Information Diffusion in Institutional Investor Networks

Edwin Hu*

October 7, 2015

Job Market Paper[1]

Abstract

This paper finds evidence that institutional investors share valuable information with one another in a manner consistent with a stylized model of network information diffusion. I develop a new measure of information asymmetry between investors resulting from network information diffusion which I call Information Diffusion Centrality. Using my measure, I show that the interim trading performance of central investors is 60% higher than that of the average investor. Furthermore, central investors' round-trip trading performance in target stock around merger announcements is more than double that of the average investor, and is independent of connections to merger advisors. In addition, centrality is distinct from expertise, plays no role in trading performance around sudden deaths, and central investors in a propensity score matched sample have higher trading performance. My findings suggest that the structure of information networks may provide insights into how information is distributed and dispersed in financial markets.

*Jesse H. Jones School of Business at Rice University. I thank Jefferson Duarte, Alex Butler, and Kerry Back for valuable comments and feedback. All errors are my own. Email: eddyhu@gmail.com.

[1]Latest version available at `http://edwinhu.github.io/idc.pdf`

How information diffuses among investors has important implications for the efficiency of financial markets. Shiller (2000) writes, "Word-of-mouth transmission of ideas appears to be an important contributor to day-to-day or hour-to-hour stock market fluctuations..." A recent literature including Colla and Mele (2010), Ozsoylev and Walden (2011), Han and Yang (2013), and Walden (2014) examines the asset pricing implications of trading in "social" or "information" networks. The common assumption in these papers is that investors share information with one another through network connections. However, there is surprisingly little evidence in the literature that informed investors share *valuable* information with one another. Shiller and Pound (1989) finds survey evidence that the majority of institutional investors purchase stocks based on conversations with other investment professionals. Therefore, an important question is whether these conversations contain valuable information, or if conversations consist of "cheap talk."[2]

In this paper I rely on the premise that information is a key determinant of trading performance to provide evidence that institutional investors share valuable information with one another. I show that the cross-sectional variation in institutions' trading performance is related to the way in which information is distributed in a stylized model of network information diffusion. I exploit the fact that even if information arrives uniformly and diffuses randomly through a network, the topological heterogeneity of the network guarantees that information sharing will result in information asymmetry between investors, such that "central" investors have superior access to information. I focus on institutional investors because they compose the majority of US stock trading volume (Boehmer and Kelley, 2009), and are likely to be informed (Hendershott, Livdan, and Schürhoff, 2015). The emphasis on trading performance is important because evidence based on similar portfolio holdings or

[2]As for why informed investors would share information, Stein (2008) argues that in a repeated game it is optimal to share truthfully so that new information may be produced and shared in return. Also, it may be valuable for a fund manager to be able take a position and credibly communicate this to other managers so that their subsequent trades will move the stock price in the right direction.

correlated trading alone does not rule out the possibility of cheap talk.

I begin by developing a new measure of centrality, Information Diffusion Centrality, based on a model in which information diffuses through a generic network of investors as they share information with their immediate neighbors. In the model a single piece of information flows through the network following a symmetric random walk which reflects the bilateral sharing of information. Central investors have superior access to information because they have a high probability of receiving information that is diffusing through the network. Although the model is very simple, it is able to provide a theoretically justifiable measure of information asymmetry between investors resulting from network information diffusion that can be taken to the data. This exercise is important because although many different definitions of information diffusion exist in the finance literature, there is no clear definition of centrality.[3]

To compute centrality empirically, I develop a new methodology to estimate information linkages between investors using daily trades. The theory suggests that information sharing leads to correlated trading (Colla and Mele, 2010), therefore I use empirically observed pairwise trading correlations to construct quarterly Weighted Trading Correlation Network snapshots. Central investors tend to have high trading correlations with many other investors, whereas investors on the periphery tend to have low trading correlations with other investors.

To see if central investors have superior access to information I compute quarterly interim trading performance using fund-level data by tracking the abnormal performance of all stocks bought and sold by a fund from at the trade execution date, using the execution price, until the end of the quarter. Following Puckett and Yan (2011) I take the weighted average of buy and sell trade performance and call the difference the interim trading per-

[3]For more on information diffusion see Hong and Stein (1999); Duffie and Manso (2007); Duffie, Malamud, and Manso (2009); Duffie, Giroux, and Manso (2010); Hong, Hong, and Ungureanu (2011); Manela (2014).

formance. I then regress quarterly interim trading performance on Information Diffusion Centrality. If investors share valuable information through their network connections, then central investors should, on average, have superior trading performance.

Because average trading performance includes many non-information (e.g. liquidity) trades, I shift my attention to a setting in which access to information should be the sole determinant of trading performance. Specifically, I examine the performance of round-trip trades in target stock made around merger announcements. It is well-documented that there are information leakages prior to merger announcements (Betton, Eckbo, and Thorburn, 2008), therefore it is reasonable to expect that some institutions may be trading on private information about these events.[4] Round-trip trading performance is well-suited for measuring informed trading around events because investors trading on short-lived information may want to reverse their positions in order to lock in gains or take advantage of overreactions (Hirshleifer, Subrahmanyam, and Titman, 1994; Brunnermeier, 2005). To see if central investors have superior information about mergers, I regress round-trip trading performance on Information Diffusion Centrality.

A potentially important source of information leakages prior to mergers could be through funds' connections to brokerage houses. Jegadeesh and Tang (2011) find that institutions connected to brokerage houses that serve as merger advisors are net buyers of target stocks and make profitable trades in aggregate. Brokerage house connections are unlikely to explain the dispersion in centrality, but may be contemporaneously correlated to the dispersion in trading performance. Following Jegadeesh and Tang (2011), I estimate funds' connections to brokerage houses who serve as merger advisors. I then regress round-trip trading performance on Information Diffusion Centrality, an indicator for merger advisor connections, and the interaction term which captures the returns to being both central and connected.

[4]There is also evidence that there is insider trading around merger announcements (Meulbroek, 1992; Ahern, 2014), and increased information asymmetry leading up to announcements (Aktas, De Bodt, Declerck, and Van Oppens, 2007; Duarte, Hu, and Young, 2015).

I lack a good instrument for centrality that would allow me to cleanly identify the effects on trading performance. As a result, I rely on three additional empirical tests to provide further indirect evidence that trading performance is driven by information and to rule out alternative hypotheses. First, industry expertise may be related to informed trading (Tookes, 2008; Peress, 2010). I examine round-trip trading around new product announcements where I expect product market expertise to be essential to information processing and trading performance. Second, as a falsification test, I examine sudden deaths of board directors and key executives to see if centrality is spuriously predicting trading performance. Sudden deaths are used as "exogenous shocks" to corporate governance (see for example, Nguyen and Nielsen, 2010), therefore I expect to find no returns to being central in this setting. Finally, I repeat my main regressions within a propensity score matched sample to try and mitigate biases due to endogeneity.

My results are as follows. In regressions of interim trading performance on Information Diffusion Centrality, a one standard deviation increase in Centrality corresponds to a 0.30% (0.34%) increase in principal-weighted (equal-weighted) average excess interim trading performance. To put this estimate in perspective, the average fund's principal-weighted (equal-weighted) average excess trading performance in my sample is 0.52% (0.65%). In other words, the trading performance of "central funds" with one standard deviation higher centrality, is approximately 60% higher than that of the average fund.

Central funds also have a significant information advantage trading around merger announcements. In regressions of round-trip trading performance on Information Diffusion Centrality, a one standard deviation increase in Centrality corresponds to a 1.48% (1.54%) increase in principal-weighted (equal-weighted) average excess round-trip trading performance. Central funds' trading performance is more than double that of the average fund.

Merger advisor connections do not play an important role in trading performance around merger announcements. In regressions of round-trip trading performance on merger advisor connections, connected funds have a 0.38% higher (-0.08% lower) principal-weighted (equal-weighted) average excess round-trip trading performance. The estimates are statistically indistinguishable from zero. In contrast, central funds not connected to merger advisors still have trading performances that are more than double that of the average fund, which suggests that information heard through the "grapevine" may be more valuable than information gained from merger advisors.[5]

Last, I find that centrality is unrelated to product market expertise, plays no role in trading performance around sudden deaths, and predicts trading performance within my propensity score matched sample.

In summary, this paper provides three contributions to the study of information diffusion in financial markets. First, I find evidence consistent with the network diffusion of valuable information among institutional investors. If institutional investors rely on word-of-mouth communication as Shiller and Pound (1989) suggests, then these results may help explain certain behavioral biases in a rational way. Shiller (1995) argues that many behavioral biases may be attributed to the way in which economic agents converse and Hirshleifer (2015) argues that the next step in behavioral finance is to study *social* finance—in other words, to incorporate theories of sociology into our understanding of financial markets. Herding, home bias, and underreaction, three of the most well-documented behavioral biases, are all arguably related to the localization of information within the "community" structure of networks. Second, I introduce Information Diffusion Centrality and show that it predicts trading performance in a variety of settings. In this paper I show that there are subtle but important distinctions between centrality measures of "importance" borrowed from the sociology literature and Information Diffusion Centrality. I show that using measures of importance as proxies for information may lead to different empirical conclusions. Finally,

[5]It could also be that funds connected to merger advisors avoid making "insider" trades because they are aware that their trades may be traced back to their desks (Griffin, Shu, and Topaloglu, 2012).

the weighted network approach used in this paper may be applicable to more general studies of social interactions beyond binary connections. Future empirical network analysis should not throw away potentially valuable information contained in weighted networks.

This paper is related to the growing empirical literature on word-of-mouth effects in investor behavior (Hong, Kubik, and Stein, 2005; Pool, Stoffman, and Yonker, 2015; Ivković and Weisbenner, 2005, 2007; Brown, Ivković, Smith, and Weisbenner, 2008; Cohen, Frazzini, and Malloy, 2008, 2010). With several significant differences, my paper is most closely related to Ozsoylev, Walden, Yavuz, and Bildik (2014) (OWYB). OWYB estimates information networks based on overlapping trades in the Istanbul Stock Exchange during 2005. The OWYB sample is comprised of 99.9% individual investors, and OWYB use Eigenvector Centrality to show that more central investors earn higher returns. My paper complements the work of OWYB by showing that central institutional investors in the US, with a longer panel, have superior trading performance. I use Information Diffusion instead of Eigenvector Centrality because Eigenvector Centrality is a measure of importance and produces inconsistent estimates as a proxy for information. I introduce Weighted Trading Correlation Networks because institutional investors trade more frequently than individuals, and many of their overlapping trades are non-informative resulting in unreliable centrality estimates.

The remainder of this paper is organized as follows. Section 1 describes the framework including Weighted Trading Correlation Networks and Information Diffusion Centrality. In Section 2 I discuss the sample selection process. Section 3 demonstrates that central funds have superior trading performance. Section 4 examines round-trip trading around merger announcements and board director and key manager deaths. Section 6 concludes.

1 Framework

Throughout the paper I run regressions of the form:

$$r_{it} = \gamma_t + \alpha_i + \beta_0 c_{it} + \beta_1 X_{it} + \epsilon_{it}, \tag{1}$$

where r_{it} is a measure of abnormal trading performance, and c_{it} is the appropriately chosen measure of centrality. γ_t and α_i are time and fund effects, and X_{it} contains controls for trading activity. I hypothesize that if central investors have superior access to information, then *ceteris paribus* they should have superior abnormal trading performance.

In order to define the appropriate measure of centrality c_{it}, I begin with a brief introduction to networks in Section 1.1, and describe the empirical estimation of Weighted Trading Correlation Networks in Section 1.2, before defining Information Diffusion Centrality in Section 1.3.

I focus on measures of abnormal trading performance using actual execution prices of trades. If institutional trading is motivated by short-lived private information, then timing is critical. I discuss abnormal trading performance in more detail in Section 1.4.

1.1 Networks Preamble

Most network analysis focuses on collections of nodes joined by edges in a binary fashion. These networks are typically represented by (0,1) binary "adjacency" matrices:

$$A_{ij} = \begin{cases} 1 & \text{if } i \text{ is connected to } j \\ 0 & \text{otherwise} \end{cases}. \tag{2}$$

However, many settings, especially social and economic ones, may include inherently "stronger" ties and "weaker" ties (Granovetter, 1973). Both strong and weak ties are important, and focusing on only the binary adjacency matrices may result in throwing out important data. Therefore, we can also represent networks in terms of "weighted" adjacency matrices:

$$A_{ij} = w_{ij}, \tag{3}$$

where w_{ij} is the weight or strength of the connection between i and j.

For example, we can think of a simple triangle network with three nodes A, B, and C represented in geometric and adjacency matrix form below in Figure 1. Connections are bilateral, so that the adjacency matrix is symmetric. The connection between A and B is strongest with a weight of 1, followed by the connection between B and C with a weight of 0.5. The connection between A and C has the lowest weight of 0.25. In Figure 1 Panel A (below), edge thickness corresponds to the strength of the connections. The same data is encoded in the off-diagonal terms of the adjacency matrix in Panel B.

Insert Figure 1

1.2 Estimating Weighted Trading Correlation Networks

In practice we often do not know the strength of the connections between investors and hence the adjacency matrix must be estimated from trading data. The basic premise of information networks is that information sharing leads to correlated trading (Colla and Mele, 2010). A natural way to estimate the strength of connections in information networks is to estimate the pairwise trading correlations between investors.

