Information Networks: Implications for Mutual Fund Trading Behavior and Stock Returns

Ankur Pareek[1]

Rutgers University

February 2012

Abstract

This paper examines the effect of information networks on the trading behavior of mutual funds and on stock returns. An information or stock ownership linkage between two mutual funds is defined by large positions in the same stock. Mutual funds trade together with other funds in their information network after controlling for the overall trading behavior of the mutual fund sector. The effect is robust and cannot be explained by style investing or geographic location. The paper also examines the effect of the structure of information networks on stock returns and stock volatility. Using network density as a measure for the speed of information diffusion in a network of investors, I find that stocks with a lower network density demonstrate stronger return momentum over medium horizons and also show a delayed response to the market-wide information. The evidence is consistent with the gradual information diffusion model of Hong and Stein (1999). Finally, I provide empirical evidence in support of recent theoretical models that study the asset pricing implications of social networks. I show that centralized information networks lead to a higher volatility of individual stocks in the cross-section and also explain the variation in average stock idiosyncratic volatility over time.

[1] This paper is based on a chapter of my dissertation at Yale University. I am grateful to my advisors Andrew Metrick (Dissertation Chair), Nick Barberis and Martijn Cremers for many helpful discussions. I also thank Jonathan Berk, James Choi, Lauren Cohen, Will Goetzmann, Roger Ibbotson, Jonathan Ingersoll, Simi Kedia, Dasol Kim, Panagiotis Patatoukas, Antti Petajisto, Geert Rouwenhorst, Heather Tookes, Johan Walden, Roy Zuckerman and seminar participants at the Cornell University, Rutgers University, Yale University, 2010 American Finance Association Annual Meeting, 2010 Midwest Finance Association Annual Meeting and 2010 Eastern Finance Association Annual Meeting for their valuable suggestions. In addition, I am grateful to Antti Petajisto for providing the data on benchmark index holdings.

Author's contact information: ankur.pareek@business.rutgers.edu ; tel. +1-973-353-1646

Networks play an important role in the flow of information through financial markets. Investors often rely on their professional and social ties to access information needed to make their trading decisions (e.g., Cohen, Frazzini and Malloy (2008), Hong, Kubik and Stein (2005)). Although the existence of information networks is well established, little is known about their effect on stock prices[2]. If investors learn from each other through their information networks, the structure of these networks can help us understand the process of information incorporation into stock prices. Network structure could also provide an explanation for anomalous stock price behavior, such as stock return momentum (Jegadeesh and Titman (1993)).

In this paper, I study the effect of information networks on stock returns and volatility. I introduce a novel approach for identifying the information networks of mutual fund managers based on the large portfolio holdings they have in common. Specifically, I show that mutual fund managers who invest a large portion of their portfolios in the same stocks are likely to be informationally connected to each other. My new holdings-based definition of information networks allows me to test some theoretical predictions concerning the effect of network structure on stock returns and stock volatility. Using size-adjusted network density as a measure for the speed of information diffusion, I find that stocks whose information network has a lower density demonstrate stronger return momentum in the short to medium horizon and their prices show a delayed response to the market-wide information. I also show that the centralization of information networks can explain the variation of idiosyncratic volatility in the cross-section of individual stocks and over time for the aggregate economy.

[2] In general, information network refers to a group of investors who share common information either through their social connections as in a social network or by accessing private information through the same sources.

The intuition behind the holdings-based definition of information networks is that the large positions of mutual funds managers are likely to be based on their private information signal rather than on their investment style or on noise. Therefore, the endowments of private information of two mutual fund managers investing a large portion of their portfolios in the same stock are likely to be correlated with each other. Recent empirical studies provide evidence in support of the above hypothesis. For example, Bushee and Goodman (2007) show that institutional investors holding large positions in a given stock posses private information about that stock. In another recent study, Cohen, Polk and Silli (2010) show that the best ideas of mutual fund managers (defined by portfolio holdings with large deviation from a benchmark index) outperform rest of their portfolio and other benchmarks. These best ideas are thus likely to be based on fund manager's private information.

