, and 1% level is indicated by *, **, and ***, respectively.

| | | Before Matching | | | After Matching | | |
|---|---|---|---|---|---|---|---|
| Variable | Treated | Control | Difference | (p-value) | Control | Difference | (p-value) |
| Panel A: 2000 Brokerage Mergers (Number of treated funds = 102) | | | | | | | |
| *Pre-treatment outcomes* | | | | | | | |
| Degree Centrality | 0.21 | 0.16 | 0.05*** | (< 0.001) | 0.20 | 0.001 | (0.49) |
| Return Gap (%) | −0.10 | −0.08 | −0.02 | (0.84) | −0.11 | 0.01 | (0.54) |
| *Covariates* | | | | | | | |
| log(Fund TNA) | 5.67 | 5.47 | 0.19 | (0.35) | 5.99 | −0.33 | (0.18) |
| Expense Ratio (%) | 0.01 | 0.01 | −0.001*** | (0.01) | 0.01 | −0.0001 | (0.60) |
| Commission Rate (%) | 0.12 | 0.14 | −0.01 | (0.42) | 0.14 | −0.01 | (0.57) |
| Trading Volume, as % of TNA | 160.99 | 174.71 | −13.72 | (0.24) | 178.77 | −17.78 | (0.28) |
| Index Fund (Yes=1) | 0.07 | 0.04 | 0.03 | (0.33) | 0.04 | 0.03 | (0.32) |
| Size Percentile | 90.20 | 88.03 | 2.17** | (0.04) | 88.77 | 1.43 | (0.22) |
| Value Percentile | 26.92 | 29.75 | −2.83** | (0.02) | 26.86 | 0.06 | (0.82) |
| Momentum Percentile | 66.37 | 64.42 | 1.95 | (0.13) | 66.69 | −0.32 | (0.77) |
| Panel B: 2008 Brokerage Mergers (Number of treated funds = 160) | | | | | | | |
| *Pre-treatment outcomes* | | | | | | | |
| Degree Centrality | 0.17 | 0.15 | 0.01*** | (< 0.001) | 0.17 | −0.001 | (0.63) |
| Return Gap(%) | 0.15 | 0.14 | 0.01 | (0.49) | 0.15 | 0.003 | (0.62) |
| *Covariates* | | | | | | | |
| log(Fund TNA) | 6.35 | 5.82 | 0.53*** | (< 0.001) | 6.23 | 0.12 | (0.18) |
| Expense Ratio(%) | 0.01 | 0.01 | −0.0005* | (0.09) | 0.01 | −0.0002 | (0.44) |
| Commission Rate(%) | 0.08 | 0.22 | −0.14*** | (< 0.001) | 0.08 | −0.003 | (0.35) |
| Trade Volume, as % of TNA | 130.52 | 106.52 | 23.99** | (0.01) | 126.83 | 3.69 | (0.21) |
| Index Fund (Yes=1) | 0.11 | 0.10 | 0.003 | (0.91) | 0.14 | −0.03 | (0.20) |
| Size Percentile | 85.26 | 83.84 | 1.42 | (0.12) | 85.92 | −0.66 | (0.19) |
| Value Percentile | 40.23 | 41.35 | −1.12 | (0.17) | 40.03 | 0.19 | (0.40) |
| Momentum Percentile | 60.89 | 60.69 | 0.20 | (0.76) | 61.15 | −0.26 | (0.48) |