An investor's buying activity of n stocks over T periods can be represented by a (0,1)

nT-vector stacked by stock:

$$\boldsymbol{b}_i = \begin{bmatrix} \boldsymbol{b}_{i1} \\ \boldsymbol{b}_{i2} \\ \boldsymbol{b}_{i3} \\ \vdots \\ \boldsymbol{b}_{in} \end{bmatrix} \tag{4}$$

where $\boldsymbol{b}_{in}$ is a T-vector which takes on values of one if if investor i bought stock n in period t and is zero otherwise.

The buy correlation between investor i and j is the Pearson correlation:

$$\rho^b_{ij} = \frac{cov(\boldsymbol{b}_i, \boldsymbol{b}_j)}{\sigma_{\boldsymbol{b}_i}\sigma_{\boldsymbol{b}_j}}, \tag{5}$$

and the sell correlation, ρ^s_{ij}, is defined analogously.

I estimate the weight of the connections between investors by taking the average of buy and sell trading correlations:

$$\hat{w}_{ij} = \frac{\rho^b_{ij} + \rho^s_{ij}}{2}. \tag{6}$$

I call the resulting adjacency matrix with elements $\hat{A}_{ij} \in [0, 1]$ the Weighted Trading Correlation Network (WTCN).

The primary motivation behind Weighted Trading Correlation Networks, besides preserving the continuous nature of the strength of connections, is that correlation weighting mitigates the influence of active liquidity traders on centrality estimates. For example, a related trading network estimation procedure, Empirical Investor Networks (Ozsoylev, Walden, Yavuz, and Bildik, 2014), assigns links in a binary fashion based on the presence of overlapping trades. If two investors trade the same stock on the same side in the same period then they are "connected." However, in the presence of active liquidity traders this methodology tends to identify many non-information links and can assign undue centrality to un-informed investors.

As a hypothetical example suppose the "true" information network is the "star" network in Figure 2 (below) of four investors who all share information with one central investor. The five "informed" investors are represented by dark nodes (●). Now suppose that these five investors only trade based on the shared information. As a result, they will have perfectly correlated trades as indicated by the dark lines connecting them. There is also one very active liquidity trader (○) who does not have or share any information but trades frequently. Because the liquidity trader has overlapping trades with all five informed investors we observe that the liquidity trader is part of the information network. However, because the liquidity trader trades in periods when the informed traders are not trading, the overall trading correlation with the informed investors is low, and the liquidity trader's "centrality" is also low. If we ignore the trading correlations and connect investors based only on the existence of overlapping trades the liquidity trader seems to be most important, with five connections when the bona fide central informed investor has only four connections. Trading correlation weighting significantly reduces the influence of non-information motivated trading and produces cleaner network estimates.

Insert Figure 2

1.3 Information Diffusion Centrality

Many measures of centrality exist in the finance literature. For instance Bonacich, Katz, and Eigenvector Centrality are popular in the asset pricing literature.[6] Betweenness, Closeness,

[6]Cohen-Cole, Kirilenko, and Patacchini (2014); Ozsoylev, Walden, Yavuz, and Bildik (2014); Ahern (2013).

and Degree Centrality are popular in the venture capital, corporate finance, and accounting literatures.[7] The question of what measure is "correct" in a particular setting is often difficult to answer because most measures of centrality are highly correlated with one another (see for example Valente, Coronges, Lakon, and Costenbader, 2008). Therefore, the best way to distinguish between measures of centrality is to examine the differences in their economic content.

The simplest measure of centrality is degree centrality which measures a nodes' importance by the sum of its direct connections. The degree centrality for node i is defined as the row sum of the adjacency matrix A:

$$d_i = \sum_j A_{ij}, \tag{7}$$

or in vector form $\boldsymbol{d} = \boldsymbol{A}\mathbf{1}$. We can also define $\boldsymbol{D}$, the diagonal matrix of each agents' degrees.

Eigenvector Centrality, which has been used as a proxy for information diffusion (Ozsoylev, Walden, Yavuz, and Bildik, 2014), is a recursively defined measure of centrality in which the measure of a node's importance depends on its connection to other important nodes. Eigenvector Centrality is defined recursively in vector form:

$$\boldsymbol{\varepsilon} = \frac{1}{\lambda}\boldsymbol{A}\boldsymbol{\varepsilon}. \tag{8}$$

Bonacich and Katz Centrality are generalizations of Eigenvector Centrality, which share the recursive definition of "importance." However, importance is not equivalent to being well-positioned to receive valuable information (see Appendix 7.2).

The ideal measure of access to information should capture not only how much information an investor receives, but how "delayed" the information is as it diffuses through the network.

[7] Hochberg, Ljungqvist, and Lu (2007, 2010); Hochberg, Lindsey, and Westerfield (2015); El-Khatib, Fogel, and Jandik (2015); Engelberg, Gao, and Parsons (2013); Larcker, So, and Wang (2013).

At the micro-level, the contribution to an investor's centrality from a particular connection should reflect the likelihood that information will diffuse through that channel. At the macro-level, centrality should also incorporate indirect connections, as Eigenvector does, since information is often "heard through the grapevine."

In order to capture this intuition I introduce Information Diffusion Centrality defined as:

$$\boldsymbol{c}(t, \boldsymbol{p}) = \boldsymbol{\mathcal{S}}_t \boldsymbol{p}. \tag{9}$$

The heart of Information Diffusion Centrality is the symmetric information sharing matrix $\boldsymbol{\mathcal{S}}_t = e^{-\boldsymbol{D}^{-1/2}(\boldsymbol{D}-\boldsymbol{A})\boldsymbol{D}^{-1/2}t}$. The information sharing matrix describes the dynamics of continuous bilateral information sharing (see Appendix 7.1).[8] Information Diffusion Centrality is parameterized by t and $\boldsymbol{p}$. $t > 0$ describes the time scale of information diffusion, such that when $t \to \infty$ information diffuses completely. $\boldsymbol{p}$ represents the probability that an investor receives new information. In empirical applications, I assume that $t = 1$ and p is uniform.

Central investors, according to Information Diffusion Centrality, have superior access to information because they are on the paths of many random walks. To see the random walk interpretation, we can rewrite Information Diffusion Centrality as an exponential sum of symmetric random walks $\boldsymbol{\mathcal{W}} = \boldsymbol{D}^{-1/2}\boldsymbol{A}\boldsymbol{D}^{-1/2}$:

$$\boldsymbol{c}(t, \boldsymbol{p}) = e^{-t} \sum_{k=0}^{\infty} \frac{t^k}{k!} \boldsymbol{\mathcal{W}}^k \boldsymbol{p}, \tag{10}$$

which follows from the fact that $\boldsymbol{D}^{-1/2}(\boldsymbol{D} - \boldsymbol{A})\boldsymbol{D}^{-1/2} = \boldsymbol{I} - \boldsymbol{\mathcal{W}}$. The exponential sum captures the contribution of higher-order random walks $\boldsymbol{\mathcal{W}}^k$ to information much like Eigen-

[8] It is worth mentioning that Information Diffusion Centrality is distinct from the Diffusion Centrality defined from Banerjee, Chandrasekhar, Duflo, and Jackson (2013) which is essentially a finite sum version of Katz Centrality. In Appendix 7.2 I show that Katz Centrality can be derived as a special case of a proxy for Information Diffusion Centrality, hence Diffusion Centrality can be derived in a similar fashion.

vector Centrality captures the contribution of higher-order connections to importance.[9]

More importantly, the random walk matrix reweights edges to reflect the fact that better connected nodes may actually receive information at a "delay." To see this we can look at the typical element of the random walk matrix:

$$\mathcal{W}_{ij} = \begin{cases} \frac{w_{ij}}{\sqrt{d_i}\sqrt{d_j}} & \text{if } i \text{ is connected to } j \\ 0 & \text{otherwise} \end{cases}. \tag{11}$$

Recall that in the adjacency matrix A_{ij}, edges are weighted according to w_{ij} the strength of the connection between node i and j. In the random walk matrix edges are reweighted such that $\mathcal{W}_{ij}$ represents the probability of a random walker jumping from node with degree d_i to a node with degree d_j through a connection with strength w_{ij}. This implies that low degree nodes can actually provide more "centrality" than high degree nodes. Eigenvector Centrality, by its recursive construction, predicts the opposite—connections to higher degree nodes defines importance.

To see the difference between Eigenvector Centrality and Information Diffusion Centrality more clearly we can look at a simple, but non-trivial, un-weighted network with 16 nodes represented below in Figure 3. The network is hierarchically constructed following Dorogovtsev, Goltsev, and Mendes (2002) such that there are three "core" nodes connected to 12 "peripheral" nodes with gray lines. I add an additional node C and connect it to eight peripheral nodes with black lines so that C has exactly as many connections as the three core nodes. Node size is proportional to the number of connections and node color correlates with Eigenvector Centrality on the left and Information Diffusion Centrality on the right. According to Eigenvector Centrality, C is not central because it is not connected to any of

[9]Eigenvector is computed using power iterations such that the $k+1$ iteration is $\boldsymbol{\varepsilon}_{k+1} = \frac{\boldsymbol{A}^{k+1}\boldsymbol{\varepsilon}_0}{\|\boldsymbol{A}^{k+1}\boldsymbol{\varepsilon}_0\|}$ which depends on the $k+1$ order adjacency matrix $\boldsymbol{A}^{k+1}$. Bonacich which is a generalization of Eigenvector can also be written as the geometric sum of $\boldsymbol{A}^k$ order connections: $\boldsymbol{\varepsilon}(\alpha,\beta) = \beta\sum_{k=0}^{\infty}\alpha^k\boldsymbol{A}^k\mathbf{1}$.

the three core nodes. In contrast, according to Information Diffusion Centrality, C is the most central because if information arrives at any of the eight peripheral nodes, it has a high likelihood of diffusing to C.

Insert Figure 3

1.4 Trading Performance

In the main analysis I compute quarterly interim trading performance by tracking the abnormal performance of all stocks bought and sold by a fund from the execution date, using the execution price, until the end of the quarter. Following Puckett and Yan (2011) I take the weighted-average of buy and sell trade performance and call the difference the interim trading performance.

For example, if a fund buys 90 shares of `AAPL` on January 20th at \$100 and `AAPL` closes at \$105 at the end of March, then the interim trading performance is 5% (assuming no dividends). If the fund also buys 100 shares of `MSFT` on February 17th at \$80, and `MSFT` closes at \$85, the principal-weighted buy performance for the fund is (9,000 $\times$ 5% + 8,000 $\times$ 6.25%) / (17,000) = 5.59%. If the fund makes one sell trade of 70 shares of `IBM` on March 4th at \$150 and `IBM` closes at \$153, then the principal-weighted interim trading performance for the fund is 5.59% - 2% = 3.59%. The equal-weighted interim trading performance is (5% + 6.25%)/2 = 5.625%. The sequence of trades is illustrated below in Figure 4.

Insert Figure 4

Interim trading is based on all trades, including many trades motivated by liquidity reasons. As a result the returns to being central using interim trading performance is likely to be biased downwards. Furthermore, as Fama (1970) points out, any measure of abnormal performance depends on the risk adjustment model. To get a cleaner estimate of the information advantage associated with being central, I focus my attention on a setting in which

I expect access to information to directly impact trading performance. Specifically, I examine the average abnormal round-trip trading performance in target stocks around merger announcements. Round-trip trading performance is well-suited for measuring informed trading around events because investors trading on short-lived information may want to reverse their positions in order to lock in gains or exploit overreactions (Hirshleifer, Subrahmanyam, and Titman, 1994; Brunnermeier, 2005). Furthermore, trading performance around mergers should be mostly "abnormal," regardless of the risk-adjustment, and directly related to information about the events.

To measure round-trip trading around mergers I restrict my sample to trades made 60 days before and 30 days after the announcement date. Since target CARs can be found up to 40 days before M&A announcements I select a 60-day window to allow for the possibility that funds have obtained information in the prior month. Therefore, I define a round-trip trade as a buy (sell) trade in the [-60,-1] window followed by a subsequent sell (buy) in the [0,30] window. A fund's round-trip trading performance for a given event is computed as the principal- or equal-weighted average of all signed holding period returns using the actual execution prices and volume traded.

As an example of round-trip trading, consider a hypothetical fund trading around the Disney-Pixar deal. If a fund bought 100 shares of Pixar stock on December 15th, 2005 at $56 per share prior to the announcement of the takeover by Disney, and sold on January 24th, 2006 at $58 once the deal was made public then the trading performance of that particular round-trip "buy" would have been 3.5%. If the fund also made a round-trip "sell" trade beginning January 11th at $58, ending February 1st at $57, the equal-weighted round-trip trading performance would have been (3.5%+1.7%)/2 = 2.6%. Figure 5 summarizes the sequence of trades below.

Insert Figure 5

For both interim and round-trip trading performance I adjust prices and share volume for stock splits. For interim trading performance I also cumulate dividends over the interim holding period and include them in the return calculations. I also compute "excess" or abnormal trading performance using Daniel, Grinblatt, Titman, and Wermers 1997 (DGTW) size, book-to-market, and momentum characteristic-matched value weighted portfolio returns computed over the corresponding (interim or round-trip) holding period and adjusted for delistings.