The empirical approach in this paper is as follows. First, I define the stock ownership or information network for a fund in a given quarter as all other funds which hold more than 5% of their portfolio in any stock in which the fund has also invested at least 5% of its portfolio. The results are robust to an alternative definition of information linkages where two funds with large active positions (measured as deviation from their benchmark index) in the same stock are hypothesized to be informationally connected to each other. I check empirically whether fund managers are likely to exchange ideas or share correlated sources of private information with other managers in their information network. I examine the existence of information networks by measuring the correlation of trades of fund managers with other managers in the same stock ownership network after controlling for the average trading behavior of *all* mutual funds. I find that after controlling for the aggregate trading of the mutual fund sector, a given fund manager increases her holdings of a stock as a fraction of her total portfolio by 0.22% when other managers in her stock ownership network increase the average holding of that stock by 1 percentage point of their portfolio. This result strongly confirms the existence of holdings-based information networks introduced in this paper.

The most likely alternative explanation for the correlated trading within the same network is style investing. I show that the results in this paper remain robust after controlling for style investing. The correlation of the trades of a mutual fund with the average trades of both the same-style funds and different investment style funds in its stock ownership network is significant after controlling for the aggregate mutual fund trading. The same-style funds are defined as funds holding stocks with similar size and book-to-market characteristics, or as funds which have the same benchmark index as the given fund. I also include fund-level style controls corresponding to book-to-market, size, momentum and industry characteristics to control for trading due to style investing. Further, I examine the interaction between a fund manager's location network and stock ownership network. I find that the correlated trading by a mutual fund manager with other fund managers in the same geographic location (Hong, Kubik and Stein (2005)) is limited to the fund managers who are also in her stock ownership network i.e., who also share a large position in at least one stock with that fund manager. This empirical observation combined with the result in the previous literature that correlated trading in the same geographic location is due to information diffusion shows that at least part of the trading within the stock ownership network is due to information diffusion.

Having shown that the holdings-based networks I define can capture the flow of private information, I then use these networks to test some empirical predictions from recent theoretical models of information diffusion. Next, I examine the effect of information network structure on future stock returns. I aggregate the information networks of the funds holding a large position (greater than or equal to 5%) or a large active position in a given stock to determine the information or ownership network corresponding to that stock. This information network for the stock corresponds to the set of fund managers amongst whom information about that stock diffuses through word of mouth or through common sources of private information. I then use these networks to test a prediction of Hong and Stein (1999) that momentum comes from the slow diffusion of information across a heterogeneous set of investors. If information diffuses gradually through the stock ownership network, we should expect this process of information diffusion to take longer if the information network for that stock has a lower "density". Network density is widely used in the social-network literature to measure the efficiency of information diffusion and is defined by the number of connections in a network as a proportion of the maximum number of possible connections (Yamaguchi (1994)). To control for mechanical correlation with stock market capitalization, I regress network density on stock market cap and use the residual ("residual network density") in my return predictability analysis. I find that the equal-weighted, long-short 3-factor (Fama-French) momentum returns for a six month holding period are a significant 0.34% per month higher for stocks in the lowest residual network density group compared to stocks in the top residual network density group. I also find that stocks with lower network density show a delayed response to market wide information measured as lagged market return or lagged industry return (see Hou and Moskowitz (2005), Hou (2007)) and thus have stronger lead-lag effects. This result provides further support to the hypothesis that information diffuses gradually through the holdings-based networks and validates network density as a measure for the speed of information diffusion.

The analysis in this paper is related to the recent theoretical papers (Cao and Xia (2006), Ozsoylev (2007), Colla and Mele (2010), Ozsoylev and Walden (2010)) that examine the effect of information networks on asset prices. Ozsoylev (2007) proposes a rational expectations model in which agents learn from each other through an exogenous social network. The model predicts that centralized social networks can lead to excess volatility in financial markets. The intuition behind this prediction is as follows. In a network with a high degree of centralization, changes in price are significantly dependent on changes in central nodes' signals. Whereas in a network with lower centralization, changes in prices are relatively equally dependent on changes in each mutual fund's private signal, and, given the independence between the error terms of their private signals most of these changes wash out each other. Therefore, the information-driven volatility component increases with an increase in network centralization. I test this prediction by examining the relationship between measures of network centralization and stock volatility.

For individual stocks, I measure the centralization of their information network as the coefficient of variation of the number of connections of the mutual funds present in the stock ownership network of that stock at the end of each quarter. Similarly, to estimate the economy-wide centralization measure, I calculate the coefficient of variation of the number of connections of all the connected mutual funds in the economy at any given point of time. I find that in the cross-section, measures of the centralization of a stock's ownership network positively predict its idiosyncratic volatility in the next period after controlling for lagged volatility and other stock characteristics known to forecast volatility. In the aggregate, the centralization measure of an information network consisting of *all* US equity mutual fund managers predicts the variation in average stock idiosyncratic volatility over time. The prediction is confirmed by a Granger causality test: the aggregate centralization measure Granger-causes average idiosyncratic volatility while the causality is insignificant the other way around.