**Table 12:** Do brokerage networks improve trading performance? DiD Results

This table reports the difference-in-differences (DiD) results for *Degree Centrality* and *Return Gap* before and after brokerage mergers for the treated mutual funds and their matched controls. The selection of treatment and control groups, the matching procedure, and the construction of pre-event outcome variables are the same as in Table 11. We choose one year immediately following the event window as the post-event period. Return gaps are averaged over the twelve months in the post-event period and mid-point values are taken for *Degree Centrality*. The event timelines are as depicted in Figure 4. If we denote the average outcome variables in the treatment (T) and control (C) groups in the pre- and post-event periods by $O_{T,1}$, $O_{T,2}$, $O_{C,1}$, and $O_{C,2}$, respectively, the partial effect of change due to the mergers can be estimated as

$$DiD = (O_{T,2} - O_{T,1}) - (O_{C,2} - O_{C,1}).$$

The t-statistics are reported in parentheses. Statistical significance at the 10%, 5%, and 1% level is indicated by *, **, and ***, respectively.

| | Treated | | Matched Control | | DiD | |
|---|---|---|---|---|---|---|
| Outcome Measures | Before | After | Before | After | Mean | (t-stat) |
| Panel A: 2000 Brokerage Mergers | | | | | | |
| Degree Centrality | 0.206 | 0.235 | 0.205 | 0.222 | 0.013** | (2.01) |
| Return Gap (%) | −0.097 | 0.063 | −0.106 | −0.038 | 0.093* | (1.67) |
| Panel B: 2008 Brokerage Mergers | | | | | | |
| Degree Centrality | 0.167 | 0.193 | 0.168 | 0.160 | 0.034*** | (8.39) |
| Return Gap (%) | 0.150 | 0.104 | 0.147 | 0.033 | 0.068* | (1.87) |

# Appendix

**Table A1:** Sample of CRSP-Thomson-NSAR Matched Funds

This table reports the total number and aggregate total net assets (TNA) of our CRSP-Thomson-NSAR matched funds each half-year.

| Year | First Half | | Second Half | |
|---|---|---|---|---|
| | Total Number of Funds | Aggregate TNA ($ billion) | Total Number of Funds | Aggregate TNA ($ billion) |
| 1994 | 421 | 243.2 | 512 | 310.7 |
| 1995 | 617 | 426.6 | 759 | 551.2 |
| 1996 | 812 | 632.2 | 868 | 759.6 |
| 1997 | 941 | 917.2 | 991 | 1,142.8 |
| 1998 | 1,065 | 1,356.2 | 1,122 | 1,559.0 |
| 1999 | 1,260 | 1,916.8 | 1,306 | 1,999.6 |
| 2000 | 1,426 | 2,162.4 | 1,476 | 2,057.5 |
| 2001 | 1,572 | 2,047.9 | 1,608 | 1,847.8 |
| 2002 | 1,679 | 1,834.2 | 1,773 | 1,608.0 |
| 2003 | 1,774 | 1,702.2 | 1,803 | 2,059.3 |
| 2004 | 1,825 | 2,293.2 | 1,845 | 2,298.6 |
| 2005 | 1,845 | 2,426.4 | 1,897 | 2,610.3 |
| 2006 | 1,950 | 2,811.4 | 2,039 | 2,947.2 |
| 2007 | 2,106 | 3,166.0 | 2,137 | 3,227.0 |
| 2008 | 2,127 | 2,852.7 | 2,089 | 2,224.0 |
| 2009 | 2,017 | 1,998.4 | 2,012 | 2,415.2 |
| 2010 | 1,968 | 2,480.1 | 1,921 | 2,567.0 |
| 2011 | 1,891 | 2,931.1 | 1,852 | 2,622.4 |
| 2012 | 1,823 | 2,798.1 | 1,757 | 2,847.8 |
| 2013 | 1,730 | 3,115.2 | 1,684 | 3,685.1 |
| 2014 | 1,649 | 4,143.0 | 1,629 | 4,072.0 |
| 2015 | 1,614 | 4,137.0 | 1,568 | 4,132.1 |
| 2016 | 1,565 | 3,880.6 | | |