2 Data

To estimate institutional investor centrality and trading performance I use high-frequency institutional trading data from a proprietary database provided by ANcerno Ltd. (a.k.a. Abel/Noser Solutions Ltd.). The ANcerno data allows me to observe the exact date, price, direction (buy or sell), and shares traded for all funds in the database from 1999Q1 to 2011Q3. My sample stops in 2011Q3 because ANcerno removed fund identifiers in 2012, preventing me from tracking subsequent trading performance or centrality. Furthermore, 2011Q4 and 2012Q1 have only about a third of the number of funds compared to 2011Q3 which results in unstable network estimates.

I restrict my sample to all trades made in the US, in US currency, of common stock listed on the NYSE, AMEX, and NASDAQ exchanges, for which I am able to identify the ANcerno fund, and for which I am able to match the ANCerno provided point-in-time CUSIP to a CRSP `PERMNO`. The 10,355 funds in my sample trade a total of 8,555 unique common stocks. Overall, the funds make 141.85 million trades of 1.11 trillion shares valued at USD 34.26 trillion dollars. Puckett and Yan (2011) estimate that the ANcerno institutions account for approximately 8% of the total dollar value of CRSP trading volume between 1999 and 2005. Table 9 in Appendix 7.3 provides yearly summary statistics on the trading activity in my sample.

Because ANcerno timestamps are incomplete (see Anand, Irvine, Puckett, and Venkataraman, 2013) I estimate Weighted Trading Correlation Networks based on daily-level trading activity. To mitigate concerns that overlapping trades are coincidental, I restrict my sample of overlapping trades to "time-sensitive" overlapping trades. To identify time-sensitive trades, I exploit the fact that ANcerno tracks the number of days over which a stock was traded as part of an order ticket. I define time-sensitive trades as trades of stocks executed within a single day. The premise is that if a fund manager receives time-sensitive information he or she would have a strong incentive to trade immediately, versus splitting up the order execution over multiple days, consequently sacrificing price impact for immediacy. On average, time-sensitive trades make up 62% of order executions. The resulting quarterly network snapshots contain on average 1,192 funds, with a minimum of 780 in 2011Q3 and a maximum of 1,639 in 2002Q1.

In the event-study portion of the paper I use merger and acquisition (M&A) announcements and news announcements of the sudden deaths of board directors and key managers (CEO, CFO, VPs, etc.). The data on mergers comes from the Thomson Reuters SDC Platinum database (SDC). I use SDC's detailed information on M&As to identify announcement dates, the identities of the target firms, and the merger advisors.

The data on sudden deaths comes from the S&P Capital IQ (CapIQ) Key Developments database. The CapIQ data includes headlines and full-text articles of over 100 types of significant corporate events including board and executive changes (codes 16, 101, and 102 respectively). To identify sudden deaths I search through headlines and full-text articles for keywords such as `death of`, `demise of`, and `passing of`. For a more detailed description of the CapIQ data and the sudden death sample selection process see Appendix 7.4.

2.1 Fund Summary Statistics

Table 1 Panel A summarizes the main variables of interest. The performance measures are the Principal-Weighted Excess Interim Trading Performance (PW Ex. Int. Perf.) and the Equal-Weighted Performance (EW Ex. Int. Perf.). The measures of centrality are Information Diffusion Centrality (Centrality), defined in Eq. 9, and Eigenvector Centrality (Eigenvector) defined in Eq. 8. I control for trading activity using Degree Centrality (Degree), which is defined in Eq. 7, the total volume traded (Volume), and the total number of trades made (# Trades) for a given fund-quarter observation.

Insert Table 1

The average fund in my sample has a principal-weighted (equal-weighted) abnormal interim trading performance of 0.52% (0.65%). For comparison, Puckett and Yan (2011) report that the average fund between 1999 and 2005 has a principal-weighted abnormal interim trading performance of 0.57%. I verify that the average fund in my sample, over the same period, has a principal-weighted abnormal interim trading performance of 0.56%.

Because I am only concerned with a fund's relative centrality within a quarter, I normalize Information Diffusion and Eigenvector Centrality, without loss of generality, such that a fund's centrality can be interpreted as the percentage of the "total centrality" in the network. Normalizing also facilitates comparing the two centrality measures simpler. The median fund has an Information Diffusion Centrality of 2.9% and an Eigenvector Centrality of 0.74%.

The median fund makes 174 trades (Trades), exchanges 680,512 million shares (Volume), and has a Degree of four. The Degree is much smaller than the number of trades because trading correlations tend to be low. The average edge weight across quarterly WTCN network snapshots is about 2.6%. If every trade made by the median fund overlaps with one fund then the Degree is roughly $174 \times 0.026 \approx 4.5$.

Most of the variables are heavy-tailed with high kurtosis, and the centrality and trading measures are positively skewed. This is to be expected because most real-world networks are characterized by power-law degree distributions (Barabási and Albert, 1999), and as a result most centrality measures have heavy right tails. Heavy tails tends to bias certain statistics which is why, one, I focus on the medians instead of means, two, I use Spearman correlations in Panel B instead of Pearson correlations, and three, I take logs of the centrality and trading measures before using them in subsequent regressions. However, it should be noted that Information Diffusion Centrality is relatively well-behaved with a small right-skew and low kurtosis. This characteristic is due to the trading correlation edge-weighting and the random walk re-weighting, both of which both help mitigate the effects of active liquidity traders on the measurement of Centrality.

To depict the effect of edge-weighting more clearly, Figure 6 plots the distribution of Information Diffusion Centrality over the sample period at the fund-quarter level using both unweighted (–) and trading correlation weighted edges (–). The solid lines are the median values of Information Diffusion Centrality unweighted and weighted, which are highly correlated over time. The shaded regions on the other hand show that unweighted centrality estimates are heavily right-skewed and have high variance much like Degree and Eigenvector Centrality due to the influence of active traders who have many overlapping trades.[10] In contrast, the weighted version of Information Diffusion Centrality is far less right-skewed and also has a significantly smaller variance.

Insert Figure 6

Table 1 Panel B reports Spearman correlations for the same variables and includes auto-correlations on the diagonal. The two performance measures are highly correlated with one another, and are weakly negatively correlated with the measures trading activity and Eigenvector and Degree Centrality. There is a small positive correlation between performance and

[10]For example, the Pearson correlation between unweighted Information Diffusion, Eigenvector, and Degree Centrality is around 90% which is mostly driven by the outliers in the right tail.

Information Diffusion Centrality. The centrality measures are positively correlated with one another and the trading measures, and are also highly autocorrelated.

Figure 7 plots an example of a Weighted Trading Correlation Network in 2002Q2. Nodes are arranged using a force-directed layout with edge weights corresponding to the estimated trading correlations. Adjacent nodes are "attracted" to one another, while all nodes have "repulsive" forces. Lighter edges indicate lower trading correlations and darker edges indicate higher trading correlations. Node size corresponds to the number of connections (Degree) and node color corresponds to Information Diffusion Centrality. The most central node (○) using the force-directed layout appears on the periphery because it has many distinct connections which provide it superior access to "localized information."

Insert Figure 7

2.2 Event Summary Statistics

Table 2 reports the distribution of M&As, director and manager deaths, and new product announcements in my sample, as well as the average event returns. Panel A shows the number of events per year, and the number of funds making round-trip trades around these events. On average there are 633 mergers, 28 deaths, and 7,319 new product announcements per year for which at least two funds in my sample make round-trip trades. The CapIQ data starts in 2001, which is why the sample of deaths and new products is sparse in the early sample, and the trades data end in 2011Q3 which is why 2011 has fewer events.

Insert Table 2

Panel B includes average cumulative abnormal returns (CARs) from -30 to +30 based on a one-factor CAPM model as well as average principal-weighted (PW) and equal-weighted (EW) round-trip trading performance. The mean CAR for target firms around merger announcements is 5.47%, the mean CAR around director and manager deaths is -3.12%, and the mean CAR around new product announcements is -0.90%. All three are statistically different from zero, although the CAR around new product announcements is economically small.[11] The mean principal-weighted (equal-weighted) average trading performance around merger announcements is 1.23% (1.19%) and is statistically different from zero, consistent with the observation that there are information leakages prior to merger announcements (see Betton, Eckbo, and Thorburn, 2008). The mean principal-weighted (equal-weighted) average trading performance around deaths is -0.27% (-0.25%) and is statistically indistinguishable from zero, consistent with the notion that these sudden deaths are unpredictable (e.g. Nguyen and Nielsen, 2010). The mean principal-weighted (equal-weighted) average trading performance around new product announcements is -0.03% (-0.05%) and is statistically significant but economically small. The average fund has a near-zero average trading performance around new product announcements, which is to be expected if new product information requires expertise to process.

3 Do Central Investors Have Superior Access to Information?

The premise of Information Diffusion Centrality is that central investors have superior access to information by virtue of being on the paths of many "random walks" as information diffuses through the network. If central funds have superior access to information, then they might also have superior trading performance.

As formal test of the main hypothesis, I estimate multiple regressions of the form:

$$r_{it} = \gamma_t + \alpha_i + \beta_0 c_{it} + \beta_1 X_{it} + \epsilon_{it}. \tag{12}$$

[11]The one-factor model is estimated over a year of data (252 days) with a 30-day gap between the estimation window and the event window. Other factor models produce similar results since most of the return is due to the announcement effect.

r_{it} is the excess interim trading performance computed using DGTW-adjusted returns at the fund-quarter level (denoted by the subscripts i and t respectively). γ_{t} is a quarter fixed effect intended to control for common trends in fund performance, i.e. bull- and bear markets. α_{i} is the unobservable component of trading performance which may be related to "manager skill." Because centrality and many of the other covariates are persistent over time, I make the identifying assumption that α_i is orthogonal to the other covariates and estimate it as a random effect. c_{it} is the fund's log rescaled centrality. Finally, X_{it}, contains the set of controls for the amount of trading a fund does in a quarter: these controls are quarterly trading volume, the number of trades executed, and the fund's degree centrality all in logs.

In every specification I control for Degree for two reasons. First, I want to capture the effect of Centrality above and beyond Degree to isolate the effect of information diffusion beyond first-order connections. Second, because connections are estimated based on observed trades rather than the "true" information linkages, the centrality measures are likely to suffer from "errors-in-variables" problems. Because multiple regression is equivalent to a series of orthogonal projections, "controlling for" Degree allows me to estimate the effect of being central orthogonal to Degree and the edge estimation error.[12]

Table 3 presents the main results of the paper. Columns 1–3 (5–7) of Table 3 show OLS estimates from regressions of principal-weighted (equal-weighted) average excess interim trading performance on measures of centrality without fund effects. In Column 1 (5) a one standard deviation increase in Information Diffusion Centrality corresponds to a 0.300% (0.336%) increase in principal-weighted (equal-weighted) average excess trading performance controlling for Degree. The effect is statistically different from zero at the one percent level, based on standard errors clustered by fund and quarter. For comparison, the average fund in my sample has a 0.52% (0.65%) principal-weighted (equal-weighted) average excess

[12] Suppose that each edge is estimated with errors that are i.i.d. with respect to other edges, i.e. $\tilde{w}_{ij} = w_{ij} + u_{ij}$. Then the estimated degree of each node i is: $\tilde{d}_{ij} = \sum_j \tilde{w}_{ij} + u_{ij} = d_{ij} + \sum_j u_{ij}$. "Controlling" for $\tilde{d}_{ij}$ is equivalent to orthogonalizing with respect to the edge errors represented by the second term.

interim trading performance. In other words, the trading performance of "central funds" is approximately 60% higher than that of the average fund.

In contrast, in Column 2 (6) a one standard deviation increase in Eigenvector Centrality corresponds to a smaller 0.102% (0.0615%) increase in trading performance. The estimate is significant at the 10% level for principal-weighted performance but insignificant for equal-weighted performance. This result is consistent with the claim that Eigenvector Centrality is a measure of importance and only a weak proxy for information diffusion because it is positively correlated with Information Diffusion Centrality (see Table 1).

In the "kitchen sink" specification using both measures of centrality and controls for trading activity the results are qualitatively similar. In Column 3 (7) a one standard deviation increase in Information Diffusion Centrality corresponds to a 0.327% (0.328%) increase in principal-weighted (equal-weighted) performance. Adding a fund random effect in Column 4 (8), the coefficient of principal-weighted (equal-weighted) performance on Information Diffusion Centrality decreases by 0.13% (0.15%). This would be consistent with persistent "skill" in trading performance which is absorbed by the random effect. Moreover, the estimate on Eigenvector Centrality shrinks to 0.0740% (0.0221%) using principal-weighted (equal-weighted) trading performance. The estimates on Information Diffusion Centrality are statistically significant at the 1% level using a block-bootstrap procedure using 100 bootstrap samples. The estimates on Eigenvector Centrality are statistically indistinguishable from zero for equal-weighted average performance.

Insert Table 3

4 Evidence from Merger Announcements

In this section I examine round-trip trading performance around merger announcements to determine whether central funds have superior access to information prior to the public announcements. To account for common sources of information, I examine funds' connections to merger advisors which may be an important determinant of trading performance that is contemporaneously correlated with centrality.