The idea that institutional investors draw other investors' attention to individual stocks through word of mouth was first studied through institutional investor surveys by Shiller and Pound (1986). Based on responses to survey questions, they provide evidence that institutional investors actively talk to other institutional investors holding the same stock. Empirically, this paper is most closely related to Hong, Kubik and Stein (2005) who show that the trading of particular stocks in a given quarter by mutual funds located in the same city is correlated. This paper makes two main contributions to the literature. First, it introduces a new methodology for identifying the information linkages between mutual funds based on their large portfolio holdings in the same stocks. Second, this paper is the first to test the effect of the structure of information networks on the behavior of stock prices.

The remainder of this paper is organized as follows. In the next section, I discuss the network analysis methodology used in this paper. In section II, I describe the data sample. In section III, I test for the existence of stock ownership based information networks by examining whether the trading of mutual funds is correlated with other funds in their network. In section IV, I examine the impact of gradual information diffusion through information networks on momentum returns and price delay. In section V, I examine the effect of network centralization on stock volatility. Finally, in section VI, I conclude with a brief summary of the results.

I. Network Analysis Methodology

The efficiency of stock prices depends on the widespread dissemination of firm specific information and on the timely incorporation of this information into prices. Social network theory provides a promising toolbox for analyzing the efficiency of information diffusion through a network of investors in financial markets. In this section, I give a formal definition of the stock ownership based information network introduced earlier and discuss the construction of the network measures used in this paper.

A. Stock Ownership Network

I first determine the set of mutual fund managers allocating more than 5% of their portfolio to a stock *i*. In this paper, I hypothesize that two fund managers allocating 5% or more of their portfolio to the same stock are connected to each other. This gives the set of mutual fund managers to which a given mutual fund manager *j* is connected in a quarter through similar large holdings. This set is denoted by $FUNDCONN_j$ and is referred to as the fund manager's information network[3]. The number of unique fund managers in a mutual fund manager *j's* information or stock ownership network, also called the *degree centrality* of fund manager *j*, is $NFUNDCONN_j$. In the appendix, I show that the results remain unchanged for an alternative network specification where in a given quarter two funds with large active positions (defined by absolute value of the portfolio weight deviation from the benchmark index) in the same stock are hypothesized to be connected to each other.

Then, for each stock, I aggregate the information networks of fund managers who invest more than 5% of their portfolio in a given stock to get the information network of fund managers for every stock ($STOCKCONN_i$). Every fund manager is included only once in the network. This network of fund managers corresponds to the set of managers among which information about that stock is likely to diffuse through word of mouth or through common information sources.

[3] I exclude funds that belong to the same management company as a given fund from its stock ownership network because trades within the same fund management company may be correlated due to reasons not related to network effects e.g. due to internal information transfer. The results remain similar if the same management company funds are also included in the ownership network of a fund.

The bigger this network, the more central a stock is and the greater the number of mutual fund managers likely to get information about that stock. The *degree* of stock *i's* ownership network is denoted by $NSTOCKCONN_i$.

To illustrate the definition of the information network of a fund, I show the information networks of two mutual funds: "Brown Capital Management Small Companies" and "FPA Paramount" at the end of December 2005 in Figures 1 and 2 respectively. As shown in Figure 1, Brown Capital Management Small Companies fund is connected to 6 other funds through an investment greater than 5% of its portfolio in 11 stocks. Similarly, FPA Paramount is connected to 4 other funds including Brown Capital Small Companies fund through investment greater than 5% in 6 stocks. Thus, the degrees of information networks (NFUNDCONN) of Brown Capital Small Companies and FPA Paramount are 6 and 4 respectively. Next, I present an example of the information network of a stock. Figure 3 shows the stock ownership network of Cognex Corporation in December 2005. Brown Capital Small Companies and FPA Paramount are the two funds with more than 5% of their portfolio invested in Cognex Corp (CGNX) at the end of year 2005. All other funds in the information network of these two funds are also in the ownership network of CGNX. After combining the information networks of Brown Capital Small Companies and FPA Paramount, the ownership network of CGNX includes 10 unique funds, which is the degree of stock ownership network (NSTOCKCONN) of CGNX.