## **Table A2:** When Funds Experience Severe Redemptions: Robustness Checks

This table provides robustness checks for the results reported in Table 6. In Panel A, we use a different cutoff (10% instead of 5%) to identify large outflow events and the rest of the model is the same as in Table 6. In Panel B, we repeat our analysis in Table 6 using a sub-sample of funds with fund family TNA is below its top quartile value. Standard errors are clustered at the fund level 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: Using a larger cutoff to define large outflow events

| *Dependent variable:* | Return Gap (%) | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Degree Centrality × 𝟙(Outflow > 10%) | 0.22*** | 0.24*** | | |
| | (3.60) | (3.53) | | |
| Eigenvector Centrality × 𝟙(Outflow > 10%) | | | 0.07*** | 0.07*** |
| | | | (3.93) | (3.42) |
| Degree Centrality | 0.11*** | 0.05 | | |
| | (3.31) | (1.07) | | |
| Eigenvector Centrality | | | 0.03*** | 0.01 |
| | | | (3.09) | (1.04) |
| 𝟙(Outflow > 10%) | −0.04*** | −0.04*** | −0.05*** | −0.04*** |
| | (−3.76) | (−3.30) | (−4.00) | (−3.17) |
| log(Fund TNA) | −0.01*** | −0.03*** | −0.01*** | −0.03*** |
| | (−6.86) | (−9.69) | (−6.85) | (−9.65) |
| log(Family TNA) | 0.01*** | 0.01*** | 0.01*** | 0.01*** |
| | (6.15) | (2.58) | (6.08) | (2.59) |
| Expense Ratio (%) | −0.01 | 0.02 | −0.01 | 0.02 |
| | (−1.01) | (1.24) | (−0.95) | (1.28) |
| Commission Rate (%) | −0.04*** | −0.07*** | −0.04*** | −0.07*** |
| | (−2.95) | (−3.99) | (−2.99) | (−4.01) |
| Trading Volume, as % of TNA | 0.0000 | 0.0001** | 0.0000 | 0.0001** |
| | (0.95) | (1.99) | (0.97) | (2.00) |
| Size Percentile | −0.001*** | 0.001 | −0.001*** | 0.001 |
| | (−4.19) | (1.34) | (−4.21) | (1.35) |
| Value Percentile | −0.002*** | −0.001*** | −0.002*** | −0.001*** |
| | (−10.67) | (−2.74) | (−10.71) | (−2.75) |
| Momentum Percentile | −0.001** | −0.001** | −0.001** | −0.001** |
| | (−2.38) | (−2.09) | (−2.39) | (−2.10) |
| Time Fixed Effects | Yes | Yes | Yes | Yes |
| Fund Fixed Effects | No | Yes | No | Yes |
| Observations | 54,331 | 54,331 | 54,331 | 54,331 |
| Adjusted $R^2$ | 0.07 | 0.11 | 0.07 | 0.11 |