4.1 Round-Trip Trading Around Mergers

The finance literature offers extensive evidence for information leakages prior to merger announcements, and the existence of a large "target premium" to be earned from buying ahead of merger announcements and selling on the announcement date. Betton, Eckbo, and Thorburn (2008) survey the M&A literature and report that the average cumulative abnormal returns of target firms around announcement dates are as large as 28%.

If Information Diffusion Centrality measures access to information, then central funds may be receiving information ahead of merger announcements and earning significant returns from trading on this information. To test this hypothesis I estimate the following regression:

$$r_{it} = \gamma_t + \alpha_i + \beta_0 c_{it} + \beta_1 X_{it} + \epsilon_{it}, \tag{13}$$

where r_{it} is the principal- or equal-weighted average excess round-trip trading performance, computed from DGTW-adjusted returns at the fund-event level (denoted by the subscripts i and t respectively). γ_{t} is an event fixed effect, as there may be some inherently "good deals" or "bad deals." α_{i} is the unobservable "fixed" component of trading performance, which I estimate as a fund random effect. c_{it} is the fund's centrality estimated in the quarter in which the announcement occurred. Finally, X_{it}, contains the set of controls, specifically, quarterly trading volume, the number of trades executed, and the fund's degree centrality all in logs.

Table 4 Columns 1–4 (5–8) present estimates from regressions of principal-weighted (equal-weighted) average excess round-trip trading performance on Information Diffusion

Centrality. A one standard deviation increase in Information Diffusion Centrality in Column 1 (5) corresponds to a 1.484% (1.542%) increase in principal-weighted (equal-weighted) average excess round-trip trading performance. As before, the estimates are significant at the 1% level with standard errors clustered by fund and quarter. To put the magnitude of the estimate in perspective, the average fund in Table 2 Panel B has a 1.23% (1.19%) principal-weighted (equal-weighted) average excess round-trip trading performance around merger announcements. Therefore the trading performance of "central funds" is more than double that of the average fund. These results suggest that Information Diffusion is a good proxy for information diffusion, and that superior access to information plays a significant role in round-trip trading performance around mergers.

A one standard deviation increase in Eigenvector Centrality in Column 2 (6) corresponds to a 0.524% (0.529%) increase in round-trip trading performance. The magnitude of the coefficient is roughly one-third of that of Information Diffusion Centrality. Again Eigenvector Centrality appears to be a weak proxy for information diffusion.

In the "kitchen sink" specification with both measures of centrality and controls for trading activity the results are qualitatively similar. The coefficient on Information Diffusion Centrality in Column 3 (7) increases to 1.805% (1.929%) for principal-weighted (equal-weighted) trading performance, while the coefficient on Eigenvector Centrality decreases to 0.485% (0.475%).

Adding a fund random effect in Column 4 (8) does not considerably shrink the estimate of the coefficient on Centrality. A one standard deviation increase in Information Diffusion Centrality corresponds to a 1.131% (1.267%) increase in principal-weighted (equal-weighted) trading performance. Roughly 63% (66%) of the "returns to being central" based on principal-weighted (equal-weighted) round-trip trading performance is unrelated to the persistent component of a fund's trading performance. Compared to quarterly interim trading performance, round-trip trading performance around idiosyncratic events should be more directly related to access to information, therefore it is not surprising that the fund "fixed" effect does not absorb much of the returns to being central. In contrast, only about 30% of the returns from Eigenvector Centrality remain after adding the fund effect.

Insert Table 4

4.2 M&A Advisor "Tipping"

Central funds appear to have superior access to information about mergers through their network connections. However, central funds might receive information about mergers through their connections to merger advisors rather than other investors in their network. Large broker-dealers frequently act as merger advisors to acquiror and target firms and could potentially "tip off" their institutional clients to these deals. Jegadeesh and Tang (2011) find that institutions connected to merger advisors are net buyers of target stocks; they also make profitable trades in aggregate. Therefore, the previous section's network analysis may not capture the potential source of information leakages prior to merger announcements due to brokerage house connections.

To examine this source of information I begin by name-matching ANcerno brokers to SDC merger advisors on both the target and acquiror side. In most M&A deals there are multiple advisory firms on each side, so there are potentially many sources of informative tips. In order to try and isolate the relevant connections, I compute the total dollar value traded by each fund with each of their brokers for each quarter. I say that a fund is "connected" to an advisor for an M&A deal if in the previous quarter the advisor was one of the top-decile brokers for that fund. Jegadeesh and Tang (2011) use a similar approach, but at the yearly level. I then repeat the round-trip trading analysis with an indicator for a fund's connection to an M&A advisor (Tip), and I include an interaction term with centrality (Cent. $\times$ Tip). Including the interaction term allows me to interpret the effects of being central independently of M&A advisor connections.

Table 5 Columns 1–3 show weak evidence that funds connected to M&A advisors earn higher returns from round-trip trades of merger targets' stock. Columns 4–8 seem to indicate that funds connected to merger advisors actually underperform on average. However, in all specifications the estimate is statistically indistinguishable from zero.[13]

In contrast, the coefficient on Information Diffusion Centrality is large, significant, and unaffected by the inclusion of M&A advisor connections. A one standard deviation increase in Information Diffusion Centrality corresponds to a 1.822% (1.947%) percent increase in principal-weighted (equal-weighted) average round-trip trading performance in Column 1 (5). These results suggest that "grapevine" communication may play a more important role in information leakages than "tipping."

If we interpret Eigenvector Centrality as "importance" rather than as a proxy for information diffusion then Column 2 (6) provides evidence that important funds within the network benefit from M&A advisor connections. The interaction between Eigenvector Centrality and M&A advisor connections (Eig. $\times$ Tip) indicates that funds that are both "important" and "connected" have 1.104% (0.805%) higher principal-weighted (equal-weighted) average round-trip trading performance relative to the average fund. Important and connected funds still earn less than "central" funds according to Information Diffusion Centrality as seen in Columns 3 and 7. This finding is consistent with the notion that importance does not imply superior access to information (see Figure 3 for example).

Adding a fund effect in Columns 4 and 8 reduces the coefficient estimates on the centrality measures, M&A advisor connections, and the interaction terms. Information Diffusion Centrality remains large and significant at the 1% level, Eigenvector Centrality becomes significant at the 10% level, and its interaction with M&A advisor connections becomes

[13] Jegadeesh and Tang (2011) find significant results using aggregated institutional trades, whereas I use disaggregated trades which are noisier.

statistically indistinguishable from zero.

Insert Table 5

5 Ruling Out Alternative Hypotheses

In this section I conduct three additional empirical tests to rule out alternative hypotheses. First, an important concern is whether the returns to being central are related to expertise in information processing. To address this issue, I look at a setting where expertise should directly contribute to trading performance in order to disentangle the effects of centrality from expertise. I examine round-trip trading around new product announcements where I expect product market expertise to be essential to information processing and trading performance.

Second, as a falsification test, I examine sudden deaths of board directors and key executives to see if centrality is spuriously predicting trading performance. If there is no information about sudden deaths, there should be no information advantage to being central.

Finally, I repeat my main regression results within a propensity score matched sample to try and mitigate biases due to endogeneity. Wooldridge (2010) refers to propensity score matched sample regression estimation as "doubly robust" method in the sense that it only requires either a well specified propensity score model or a well specified regression model. I ensure the former by requiring that Central funds are matched to Peripheral funds with propensity scores that are within 0.1 basis points in absolute value, and I show that the matching covariates are balanced.

5.1 Expertise versus Information Diffusion

Trading performance may also be determined by superior information processing which may be related expertise in a particular industry or product market. Tookes (2008) and Peress (2010) suggest that product market imperfections create opportunities for profitable industry specific informed trading. I examine new product announcements as a setting in which I expect product market expertise to be distinguishable from information diffusion as predictors of trading performance.

To measure product market expertise I rely on the text-based network industry classifications (`TNIC`s) provided by Hoberg and Phillips (2010, 2015). Hoberg and Phillips (2010, 2015) classify firms into `TNIC`s industries using the product descriptions found in 10-K filings. Each product announcement in my sample is assigned to an industry based on yearly `GVKEY-TNIC` pairs. I compute a fund's product market expertise as the trade-weighted average cosine similarity between the announcing firm's product description and the product descriptions of all of the other firms in which the fund traded stocks. The resulting measure is bounded between zero and one, where larger values indicate that a fund's trading is concentrated in the announcing firm's industry. Intuitively, if a fund exclusively trades in technology stocks such as `AAPL`, `MSFT`, and `IBM` then the fund would be classified as an expert in the `TNIC` technology industry.[14]

Table 6 presents the results of regressions of average excess round-trip trading performance on my measures of Expertise, Centrality, and the interaction which captures the effect of being both an Expert and Central. Column 1 (3) shows that a one standard deviation increase in Expertise corresponds to 0.044% (0.032%) higher principal-weighted (equal-weighted) average excess round-trip trading performance. A one standard deviation increase in Centrality corresponds to 0.392% (0.367%) higher principal-weighted (equal-weighted)

[14] `TNIC` definitions change over time as product categories change. Because I am only concerned with whether a fund trades stocks similar to the announcing firm within a given year, the time series variation in `TNIC`s does not affect my definition of expertise. See Hoberg and Phillips (2010, 2015) for more details.

average excess round-trip trading performance.

Adding a fund random effect in Column 2 (4) shrinks the estimates on Expertise and Centrality. A one standard deviation increase in Centrality corresponds to 0.130% (0.190%) higher principal-weighted (equal-weighted) average excess round-trip trading performance, but the estimate is statistically indistinguishable from zero. However in the case of new product announcements, the random effects specification is overly conservative because the within fund autocorrelation for centrality is 97%; hence the random effect precludes parameter estimation for centrality. In contrast, the autocorrelation for my measure of expertise is only 61%, so the estimate is not affected as much by the addition of the random effect.

Overall, these results suggest that Centrality and Expertise are distinct from one another. The fact that Central investors are different from Experts is not surprising because Experts' are by construction concentrated within an industry and therefore Experts are well-connected to the general population on average.

The more surprising result is that Centrality seems to be more important than Expertise as a determinant of trading performance around new product announcements. That is not to say that Expertise is not important. Considering that the average fund has a near-zero trading performance around new product announcements, an additional 4 basis points from being an Expert is a non-trivial increase in trading performance.

5.2 (No) Timing Sudden Director and Manager Deaths

An important concern is determining whether central fund managers have genuine information or whether centrality is somehow correlated with trading performance due to pseudo event timing. For example Butler, Grullon, and Weston (2005) finds that managers pseudo market time equity markets and are not "forecasting" when their stock is overvalued in order to issue equity.

As a falsification test, I look at round-trip trading around sudden deaths of directors and managers. Because sudden deaths are unpredictable, centrality, as a proxy for access to information, should play no role in the performance of trades made around these events.

Consistent with this intuition, Table 7 shows that central funds do not have statistically significant increased trading performance relative to the average fund. When there is no information to share, there is no information advantage to being central.

Insert Table 7

5.3 Matched Sample Regressions

A final concern is that my results may be driven by unobserved heterogeneity between central funds and peripheral funds. Michaely and Roberts (2012) uses a propensity score matched sample regression to account for the endogenous selection of being a private firm in order to study the dividend payout policy of private firms versus public firms. Wooldridge (2010) refers to propensity score matched sample regression estimation as "doubly robust" method in the sense that it only requires either a well specified propensity score model or a well specified regression model.

To account for differences between Central and Peripheral funds I compute the probability (propensity) of having above median Centrality in a given quarter (Central) conditional on Degree, Eigenvector, # of Trades, and Volume.[15] To ensure that Central funds are matched to sufficiently similar Peripheral funds I require that the difference in propensity scores does not exceed 0.1 basis points in absolute value. I report the results of regressions within the full-sample and matched-sample in Table 8 Panel A and I report normalized differences in Panel B (see Imbens and Wooldridge, 2009).

Insert Table 8

[15]As before, I orthogonalize Information Diffusion Centrality with respect to Degree to account for estimation error and to isolate the effect of Centrality from Degree.

The results are qualitatively similar to the main results in Table 3. In Column 1 (4) of Panel A, Central funds have a 0.446% (0.491%) higher principal-weighted (equal-weighted) average excess trading performance in the full sample. Within the matched sample, shown in Column 2 (4), the effect is approximately 50 basis points larger. Adding fund random effects in Column 3 (6) reduces the coefficient to 0.250% (0.241%).

In the full sample Eigenvector Centrality has a positive coefficient, but in the matched sample higher Eigenvector Centrality actually corresponds to lower trading performance. However, the point estimate is only significant at the 10% level for the full-sample principal-weighted average excess trading performance.

The normalized differences in Panel B are smaller than 0.12 standard deviations for all covariates which indicates a good match. Peripheral and Central funds appear to have different average trading activity in Panel B but the standard deviations for trading activity are very large which is why the normalized differences are small (see 1).

6 Conclusion

Overall my findings indicate that institutional investors share valuable information with one another. My findings further indicate that word-of-mouth information sharing may be a stylized fact of investor behavior which is not limited to individual investors (Ivković and Weisbenner, 2007) or local institutions (Hong, Kubik, and Stein, 2005; Pool, Stoffman, and Yonker, 2015).