B. Information Diffusion Measure

A number of papers in the social-network literature study the efficiency of information diffusion through networks. The term "efficiency" refers to how fast information is expected to flow through a network with a particular structure. Yamaguchi (1994) and Buskens and Yamaguchi

(1998) show that the structural characteristics of a social network strongly affect the efficiency of information flow. Yamaguchi (1994) develops a Markov chain based model of information flow and then relates the theoretical time taken for information to flow across a network with simple measures of its structure (e.g., network density). He shows that *increasing network density leads to faster information diffusion* through the network. Based on this research, I use network density as a measure of the speed of information diffusion. Network density or global density is defined as the number of ties in a network divided by the maximum number of possible ties. The maximum number of ties possible in a network is n(n-1)/2, where n is the number of funds in the network. For example, the number of links in the ownership network of the stock Cognex Corp shown in Figure 3 is 14 and the corresponding maximum number of possible links is 45. Therefore, the density of the stock ownership network of Cognex Corp in December 2005 is 0.31.

Consistent with the social network theory literature, I find that the network density of a stock's information network is a decreasing function of its network size and thus of the stock's market capitalization. As shown in Panel B of Table I, the density of a stock's information network (NETDENSITY) is highly negatively correlated (correlation coefficient of -0.40) with its market cap. This makes comparison of network density across stocks of different sizes difficult. To eliminate the mechanical effect of stock market capitalization and to allow for comparison across the stock ownership networks of different stocks, each quarter I regress the logit transformation of network density ($NETDENSITY_i$) of stock i on the logarithm of its market capitalization (Equation 1).

$$\ln\left(\frac{NETDENSITY_{it}}{1-NETDENSITY_{it}}\right)=\alpha_t+\beta_t\log(MCAP_{it})+\varepsilon_{it}$$
$$RESDENSITY_{it}=\varepsilon_{it} \tag{1}$$

Network density is bounded between 0 and 1, therefore to map it to the real line for regression (1) to be well specified, I include the logit transformation of NETDENSITY as a dependent variable. I call the residual from this regression the "residual network density" ($RESDENSITY_i$) and use it as a measure of the speed of information diffusion in the subsequent analysis. This method of adjusting for market capitalization is similar to that used by Hong, Lim and Stein (2000) and Nagel (2005) in different contexts. As expected, the coefficient β_t in equation (1) is negative and highly significant for the majority of the quarterly regressions (95 out of 108 quarters).

C. Centralization and Stock Volatility

Network Centralization measures the extent to which an entire network is focused around a few central funds that act as information sources for the majority of other funds in the network. Ozsoylev (2007) predicts that asset volatility increases as network centralization increases, provided that agents in different networks have a similar number of information sources on average. To test this prediction, I measure network centralization by the following measures, which are widely used in the network theory literature:

1. Coefficient of Variation ratio (CVRATIO): Formally, it is defined as stdev(n_{ij})/mean(n_{ij}) where n_{ij} is the number of direct ties that a given fund j in the stock ownership network of stock i has with other funds in the network of stock i, and mean(n_{ij}) and stdev(n_{ij}) are the mean and standard deviation of n_{ij} over all the funds in the ownership network of stock i. The

coefficient of variation allows for comparison across networks with widely different mean number of connections. For example, in Figure 3 standard deviation and mean of the number of connections of funds in the stock ownership network of Cognex Corp are 1.75 and 2.80 respectively. Therefore, the coefficient of variation is 0.63.

2. Standard deviation of degrees (STDDEGREE): It is defined as the standard deviation of n_{ij} over all the funds in the ownership network of stock i, where n_{ij} is the number of connections that a given fund j in the stock ownership network of stock i has with other funds in the ownership network of stock i. For example, in Figure 3 standard deviation of the number of connections of funds in the stock ownership network of Cognex Corp is 1.75.

3. Aggregate Centralization measure: I measure the aggregate economy-wide centralization at the end of each quarter by calculating the coefficient of variation (AGG_CVRATIO) of total number of connections of all the funds in the economy.

Intuitively, the coefficient of variation is expected to be a better measure for centralization compared to standard deviation as it allows for comparison across networks with widely different values for mean number of information sources. The mean number of information sources varies over time with the variation in number of mutual funds in the economy. In the cross-section, the mean number of information sources increase in stock's market capitalization.