**Table A2**–*Continued*

Panel B: Excluding funds that belong to large fund families

| *Dependent variable:* | Return Gap (%) | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Degree Centrality × $\mathbb{1}$(Outflow > 5%) | 0.13** | 0.14** | | |
| | (2.51) | (2.40) | | |
| Eigenvector Centrality × $\mathbb{1}$(Outflow > 5%) | | | 0.05*** | 0.04** |
| | | | (2.77) | (2.23) |
| Degree Centrality | 0.13*** | 0.03 | | |
| | (3.18) | (0.54) | | |
| Eigenvector Centrality | | | 0.03*** | 0.01 |
| | | | (2.87) | (0.46) |
| $\mathbb{1}$(Outflow > 5%) | −0.03*** | −0.03*** | −0.03*** | −0.03*** |
| | (−3.05) | (−2.96) | (−3.26) | (−2.78) |
| log(Fund TNA) | −0.01*** | −0.04*** | −0.01*** | −0.04*** |
| | (−6.10) | (−8.85) | (−6.11) | (−8.81) |
| log(Family TNA) | 0.01*** | 0.01** | 0.01*** | 0.01** |
| | (3.89) | (2.07) | (3.90) | (2.09) |
| Expense Ratio (%) | −0.003 | 0.01 | −0.003 | 0.01 |
| | (−0.49) | (0.34) | (−0.43) | (0.36) |
| Commission Rate (%) | −0.04*** | −0.07*** | −0.04*** | −0.07*** |
| | (−2.84) | (−4.02) | (−2.89) | (−4.05) |
| Trading Volume, as % of TNA | −0.0000 | 0.0001 | −0.0000 | 0.0001 |
| | (−0.27) | (1.00) | (−0.25) | (1.01) |
| Size Percentile | −0.001*** | 0.001 | −0.001*** | 0.001 |
| | (−4.20) | (1.51) | (−4.22) | (1.51) |
| Value Percentile | −0.002*** | −0.002*** | −0.002*** | −0.002*** |
| | (−10.16) | (−3.13) | (−10.19) | (−3.15) |
| Momentum Percentile | −0.001*** | −0.001* | −0.001*** | −0.001* |
| | (−2.63) | (−1.89) | (−2.65) | (−1.91) |
| Time Fixed Effects | Yes | Yes | Yes | Yes |
| Fund Fixed Effects | No | Yes | No | Yes |
| Observations | 40,743 | 40,743 | 40,743 | 40,743 |
| Adjusted $R^2$ | 0.06 | 0.10 | 0.06 | 0.10 |

## **Table A3:** Testing for Matching Balance

This table reports the cross-sectional means and differences in means of the pre-treatment outcome variables and other pre-event fund-characteristics for the treated mutual funds and (matched) controls before and after the matching. We take top ten percent of funds with the largest expected changes in *Eigenvector Centrality* as the treatment group. Among the remaining 90% of the sample, we construct the control group by matching on pre-treatment (pre-event) outcome variables and fund characteristics, using *Genetic Matching* algorithm proposed by Diamond and Sekhon (2013). The pre-treatment outcome variables include *Eigenvector Centrality* and *Return Gap* and pre-event fund characteristics include *Log(Fund TNA)*, *Expense Ratio*, *Commission Rate*, *Trading Volume, as % of TNA*, *Index Fund (Yes=1)*, *Size Percentile*, *Value Percentile*, and *Momentum Percentile*. We choose one year just prior to the event window as the pre-event period. Return gaps are averaged over the twelve months in the pre-event period and mid-point values are taken for other variables. The event timelines are as depicted in Figure 4. The t-statistics are reported in parentheses. Statistical significance at the 10%, 5%, and 1% level is indicated by *, **, and ***, respectively.