Broadly speaking, this paper provides evidence that the structure of information networks is related to information diffusion, which in turn supports the view that the "microstructure of social transactions" affects investment decisions (Hirshleifer, 2015). Hirshleifer (2015) argues that the next step in behavioral finance is to study *social* finance—in other words, to incorporate theories of sociology into our understanding of financial markets. Shiller (1995) argues that many behavioral biases may be attributed to the way in which economic agents converse. Herding, home bias, and underreaction, three of the most well-documented behavioral biases, are all arguably related to the localization of information within the "community" structure of networks. The sociology literature has long recognized that economic actions such as learning may be "embedded" in social networks (Granovetter, 1985), but network analysis is relatively new to the finance literature. A more detailed subsequent analysis of network structure, beyond centrality, may provide a fuller understanding of the distribution and dispersion of information in financial markets.

In addition, I find that there is an economically meaningful difference between Eigenvector Centrality, a network measure of importance from the sociology literature, and Information Diffusion Centrality, in terms of their ability to predict trading performance. Therefore empirical researchers in finance should be careful in using sociological measures of centrality as proxies in finance applications. However, importance and information diffusion are likely to be more intertwined than indicated by my analysis. For example, (DeMarzo, Vayanos, and Zwiebel, 2003) argues that important individuals can have undue influence on the formation of opinions through repeated interactions and naive social learning. Detailed data on actual communications between investors and more sophisticated models of information diffusion are likely to yield new insights.

Finally, Weighted Trading Correlation Network estimation may be useful as more high-frequency institutional trading data becomes available.

7 Appendix

7.1 Information Sharing Matrix

Suppose that each agent $i = 1, \ldots, N$ is endowed with a private signal s_i, and that agents continuously exchange a fraction $w_{ij} < 1$ of their information with their neighbors, and update their information based on the differences between their signals and their neighbors' signals. We can represent the evolution of agent i's signal as:

$$\begin{aligned}\frac{ds_i}{dt} &= -\sum_j A_{ij}(s_j - s_i) \\ &= -d_i s_i + \sum_j A_{ij} s_j, \end{aligned} \tag{14}$$

such that agent i "sends" d_i of his signal to his neighbors, and "receives" $\sum_j A_{ij} s_j$ from his neighbors. We can write the above in vector form,

$$\frac{d\boldsymbol{s}}{dt} = -(\boldsymbol{D} - \boldsymbol{A})\boldsymbol{s} = -\boldsymbol{L}\boldsymbol{s}, \tag{15}$$

where $\boldsymbol{D}$ is the diagonal matrix of each agents' degrees and $\boldsymbol{L}$ is the graph Laplacian. The differential equation has the solution:

$$\boldsymbol{s}_t = e^{-\boldsymbol{L}t}\boldsymbol{s}_0, \tag{16}$$

given the initial distribution of signals $\boldsymbol{s}_0$. Equation (15) is called the heat equation, as it describes the behavior of heat diffusion on a square lattice, and $\boldsymbol{S}_t = e^{-\boldsymbol{L}t}$ is the non-symmetric information "sharing" matrix.[16] To make information sharing symmetric we can replace $\boldsymbol{S}_t$ with its symmetric counterpart $\boldsymbol{\mathcal{S}}_t = e^{-\boldsymbol{D}^{-1/2}\boldsymbol{L}\boldsymbol{D}^{-1/2}t}$.

7.2 Eigenvector centralities

One measure which has not been popular in the finance literature, Personalized PageRank, is useful because it is a close proxy for Information Diffusion Centrality and can be used to

[16] $\boldsymbol{S}_t$ is also called the "heat kernel" in the spectral graph theory literature. Chung and Yau (1998) introduces the symmetric heat kernel, and Chung (2007) develops a centrality measure based on the non-symmetric version $e^{-\boldsymbol{L}t}$.

derive the "eigenvector centralities" popular in the finance literature (Bonacich, Katz, and Eigenvector). Recall that Information Diffusion Centrality is defined as the exponential sum of symmetric random walks (9). Analogously, Personalized PageRank can be written as the *geometric* sum of random walks:

$$\boldsymbol{\varepsilon}(\alpha, \beta) = \beta \sum_{k=0}^{\infty} \alpha^k \boldsymbol{W}^k \boldsymbol{p}. \tag{17}$$

Or in vector form, $\boldsymbol{\varepsilon} = (\boldsymbol{I} - \alpha \boldsymbol{W})^{-1} \beta \boldsymbol{p}$. Here, $\boldsymbol{p}$ represents the "preference" vector or probability that a random walker will start from any given node which is analogous to the arrival of information above. α acts as a damping factor in the geometric series, and β is a positive constant.[17] The standard random walk matrix $\boldsymbol{W}$ can also be replaced with its symmetric counterpart to be consistent with the bilateral exchange of information.

PageRank, named after Google co-founder Larry Page, is designed to measure the importance of websites based on the notion that a website is important if other important websites link to it. As we have shown, it can also be expressed in terms of a "random (web) surfer" clicking links on pages such that important sites tend to be on the paths of many random walks (see Brin and Page, 1998; Page, Brin, Motwani, and Winograd, 1999). However it would be a mistake to call PageRank a measure of information diffusion as PageRank contains no description of the underlying information sharing process. Hence the canonical form of PageRank (with $\boldsymbol{p} = \boldsymbol{1}$) is not the random walk form shown above, but the recursive form:

$$\boldsymbol{\varepsilon} = \alpha \boldsymbol{W} \boldsymbol{\varepsilon} + \beta \boldsymbol{1}. \tag{18}$$

Bonacich, Katz, and Eigenvector centrality are also defined in recursive form, and can

[17] β is sometimes set to $(1 - \alpha)$ with $\alpha < 1$ such that the above random walk can be interpreted as a "teleporting random walk."

in fact be derived as special cases of PageRank. To see this, we only need to examine the typical element of the PageRank vector:

$$\varepsilon_i = \alpha \sum_j A_{ij} \frac{\varepsilon_j}{d_j} + \beta. \tag{19}$$

PageRank can be viewed as a scaled or random walk version of Bonacich centrality. When ε_j is not scaled by the degree d_j inside the summation (i.e. if $d_j = 1$ above) PageRank becomes Bonacich centrality. Katz centrality is a special case of Bonacich centrality when $\beta = 1$, and the standard Eigenvector centrality is a special case of Bonacich centrality with $\alpha = 1$ and $\beta = 0$.

Although these "eigenvector centralities" may be positively correlated with Information Diffusion Centrality, only PageRank has the random walk interpretation, and even then PageRank lacks the information sharing foundations of information diffusion.

7.3 ANCerno Data

Each observation in the ANcerno dataset consists of the execution of a trade along with the corresponding order-level identifiers, and trade-day liquidity measures.

Stocks are identified by `CUSIP` at the time of execution, ANcerno clients are identified by *clientcode*, their money managers by *managercode*, and the broker-dealer responsible for each execution by *brokercode*. Following Puckett and Yan (2011) I define a fund as a client-manager pair. Using the ANcerno provided cross-reference metadata, I identify 10,355 unique funds, which are composed of 1,358 clients and 1,007 money managers. A sample of the cleaned data is shown below.

Obs	Order	Date	Fund	Broker	Stock	Price	Volume	Side	Days
1	0	2002-01-23	1	5	75577	22.529	8100	-1	1
2	784	2002-01-23	5	7486	77472	38.591	3550	-1	1
3	1138	2002-01-23	5	7486	83597	35.960	3645	-1	1
4	1660	2002-01-29	5	6826	77662	50.960	5775	-1	1
5	521	2002-02-01	5	7486	79133	31.955	2600	1	3
6	521	2002-02-05	5	7486	79133	31.560	10700	1	3
7	1547	2002-02-19	5	13	12490	100.153	2600	-1	1
8	621	2002-03-07	5	161	43772	57.025	1765	-1	1
9	621	2002-03-07	5	26438	43772	57.025	1765	-1	1
10	1996	2002-03-12	5	50	10104	14.313	5800	1	1

Order identifies an order ticket, Date indicates when the trade was made, Fund is the composite *clientcode-brokercode* identifier, Broker is the broker identifier, Stock is the `PERMNO` corresponding to the provided `CUSIP`, Price and Volume are at the trade level, Side is 1 if the trade was a buy and -1 if the trade was a sell, and Days refers to the number of days over which the stock was traded as part of the Order.

The funds in my sample trade a total of 8,555 unique common stocks listed across the NYSE, NASDAQ, and AMEX exchanges. Overall, the funds make 141.85 million trades of 1.11 trillion shares valued at 34.26 trillion dollars USD. Puckett and Yan (2011) estimate that the ANcerno institutions account for approximately 8% of the total dollar value of CRSP trading volume between 1999 and 2005. More detailed summary statistics are provided in Table 9 below.

Insert Table 9

Puckett and Yan (2011) obtains a sample of 64 ANcerno client institution names which include pension plan sponsors such as CalPERS, the Commonwealth of Virginia, the YMCA retirement fund, money managers such as Massachusetts Financial Services, Putnam Investments, Lazard Asset Management, and several broker-dealers. Although ANcerno does not report the identities of their clients, the dataset does include the names of their clients' money managers and brokers. These money managers include investment management companies such as AllianceBernstein, BlackRock, and Wellington as well as investment banks such as Goldman Sachs, JP Morgan, and Deutsche Bank.[18] Broker-dealers include independent broker-dealers such as LPL Financial, Raymond James Financial Services, Ameriprise Financial Services as well as dealer banks such as Bank of America, Citigroup, and Credit Suisse.

7.4 Capital IQ - Sudden Deaths

The CapIQ Key Developments data includes over 100 types of significant corporate events including executive changes, M&A rumors, changes to guidance, delayed filings, SEC inquiries and more.[19] The dataset includes over 8.5 million events involving over 800,000 companies worldwide. CapIQ claims that they review between 14,000-16,000 news stories from 165 countries on a daily basis in addition to press releases, regulatory filings, company websites, investor conference organizer websites, and call transcripts. CapIQ also claims that equity research analysts, investment bankers, and portfolio managers use their data to predict the impact of events on stock performance, i.e. to conduct event studies.

I use their data on director and manager changes to identify changes related to sudden deaths. I search headlines for `death of`, `passing of`, `demise of`, `dies`, `died`, `dead`, `demise`, `passed away`, and `killed` and find 888 observations. I also search the full text article and remove mentions of `suicide`, which eliminates two observations. I also try searching the full text for the same keywords, which but most of the additional events are misclassified and need to be removed manually. For example there are a few observations in which the

[18]The database also includes the Carlyle Group, KKR & Co., and Blackstone Group LP which were formerly external managers investing on behalf of CalPERS, an ANcerno client. Source: "Calpers to Cut External Money Managers by Half." WSJ. Web. 21 July 2015.

[19]Source: WRDS FAQs on S&P Capital IQ.

CEO stepped down following the death of plant workers. There are also several instances in which the announcement is about the newly elected board member or manager following a (previously recorded) death. Since the headline search is cleaner and the sample sizes are about the same after manual cleaning I use the headline-search only sample.

References

Ahern, Kenneth R, 2013, Network centrality and the cross section of stock returns, *Available at SSRN 2197370.*

———, 2014, Information networks: Evidence from illegal insider trading tips, *Available at SSRN.*

Aktas, Nihat, Eric De Bodt, Fany Declerck, and Herve Van Oppens, 2007, The pin anomaly around m&a announcements, *Journal of Financial Markets* 10, 169–191.

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

Banerjee, Abhijit, Arun G Chandrasekhar, Esther Duflo, and Matthew O Jackson, 2013, The diffusion of microfinance, *Science* 341, 1236498.

Barabási, Albert-László, and Réka Albert, 1999, Emergence of scaling in random networks, *Science* 286, 509–512.

Betton, Sandra, B. Espen Eckbo, and Karin S. Thorburn, 2008, Chapter 15 - corporate takeovers, in *Handbook of Empirical Corporate Finance*, ed. by B. Espen Eckbo, vol. 2 of *Handbooks in Finance* pp. 291–429.

Boehmer, Ekkehart, and Eric K Kelley, 2009, Institutional investors and the informational efficiency of prices, *Review of Financial Studies* 22, 3563–3594.

Brin, Sergey, and Lawrence Page, 1998, The anatomy of a large-scale hypertextual web search engine, *Computer Networks and ISDN Systems* 30.

Brown, Jeffrey R, Zoran Ivković, Paul A Smith, and Scott Weisbenner, 2008, Neighbors matter: Causal community effects and stock market participation, *The Journal of Finance* 63, 1509–1531.

Brunnermeier, Markus K, 2005, Information leakage and market efficiency, *Review of Financial Studies* 18, 417–457.

Butler, Alexander W, Gustavo Grullon, and James P Weston, 2005, Can managers forecast aggregate market returns?, *The Journal of Finance* 60, 963–986.

Chung, Fan, 2007, The heat kernel as the pagerank of a graph, *Proceedings of the National Academy of Sciences* 104, 19735–19740.

——— , and Shing-Tung Yau, 1998, Coverings, heat kernels and spanning trees, *Journal of Combinatorics* 6, 163–184.