D. Stability of Holdings-Based Networks

If the holdings based definition of networks introduced in this paper accurately captures the informational linkages between investors, we should expect these networks to be relatively stable over time. Panel C of Table I documents the persistence or stability of the holdings based information networks. For the fund's information network, I find that on average 58 % (33%) of the total linkages between funds in the entire mutual fund sample exist one (three) year hence. The benchmark for comparison is a random graph where the linkages between funds are assumed to be formed randomly. To construct a random network for the mutual funds in my sample, I assume that any two mutual funds are likely to be connected with a probability *p* where *p* is determined by the observed percentages of linkages connected in any given year based on the holdings-based definition. Next, for each pair of funds, a number is drawn randomly from a uniform distribution between 0 and 1. A pair of fund is assumed to be connected if the random number drawn is less than *p* and vice versa. The value of *p* varies between 0.04 and 0.33 over the sample period. This gives me a random network for all the funds in any given year. Next, similar to as done for the observed holdings-based networks above, I examine the overlap between these random networks over time. I repeat this network simulation 100 times and take an average of the network stability or persistence measures for the 100 different iterations. As shown in Panel C, I find that on average only 12% of the linkages that exist in any given year t survive one year later for a random graph compared to 54% for the holdings based fund network. This shows that holdings based networks are much more stable than what we would observe by random chance. The stability for a stock's information network is even stronger. For a stock's information network, I find that on average 77 % (58%) of the linkages exist one (three) year hence. These results show that holdings based networks are reasonably stable over time and this degree of persistence is unlikely to be observed due to random chance.

II. Data

The mutual fund holdings data in this study comes from the CDA/spectrum database, which includes stock holdings for all registered US mutual funds filing with the SEC. The data provides holdings of individual funds collected via fund prospectuses and SEC N30D filings at either quarterly or semi-annual frequency. I focus the analysis only on actively-managed US equity funds by including funds with investment objectives of aggressive growth, growth or growth and income in the CDA dataset. Additionally, I manually screen all funds and exclude index funds, as index funds are passive and their trades are mechanical and so might bias the conclusions of this study. I merge the Thomson CDA database and the CRSP mutual fund database using the MFLINKS dataset in WRDS to eliminate from my sample any mutual funds based outside the US (mostly Canadian). Fund location and fund family information is also obtained from the CRSP mutual fund database. Finally, I only include fund portfolios with a minimum of 10 stocks that can be matched to the CRSP stock database and with equity holdings exceeding $1 million.

To study the effect of network structure on stock returns and volatility (in sections IV and V), I need to synchronize the fund holdings. For funds not reporting at the end of the recent quarter, I move their most recent holdings snapshot reported within the last six months to the end of the most recent quarter. If a fund did not report its holdings within the last six months, I exclude it from the sample. For the analysis of correlated trading in section III, I calculate trades of mutual funds based on their quarterly changes in holdings. Therefore, for analyzing correlated trading behavior, I only include funds that file their holdings quarterly. Stock return data is obtained from the monthly CRSP stock data files and accounting data is from COMPUSTAT. I obtain stock recommendations and analyst coverage from the IBES database. The analysis focuses only on common stocks with stock type codes of 10 and 11. The mutual fund data in this paper is from January 1980 to December 2006. The fund family and fund location data is available only from June 2003 to December 2006.

Summary statistics for the sample used in the stock return analysis is reported in Panel A of Table I. The stocks which constitute 5% or more of the portfolio of at least one mutual fund are included in the return analysis. On average, the sample for return estimation includes 19% of the CRSP stocks, but they form a significant 70% of the market cap of the CRSP universe. The percentage of mutual funds connected to at least one other fund is around 70% on average. Therefore, using the holdings-based definition I am able to identify information networks for the vast majority of the mutual fund universe. In my sample, a mutual fund is on average connected to 54 (median) other funds. The median number of funds in the information network of a stock who are likely to receive information about that stock is 62. Finally, the mean value of network density for the stocks in the sample is 0.42, therefore on average 42% of the maximum possible links between funds in the information network of a stock are connected. Panel B reports the correlations between network centralization, network density, network size measures and other stock characteristics. The correlation between network density and coefficient of variation is negative and highly significant at -0.66. Both have been previously used in the social networks literature to measure the efficiency of information flow (Yamaguchi (1994), Buskens and Yamaguchi (1998)). Due to the high correlation between different stock-level network measures of the speed of information flow, in this paper I use network density as the primary measure for the efficiency of information diffusion. The results are robust to employing an alternative measure e.g. the coefficient of variation.