| | | Before Matching | | | After Matching | | |
|---|---|---|---|---|---|---|---|
| Variable | Treated | Control | Difference | (p-value) | Control | Difference | (p-value) |
| Panel A: 2000 Brokerage Mergers (Number of treated funds = 102) | | | | | | | |
| *Pre-treatment outcomes* | | | | | | | |
| Eigenvector Centrality | 0.68 | 0.50 | 0.18*** | (< 0.001) | 0.67 | 0.01 | (0.37) |
| Return Gap (%) | −0.03 | −0.09 | 0.05 | (0.54) | −0.03 | −0.01 | (0.82) |
| *Covariates* | | | | | | | |
| log(Fund TNA) | 5.27 | 5.52 | −0.25 | (0.18) | 5.29 | −0.02 | (0.70) |
| Expense Ratio (%) | 0.01 | 0.01 | −0.001 | (0.13) | 0.01 | −0.0001 | (0.78) |
| Commission Rate (%) | 0.18 | 0.13 | 0.06** | (0.04) | 0.18 | 0.003 | (0.31) |
| Trading Volume, as % of TNA | 158.43 | 175.00 | −16.57 | (0.18) | 152.24 | 6.19 | (0.32) |
| Index Fund (Yes=1) | 0.07 | 0.04 | 0.03 | (0.33) | 0.03 | 0.04 | (0.16) |
| Size Percentile | 89.72 | 88.09 | 1.63 | (0.13) | 89.03 | 0.69 | (0.57) |
| Value Percentile | 27.76 | 29.66 | −1.90 | (0.13) | 28.53 | −0.77 | (0.47) |
| Momentum Percentile | 65.69 | 64.49 | 1.19 | (0.36) | 64.89 | 0.80 | (0.30) |
| Panel B: 2008 Brokerage Mergers (Number of treated funds = 161) | | | | | | | |
| *Pre-treatment outcomes* | | | | | | | |
| Eigenvector Centrality | 0.60 | 0.49 | 0.10*** | (< 0.001) | 0.60 | 0.001 | (0.71) |
| Return Gap(%) | 0.17 | 0.13 | 0.03 | (0.11) | 0.16 | 0.01 | (0.41) |
| *Covariates* | | | | | | | |
| log(Fund TNA) | 6.50 | 5.80 | 0.70*** | (< 0.001) | 6.46 | 0.03 | (0.45) |
| Expense Ratio(%) | 0.01 | 0.01 | −0.0005 | (0.11) | 0.01 | 0.0001 | (0.85) |
| Commission Rate(%) | 0.08 | 0.22 | −0.15*** | (< 0.001) | 0.13 | −0.05 | (0.13) |
| Trade Volume, as % of TNA | 132.59 | 106.27 | 26.32*** | 0.003 | 131.14 | 1.45 | (0.53) |
| Index Fund (Yes=1) | 0.11 | 0.10 | 0.01 | (0.73) | 0.12 | −0.01 | (0.82) |
| Size Percentile | 85.10 | 83.85 | 1.25 | (0.16) | 85.64 | −0.54 | (0.66) |
| Value Percentile | 40.08 | 41.37 | −1.29 | (0.12) | 40.18 | −0.10 | (0.60) |
| Momentum Percentile | 60.88 | 60.69 | 0.19 | (0.76) | 61.71 | −0.84 | (0.28) |

**Table A4:** Do brokerage networks improve trading performance? DiD Results

This table reports the difference-in-differences (DiD) results for *Eigenvector Centrality* and *Return Gap* before and after brokerage mergers for the treated mutual funds and their matched controls. The selection of treatment and control groups, the matching procedure, and the construction of pre-event outcome variables are the same as in Table A3. We choose one year immediately following the event window as the post-event period. Return gaps are averaged over the twelve months in the post-event period and mid-point values are taken for *Eigenvector Centrality*. The event timelines are as depicted in Figure 4. If we denote the average outcome variables in the treatment (T) and control (C) groups in the pre- and post-event periods by $O_{T,1}$, $O_{T,2}$, $O_{C,1}$, and $O_{C,2}$, respectively, the partial effect of change due to the mergers can be estimated as

$$DiD = (O_{T,2} - O_{T,1}) - (O_{C,2} - O_{C,1}).$$

The t-statistics are reported in parentheses. Statistical significance at the 10%, 5%, and 1% level is indicated by *, **, and ***, respectively.

| | Treated | | Matched Control | | DiD | |
|---|---|---|---|---|---|---|
| Outcome Measures | Before | After | Before | After | Mean | (t-stat) |
| Panel A: 2000 Brokerage Mergers | | | | | | |
| Eigenvector Centrality | 0.677 | 0.662 | 0.671 | 0.680 | 0.053** | (2.59) |
| Return Gap (%) | −0.034 | 0.026 | −0.029 | −0.076 | 0.108* | (1.68) |
| Panel B: 2008 Brokerage Mergers | | | | | | |
| Eigenvector Centrality | 0.596 | 0.679 | 0.595 | 0.592 | 0.090*** | (5.68) |
| Return Gap (%) | 0.167 | 0.107 | 0.159 | 0.024 | 0.075** | (1.98) |