Cohen, Lauren, Andrea Frazzini, and Christopher Malloy, 2008, The small world of investing: Board connections and mutual fund returns, *Journal of Political Economy* 116, 951–979.

——— , 2010, Sell-side school ties, *The Journal of Finance* 65, 1409–1437.

Cohen-Cole, Ethan, Andrei Kirilenko, and Eleonora Patacchini, 2014, Trading networks and liquidity provision, *Journal of Financial Economics* 113, 235–251.

Colla, Paolo, and Antonio Mele, 2010, Information linkages and correlated trading, *Review of Financial Studies* 23, 203–246.

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

DeMarzo, Peter M, Dimitri Vayanos, and Jeffrey Zwiebel, 2003, Persuasion bias, social influence, and unidimensional opinions, *The Quarterly Journal of Economics* 118, 909–968.

Dorogovtsev, Sergey N, AV Goltsev, and José Ferreira F Mendes, 2002, Pseudofractal scale-free web, *Physical Review E* 65, 066122.

Duarte, Jefferson, Edwin Hu, and Lance A Young, 2015, What does the pin model identify as private information?, *Available at SSRN 2564369*.

Duffie, Darrell, Gaston Giroux, and Gustavo Manso, 2010, Information percolation, *American Economic Journal: Microeconomics* 2, 100–111.

Duffie, Darrell, Semyon Malamud, and Gustavo Manso, 2009, Information percolation with equilibrium search dynamics, *Econometrica* 77, 1513–1574.

Duffie, Darrell, and Gustavo Manso, 2007, Information percolation in large markets, *The American Economic Review* 97, 203–209.

El-Khatib, Rwan, Kathy Fogel, and Tomas Jandik, 2015, Ceo network centrality and merger performance, *Journal of Financial Economics* 116, 349–382.

Engelberg, Joseph, Pengjie Gao, and Christopher A Parsons, 2013, The price of a ceo's rolodex, *Review of Financial Studies* 26, 79–114.

Fama, Eugene F, 1970, Efficient capital markets: A review of theory and empirical work, *The Journal of Finance* 25, 383–417.

Granovetter, Mark, 1985, Economic action and social structure: the problem of embeddedness, *American Journal of Sociology* 91, 481–510.

Granovetter, Mark S, 1973, The strength of weak ties, *American Journal of Sociology* 78, 1360–1380.

Griffin, John M, Tao Shu, and Selim 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.

Han, Bing, and Liyan Yang, 2013, Social networks, information acquisition, and asset prices, *Management Science* 59, 1444–1457.

Hendershott, Terrence, Dmitry Livdan, and Norman Schürhoff, 2015, Are institutions informed about news?, *Journal of Financial Economics* 112, 249–287.

Hirshleifer, David, 2015, Behavioral finance, *Annual Review of Financial Economics* 7.

———, Avanidhar Subrahmanyam, and Sheridan Titman, 1994, Security analysis and trading patterns when some investors receive information before others, *Journal of Finance* 49, 1665–1698.

Hoberg, Gerard, and Gordon Phillips, 2010, Product market synergies and competition in mergers and acquisitions: A text-based analysis, *Review of Financial Studies* 23, 3773–3811.

Hoberg, Gerard, and Gordon M Phillips, 2015, Text-based network industries and endogenous product differentiation, *Journal of Political Economy, Forthcoming.*

Hochberg, Yael V, Laura Anne Lindsey, and Mark M Westerfield, 2015, Resource accumulation through economic ties: Evidence from venture capital, *Journal of Financial Economics (JFE), Forthcoming.*

Hochberg, Yael V, Alexander Ljungqvist, and Yang Lu, 2007, Whom you know matters: Venture capital networks and investment performance, *The Journal of Finance* 62, 251–301.

———, 2010, Networking as a barrier to entry and the competitive supply of venture capital, *The Journal of Finance* 65, 829–859.

Hong, Dong, Harrison G Hong, and Andrei Ungureanu, 2011, An epidemiological approach to opinion and price-volume dynamics, *SSRN Working Paper Series.*

Hong, Harrison, Jeffrey D Kubik, and Jeremy C Stein, 2005, Thy neighbor's portfolio: Word-of-mouth effects in the holdings and trades of money managers, *The Journal of Finance* 60, 2801–2824.

Hong, Harrison, and Jeremy C Stein, 1999, A unified theory of underreaction, momentum trading, and overreaction in asset markets, *The Journal of Finance* 54, 2143–2184.

Imbens, Guido W., and Jeffrey M. Wooldridge, 2009, Recent developments in the econometrics of program evaluation, *Journal of Economic Literature* 47, 5–86.

Ivković, Zoran, and Scott Weisbenner, 2005, Local does as local is: Information content of the geography of individual investors' common stock investments, *The Journal of Finance* 60, 267–306.

———, 2007, Information diffusion effects in individual investors' common stock purchases: Covet thy neighbors' investment choices, *Review of Financial Studies* 20, 1327–1357.

Jegadeesh, Narasimhan, and Yue Tang, 2011, Institutional trades around takeover announcements: Skill vs. inside information, *SSRN Working Paper Series.*

Larcker, David F, Eric C So, and Charles CY Wang, 2013, Boardroom centrality and firm performance, *Journal of Accounting and Economics* 55, 225–250.

Manela, Asaf, 2014, The value of diffusing information, *Journal of Financial Economics* 111, 181–199.

Meulbroek, Lisa K, 1992, An empirical analysis of illegal insider trading, *The Journal of Finance* 47, 1661–1699.

Michaely, Roni, and Michael R Roberts, 2012, Corporate dividend policies: Lessons from private firms, *Review of Financial Studies* 25, 711–746.

Nguyen, Bang Dang, and Kasper Meisner Nielsen, 2010, The value of independent directors: Evidence from sudden deaths, *Journal of Financial Economics* 98, 550–567.

Ozsoylev, Han N, and Johan Walden, 2011, Asset pricing in large information networks, *Journal of Economic Theory* 146, 2252–2280.

———, M Deniz Yavuz, and Recep Bildik, 2014, Investor networks in the stock market, *Review of Financial Studies* 27, 1323–1366.

Page, Lawrence, Sergey Brin, Rajeev Motwani, and Terry Winograd, 1999, The pagerank citation ranking: bringing order to the web., .

Peress, Joel, 2010, Product market competition, insider trading, and stock market efficiency, *The Journal of Finance* 65, 1–43.

Pool, Veronika K, Noah Stoffman, and Scott E Yonker, 2015, The people in your neighborhood: Social interactions and mutual fund portfolios, *forthcoming in The Journal of Finance.*

Puckett, Andy, and Xuemin Sterling Yan, 2011, The interim trading skills of institutional investors, *The Journal of Finance* 66, 601–633.

Shiller, Robert J, 1995, Conversation, information, and herd behavior, *The American Economic Review* 85, 181–185.

——— , 2000, *Irrational exuberance* (Princeton University Press).

——— , and John Pound, 1989, Survey evidence on diffusion of interest and information among investors, *Journal of Economic Behavior & Organization* 12, 47–66.

Stein, Jeremy C, 2008, Conversations among competitors, *The American Economic Review* 98, 2150–2162.

Tookes, Heather E, 2008, Information, trading, and product market interactions: Cross-sectional implications of informed trading, *The Journal of Finance* 63, 379–413.

Valente, Thomas W, Kathryn Coronges, Cynthia Lakon, and Elizabeth Costenbader, 2008, How correlated are network centrality measures?, *Connections* 28, 16.

Walden, Johan, 2014, Trading, profits, and volatility in a dynamic information network model, *Available at SSRN 2561055.*

Wooldridge, Jeffrey M, 2010, *Econometric analysis of cross section and panel data* (MIT Press) 2 edn.

8 Tables and Figures

Table 1: **Fund Summary Statistics 1991Q1–2011Q3.** Interim trading performance is measured quarterly as the difference between a fund's principal- or equal-weighted average excess buy return minus its average excess sell return. Returns are computed as the simple return from the actual execution price to the end of quarter closing price including dividends and are also DGTW-adjusted over the same holding period. The principal weight is computed using the actual price $\times$ volume traded. Prices and volume are adjusted for stock splits, and benchmark returns are adjusted for delistings. Centrality is Information Diffusion Centrality defined in Eq. (9), and Eigenvector is Eigenvector Centrality defined in Eq. (8). Both measures are computed based on quarterly trades and reported as percentages of the "total" centrality. Degree is the sum of a fund's weighted edges (trading correlations) a given quarter. Volume refers to fund-quarter share volume traded, and # of Trades refers to the actual number of transactions. Correlations in Panel B are Spearman correlations and the diagonal entries are autocorrelations.

(a) Summary Statistics

	N	Mean	Med.	Std	Skew	Kurt
PW Ex. Int. Perf.	59,374	0.52%	0.32%	6.3%	0.03	5.51
EW Ex. Int. Perf.	59,374	0.65%	0.44%	5.8%	-0.14	6.04
Centrality	59,374	2.9%	2.8%	0.42%	1.02	2.28
Eigenvector	59,374	0.74%	0.07%	2.8%	7.31	60.23
Degree	59,374	5	4	3	2.37	7.79
Volume	59,374	18.53M	680,512	182.85M	34.07	2,382.43
# of Trades	59,374	2,340	174	20,167	25.60	1,004.18

(b) Correlations

	PW	EW	Cent	Eig	Deg	Vol	Trades
PW	0.048	.	.	.	.	.	.
EW	0.77	0.065	.	.	.	.	.
Cent	0.0021	0.0036	0.91	.	.	.	.
Eig	-0.051	-0.070	0.52	0.72	.	.	.
Deg	-0.026	-0.021	0.23	0.55	0.91	.	.
Vol	-0.030	-0.023	0.14	0.20	0.21	0.84	.
Trades	-0.00048	0.014	0.31	0.25	0.37	0.66	0.87

Table 2: **Event Summary Statistics.** Merger announcements come from the Thomson Reuters SDC Platinum database and director and manager sudden deaths and new product announcements come from the S&P Capital IQ Key Developments Database (event types 101, 102, 16 for deaths and 41 for new products). CARs are measured from [-30, +30] using a one-factor CAPM model estimated over a 252-day window with a 30-day gap between the estimation window and the event window. Round-trip trading performance is measured as the principal- or equal-weighted average of all signed simple excess returns of buy (sell) trades initiated between [-60, -1] and closed between [0, +30]. t-statistics are based on un-adjusted standard errors.

(a) Distribution of Events

	Mergers		Deaths		New Products	
Year	# Events	# Funds	# Events	# Funds	# Events	# Funds
1999	890	1,102	.	.	.	.
2000	858	1,113	.	.	.	.
2001	543	1,019	1	2	112	385
2002	526	1,085	4	101	6,498	1,506
2003	610	964	13	62	6,840	1,430
2004	705	855	28	150	8,432	1,285
2005	734	832	36	216	8,952	1,144
2006	739	713	38	171	8,637	996
2007	757	714	49	210	8,526	942
2008	627	743	55	240	9,347	987
2009	510	745	42	191	8,781	996
2010	467	627	34	231	8,584	842
2011	307	487	21	142	5,806	742

(b) Event Returns

Event	# Events	# Trades	CAR	t	PW	t	EW	t
M&A Targets	8,230	181,189	5.47%	16.91	1.23%	33.12	1.19%	34.16
Deaths	317	5,512	−3.12%	−2.39	−0.27%	−1.65	−0.25%	−1.70
New Products	79,715	2,887,988	−0.90%	−10.70	−0.03%	−5.47	−0.05%	−9.07

Table 3: **Interim Trading Performance and Centrality.** The dependent variable is the interim trading performance. Centrality refers to Information Diffusion Centrality, and Eigenvector refers to Eigenvector Centrality. All independent variables are log transformed and standardized. Coefficients and R^2 values are reported in percentages. t-statistics are reported in parentheses, with standard errors clustered by fund and quarter. Standard errors for the random effects specifications are computed using a block-bootstrap with 100 repetitions. Stars indicate significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

	Principal Weighted				Equal Weighted			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Centrality	0.300***		0.327***	0.201***	0.336***		0.328***	0.180***
	(3.91)		(4.00)	(3.34)	(4.96)		(4.20)	(3.05)
Degree	-0.378***	-0.197***	-0.442***	-0.270***	-0.364***	-0.136**	-0.411***	-0.227***
	(-3.44)	(-2.91)	(-3.88)	(-3.74)	(-3.52)	(-2.12)	(-4.11)	(-3.30)
Eigenvector		0.102*	0.107*	0.0740***		0.0615	0.0710	0.0221
		(1.77)	(1.81)	(3.21)		(1.03)	(1.16)	(1.07)
# Trades (Log)			0.176**	0.174**			0.247***	0.263***
			(2.03)	(2.30)			(2.87)	(3.63)
Volume (Log)			-0.299***	-0.302***			-0.307***	-0.312***
			(-4.89)	(-5.05)			(-5.31)	(-5.14)
Qtr FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fund RE	No	No	No	Yes	No	No	No	Yes
R^2(%)	0.15	0.077	0.28	0.26	0.17	0.042	0.30	0.27
Obs.	59,374	59,374	59,374	59,374	59,374	59,374	59,374	59,374

Table 4: **Round-trip Trading around M&A Announcements — Targets.** Merger announcements come from the Thomson Reuters SDC Platinum Database. The dependent variable is the round-trip trading performance. Centrality refers to Information Diffusion Centrality, and Eigenvector refers to Eigenvector Centrality. All independent variables are log transformed and standardized. Coefficients and R^2 values are reported in percentages. t-statistics are reported in parentheses, with standard errors clustered by fund and quarter. Standard errors for the random effects specifications are computed using a block-bootstrap with 100 repetitions. Stars indicate significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

	Principal Weighted				Equal Weighted			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Centrality	1.484**		1.805***	1.131***	1.542**		1.929***	1.267***
	(2.35)		(3.25)	(3.24)	(2.44)		(3.43)	(3.66)
Degree	-2.010***	-0.919***	-2.383***	-1.312***	-2.064***	-0.921***	-2.443***	-1.370***
	(-2.74)	(-3.70)	(-2.96)	(-6.25)	(-2.84)	(-3.75)	(-3.06)	(-6.62)
Eigenvector		0.524***	0.485**	0.141*		0.529***	0.475**	0.136*
		(3.10)	(2.39)	(1.84)		(3.10)	(2.34)	(1.80)
# Trades (Log)			-0.207	0.110			-0.252	0.0852
			(-0.53)	(0.49)			(-0.66)	(0.38)
Volume (Log)			-0.174	-0.513**			-0.222	-0.559**
			(-0.61)	(-2.34)			(-0.80)	(-2.56)
Event FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fund RE	No	No	No	Yes	No	No	No	Yes
R^2(%)	0.31	0.27	0.45	0.39	0.37	0.32	0.56	0.47
Obs.	181,189	181,189	181,189	181,189	181,189	181,189	181,189	181,189

Table 5: **Tipping around M&A Announcements — Targets.** A fund is connected to a brokerage house (Tip=1) if the brokerage house was one of the funds' main brokers in the last quarter. A fund's brokers are ranked according to share volume executed in each quarter, and top decile brokers are identified as main brokers, see Jegadeesh and Tang (2011) for more details. All other independent variables are log transformed and standardized. Coefficients and R^2 values are reported in percentages. Standard errors for the random effects specifications are computed using a block-bootstrap with 100 repetitions. t-statistics are reported in parentheses, with standard errors clustered by fund and quarter. Stars indicate significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

	Principal Weighted				Equal Weighted			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Tip	0.383	0.215	0.418	-0.0159	-0.0838	-0.322	-0.0560	-0.483
	(0.57)	(0.35)	(0.62)	(-0.026)	(-0.13)	(-0.58)	(-0.085)	(-0.80)
Centrality	1.822***		1.809***	1.132***	1.947***		1.933***	1.269***
	(3.27)		(3.25)	(3.24)	(3.47)		(3.44)	(3.66)
Cent. × Tip	-0.0272		-0.447	-0.171	-0.331		-0.649	-0.377
	(-0.052)		(-0.80)	(-0.30)	(-0.69)		(-1.17)	(-0.73)
Degree	-2.073***	-0.849*	-2.385***	-1.313***	-2.139***	-0.802*	-2.444***	-1.370***
	(-2.89)	(-1.90)	(-2.96)	(-6.25)	(-3.01)	(-1.82)	(-3.06)	(-6.62)
Volume (Log)	-0.250	-0.216	-0.174	-0.513**	-0.296	-0.266	-0.221	-0.557**
	(-0.84)	(-0.71)	(-0.61)	(-2.34)	(-1.01)	(-0.89)	(-0.80)	(-2.56)
# Trades (Log)	-0.224	0.0878	-0.208	0.110	-0.268	0.0632	-0.253	0.0841
	(-0.57)	(0.19)	(-0.53)	(0.49)	(-0.71)	(0.14)	(-0.66)	(0.37)
Eigenvector		0.492**	0.481**	0.139*		0.484**	0.472**	0.134*
		(2.49)	(2.37)	(1.81)		(2.46)	(2.33)	(1.78)
Eig. × Tip		0.612**	0.835***	0.560		0.321	0.636**	0.338
		(2.45)	(2.89)	(1.20)		(1.37)	(2.06)	(0.87)
Event FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fund RE	No	No	No	Yes	No	No	No	Yes
R^2(%)	0.38	0.28	0.45	0.39	0.47	0.34	0.56	0.47
Obs.	181,189	181,189	181,189	181,189	181,189	181,189	181,189	181,189

Table 6: **Round-trip Trading around New Product Announcements.** This table includes the sample of round-trip trades made by funds around the announcements of new products. New product events are identified in the S&P Capital IQ Key Events Database as event type 41. Hoberg and Phillips (2010, 2015) classify firms into industries (`TNICs`) using the product descriptions found in 10-K filings. Each product announcement in my sample is assigned to an industry based on yearly `GVKEY-TNIC` pairs. I compute a fund's product market expertise as the trade-weighted average cosine similarity between the announcing firm's product description and the product descriptions of all of the other firms in which the fund traded stocks. Covariates are standardized, and coefficients and R^2 values are reported in percentages. t-statistics are reported in parentheses, with standard errors clustered by fund and quarter. Standard errors for the random effects specifications are computed using a block-bootstrap with 100 repetitions. Stars indicate significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

	Principal Weighted		Equal Weighted	
	(1)	(2)	(3)	(4)
Expertise	0.0442**	0.0411***	0.0320*	0.0289**
	(2.47)	(3.09)	(1.75)	(2.51)
Centrality	0.392**	0.130	0.367**	0.190
	(2.46)	(0.98)	(2.28)	(1.52)
Cent. × Expert.	0.0254*	0.0231*	0.0162	0.0134
	(1.92)	(1.81)	(1.43)	(1.27)
Eigenvector	0.145*	0.0482	0.135	0.0364
	(1.71)	(1.41)	(1.63)	(1.20)
Degree	-0.547***	-0.313**	-0.523***	-0.324***
	(-3.39)	(-2.56)	(-3.27)	(-2.71)
Volume (Log)	-0.133	-0.199*	-0.134	-0.166*
	(-1.34)	(-1.88)	(-1.37)	(-1.73)
Trades (Log)	-0.00718	0.207**	-0.0323	0.103
	(-0.065)	(2.19)	(-0.30)	(1.17)
Event FE	Yes	Yes	Yes	Yes
Fund RE	No	Yes	No	Yes
R^2(%)	0.060	0.037	0.081	0.066
Obs.	2,887,988	2,887,988	2,887,988	2,887,988

Table 7: **Round-trip Trading around Sudden Director & Manager Deaths.** Director or manager changes are identified in the S&P Capital IQ Key Events Database as event types 101 (CEO), 102 (CFO), and 16 (Board and Other Executive). I identify sudden deaths by searching the headlines for key phrases including `death of`, `demise of`, and `passing of`. All independent variables are log transformed and standardized. Coefficients and R^2 values are reported in percentages. t-statistics are reported in parentheses, with standard errors clustered by fund and quarter. Standard errors for the random effects specifications are computed using a block-bootstrap with 100 repetitions. Stars indicate significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

	Principal Weighted				Equal Weighted			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Centrality	1.017		0.960	0.826	0.792		0.733	0.683
	(1.38)		(1.41)	(1.16)	(1.10)		(1.08)	(1.04)
Degree	-1.058	-0.148	-0.929	-0.707	-0.836	-0.0630	-0.658	-0.397
	(-1.34)	(-0.73)	(-1.12)	(-1.07)	(-1.07)	(-0.30)	(-0.82)	(-0.66)
Eigenvector		0.0943	-0.0432	-0.0547		-0.0189	-0.144	-0.151
		(0.60)	(-0.25)	(-0.24)		(-0.12)	(-0.86)	(-0.83)
# Trades (Log)			0.869**	0.733			0.888**	0.610
			(2.14)	(1.31)			(2.44)	(1.22)
Volume (Log)			-1.137***	-1.049**			-1.115***	-1.002**
			(-2.95)	(-2.09)			(-2.99)	(-2.09)
Event FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fund RE	No	No	No	Yes	No	No	No	Yes
R^2(%)	0.069	0.0087	0.32	0.32	0.053	0.0056	0.35	0.32
Obs.	5,512	5,512	5,512	5,512	5,512	5,512	5,512	5,512

Table 8: **Matched Sample Regression.** Central is a dummy variable which is one if a fund has above median centrality in a given quarter and is zero otherwise. I estimate the propensity of being Central as a function of Degree, Eigenvector, # of Trades, and Volume (all in logs). I require that the maximum difference between the propensity score of a Central fund and its matching peer is no more than 0.1 basis points in absolute value. All independent variables are log transformed and standardized. Coefficients and R^2 values are reported in percentages. t-statistics are reported in parentheses, with standard errors clustered by fund and quarter. Standard errors for the random effects specifications are computed using a block-bootstrap with 100 repetitions. Stars indicate significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

(a) Regressions

	Principal Weighted			Equal Weighted		
	(1)	(2)	(3)	(4)	(5)	(6)
Central	0.446***	0.492***	0.250***	0.491***	0.545***	0.241***
	(4.46)	(4.06)	(3.09)	(5.71)	(5.53)	(2.76)
Degree	-0.175**	-0.196**	-0.164**	-0.138**	-0.184**	-0.166**
	(-2.63)	(-2.57)	(-2.23)	(-2.23)	(-2.23)	(-2.46)
Eigenvector	0.110*	-0.0754	-0.0282	0.0742	-0.127	-0.0589
	(1.88)	(-0.96)	(-0.46)	(1.23)	(-1.62)	(-1.25)
# Trades (Log)	0.182**	0.248**	0.287***	0.244***	0.266**	0.375***
	(2.15)	(2.38)	(2.88)	(2.94)	(2.50)	(3.86)
Volume (Log)	-0.278***	-0.344***	-0.378***	-0.283***	-0.283***	-0.358***
	(-4.53)	(-4.60)	(-4.48)	(-4.88)	(-3.98)	(-4.35)
Qtr FE	Yes	Yes	Yes	Yes	Yes	Yes
Matched	No	Yes	Yes	No	Yes	Yes
Fund RE	No	No	Yes	No	No	Yes
R^2(%)	0.30	0.50	0.32	0.36	0.53	0.30
Obs.	59,374	34,324	34,324	59,374	34,324	34,324

(b) Normalized Differences

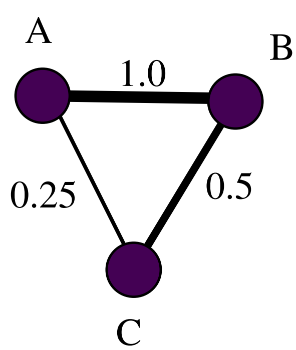
	Peripheral	Central	Norm. Diff.
Degree	4.65	4.85	0.059
Eig.	0.51%	0.46%	-0.019
# of Trades	434	2040	0.12
Volume	4.43M	20.68M	0.087
N	11,840	17,162	.

Table 9: **ANCerno Data (1991Q1–2012Q1).** A fund is defined as a unique *clientcode–managercode* pair. The sample of stocks traded includes all common stock listed on the NYSE, AMEX, and NASDAQ exchanges which I am able to match to a CRSP `PERMNO`s using CUSIPs and tickers. Brokers are identified by *brokercode*. Prices and volume are adjusted for stock splits using adjustment factors from the CRSP database.

Year	# of Funds	# Stocks Traded	# Brokers	# of Trades	Volume Traded	Dollar Value Traded
1999	1,992	4,921	651	1.98M	18.17B	904.88B
2000	1,901	4,814	641	3.15M	32.67B	1.72T
2001	1,996	4,557	693	3.59M	40.97B	1.32T
2002	2,105	4,348	747	3.57M	39.81B	1.02T
2003	2,068	4,315	729	3.77M	34.45B	859.37B
2004	1,950	4,256	671	4.42M	33.64B	985.22B
2005	1,612	4,057	687	4.39M	31.42B	1.02T
2006	1,404	3,984	665	7.32M	47.10B	1.55T
2007	1,296	3,928	638	8.64M	50.87B	1.84T
2008	1,285	3,771	668	7.32M	57.82B	1.71T
2009	1,304	3,712	648	6.49M	58.47B	1.28T
2010	1,220	3,350	652	5.19M	29.57B	870.19B
2011	1,223	3,387	659	4.80M	17.47B	561.19B
2012	238	2,657	186	506,477	1.45B	50.49B

Figure 1: **Triangle Network.** A network is defined as a collection of nodes (●) and edges (▬). Panel A shows the geometric representation of the network and Panel B shows the (weighted) adjacency matrix representation. The adjacency matrix is symmetric, and the edge weights are on the off-diagonal entries.

(a) Geometric Representation



(b) Adjacency Matrix

$$A = \begin{bmatrix} 0 & 1 & 0.25 \\ 1 & 0 & 0.5 \\ 0.25 & 0.5 & 0 \end{bmatrix}$$

Figure 2: **Weighted Trading Correlation Network with an active liquidity trader.** The "true" star network consists of five informed investors (●) connected to one another via black lines (━). Adding an active liquidity trader (●) results in many un-informative linkages to the informed investors. Using Weighted Trading Correlation Network estimation attenuates the strength of the un-informed connections (━). Node size corresponds to the sum of the weighted edges such that the liquidity trader is "unimportant" and on the periphery.

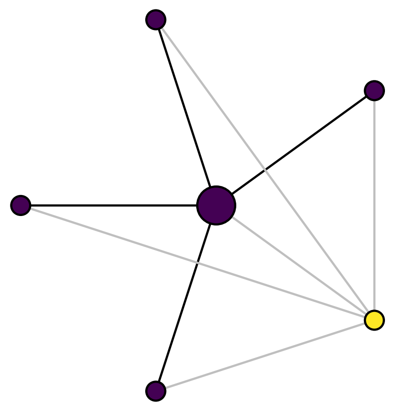
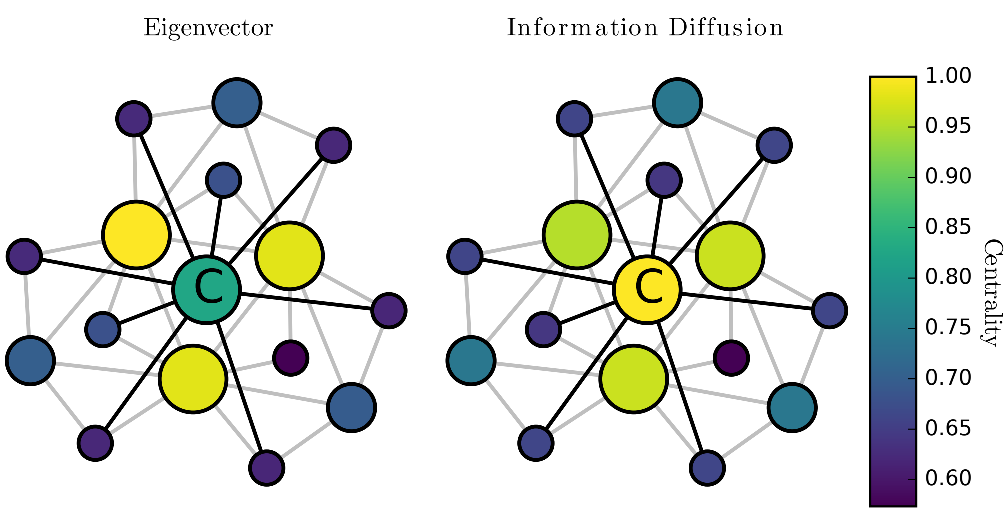


Figure 3: **Eigenvector vs. Information Diffusion Centrality.** The network is hierarchically constructed following Dorogovtsev, Goltsev, and Mendes (2002) such that there are three "core" nodes connected to 12 "peripheral" nodes with gray lines (━). I add an additional node C and connect it to eight peripheral nodes with black lines (━) so that C has exactly as many connections as the three core nodes. Node size is proportional to the number of connections and node color is proportional to Eigenvector Centrality on the left and Information Diffusion Centrality on the right. According to Eigenvector Centrality C is not central because it is not connected to any of the three core nodes. In contrast, according to Information Diffusion Centrality C is the most central.



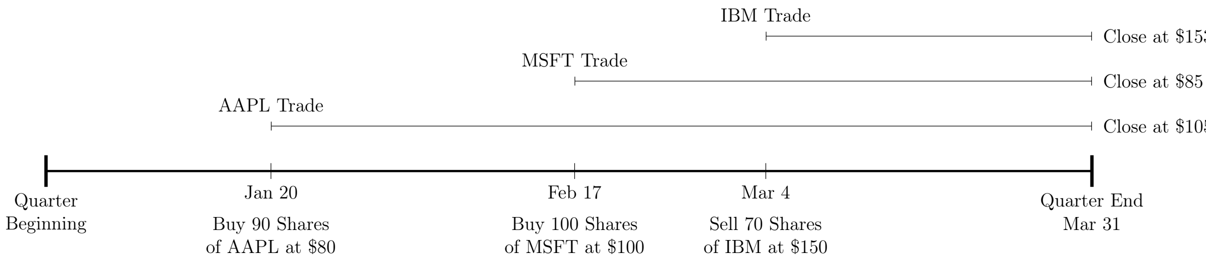


Figure 4: Interim Trading Example

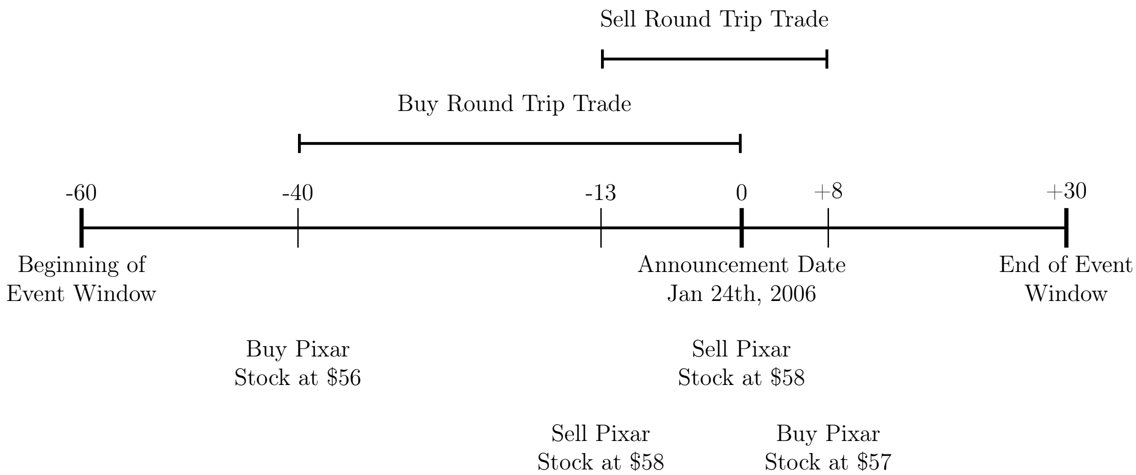
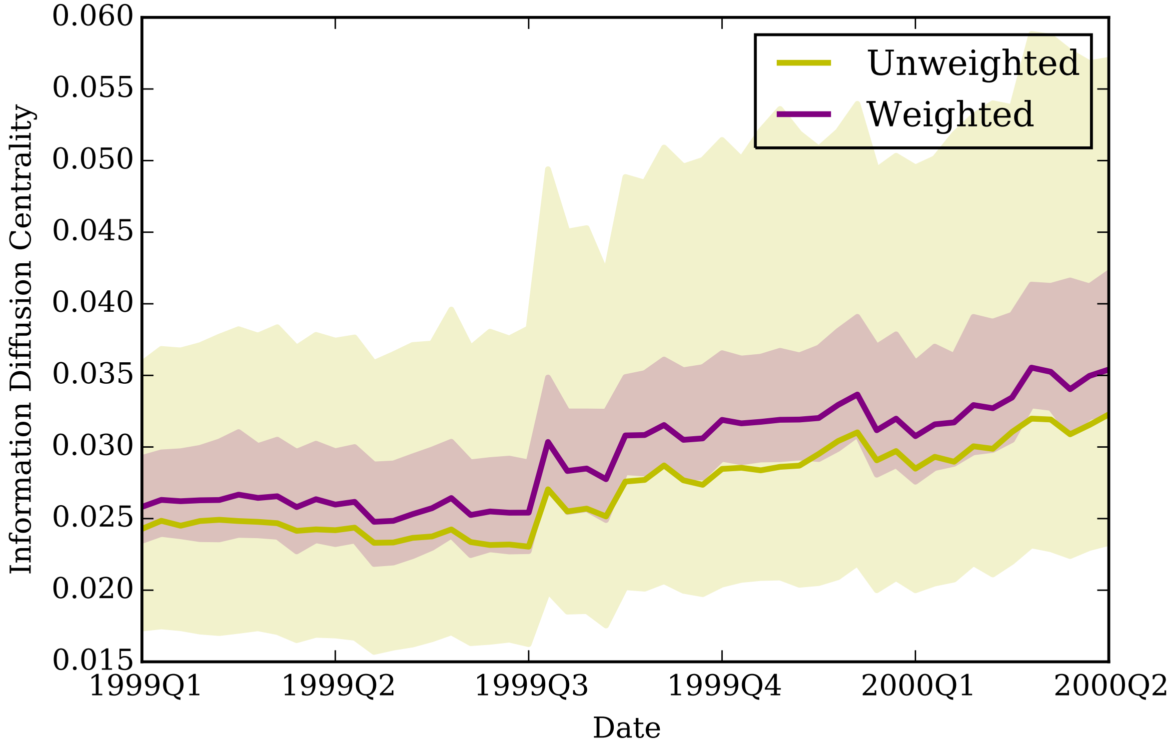


Figure 5: **Round Trip Trades around Disney-Pixar Deal.**

Figure 6: **Centrality 1991Q1–2011Q3.** This Figure plots the distribution of Information Diffusion Centrality (Eq. 9) over the sample period. The solid lines represent the median, and the shaded regions represent the 5 and 95 percentiles of Information Diffusion Centrality. The Figure includes Information Diffusion Centrality based on Unweighted (–), and Weighted (–) edges. Edge-weighting based on trade correlations significantly reduces the right skew and kurtosis of Information Diffusion Centrality.



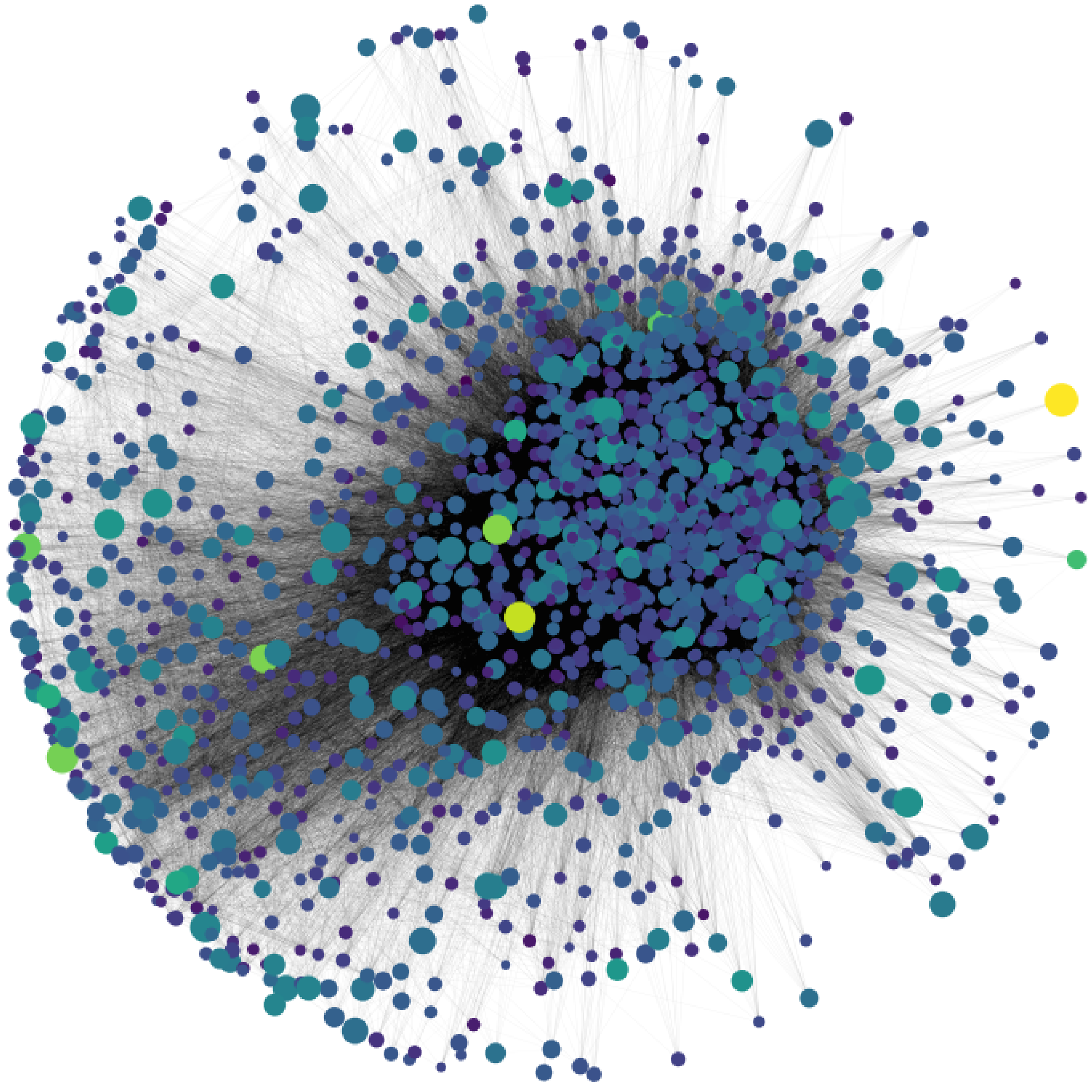


Figure 7: **Weighted Trading Correlation Network (2002Q2).** Nodes represent funds and edges represent trading correlations. Nodes are placed using a force-directed layout with edge weights corresponding to trading correlations. Lighter edges indicate lower trading correlations and darker edges indicate higher trading correlations. Node size is proportional to the number of edges, and node color is proportional to Information Diffusion Centrality. Lighter colored nodes (○) are "central" while darker colored nodes (●) are "peripheral."