III. Information Networks and Correlated Trading

A. Correlated Trading

Following the methodology in Hong, Kubik and Stein (2005) and Cohen, Frazzini and Malloy (2008), I first test the existence of stock-ownership based information networks by examining whether funds trade together with other funds in their stock-ownership network. If fund managers in the same network communicate with each other or share the same information sources, they are more likely to buy or sell a particular stock in any quarter if other managers in the same network are also buying or selling that same stock. This provides initial evidence in support of the stock-ownership based information networks. I use a Fama-MacBeth (1973) methodology in my analysis of trades. First, I define $h^i_{j,t}$ as the percentage of fund *j*'s portfolio invested in stock *i* in quarter *t*. Next, I calculate $H^i_{N,j,t}$ as the average holding of stock *i* by all the funds in the stock ownership network of fund *j* in quarter *t* except for the fund *j* itself. Similarly, I define $H^i_{R,j,t}$ as the average percentage holding of stock *i* in quarter *t* by all mutual funds in the sample except for fund *j*. For calculating $H^i_{N,j,t}$ and $H^i_{R,j,t}$, I also include funds with a zero holding of stock *i* in quarter *t*. Our sample for examining correlated trading excludes trades of stocks in the bottom NYSE size quintile and by funds with less than six other funds in their stock ownership network.

To examine the effect of stock-ownership networks on the trading behavior of mutual funds, I regress the change in the weight of stock *i* in mutual fund *j*'s portfolio between time t-1 and t ($\Delta h^i_{j,t}$) on the change in the average weight of stock *i* in the portfolios of other mutual funds in fund *j*'s stock ownership network ($\Delta H^i_{N,j,t}$) and on the change in the average weight of stock *i* in the portfolios of all mutual funds except for the fund itself ($\Delta H^i_{R,j,t}$). I run the following regression each quarter from June 1980 to December 2006:

$$\Delta h^i_{j,t} = \alpha + \beta \cdot \Delta H^i_{N,j,t} + \gamma \cdot \Delta H^i_{R,j,t} + \varepsilon^i_{j,t} \qquad (2)$$

The key hypothesis is that the effect of trading decisions of other mutual funds in a fund's stock ownership network on a fund's trades is significant after controlling for the average trading behavior of the entire mutual fund sector. The null hypothesis is that information diffusion effects do not exist, i.e., coefficient β is zero. For information diffusion to exist through the stock ownership ties, we should have $\beta>0$. I calculate the average coefficients and corresponding t-statistics by averaging regression coefficients across all 108 quarters from March 1980 to December 2006.

The results are presented in Table II. As shown in the table, the correlation of the trades of a mutual fund with the trades of other mutual funds in its stock ownership network is positive and highly significant after controlling for the aggregate trades of the mutual fund industry. In all the regression specifications, the coefficient β is positive and highly significant after controlling for the overall trading behavior of the mutual fund sector. As shown in specifications 1 in Table II, a one percentage point increase in the average weight of a stock within a fund's stock ownership network leads to a highly significant 0.22% increase in the portfolio weight of that stock in the fund's portfolio. In terms of economic significance, a one standard deviation increase in the average weight of a stock in a fund's information network leads to a 0.06 standard deviation increase in the portfolio weight of that stock in the fund's portfolio. Similarly, a one standard deviation increase in the average weight of a stock by the entire mutual fund sector leads to only a 0.03 standard deviation increase in the portfolio weight of the stock in the fund's portfolio. Therefore, the economic impact of other funds in a given fund's information network is twice that of other mutual funds in general.

In specification 2, I examine whether the correlated trading behavior examined in equation 2 is stronger if the mutual fund also belongs to the information network of the stock it is trading. The dummy variable IN is equal to 1 if the mutual fund belongs to the network of the given stock and is 0 otherwise. A shown in the second column of Table II, the interaction term between IN and average trades within the stock ownership network is positive and highly significant. This shows that a fund is more likely to follow other funds in its network while trading a stock if it also belongs to the information network of that stock. On the other hand, the coefficient of the interaction term between IN and trades of all other funds in the mutual fund universe is insignificant because as expected the information network definition of a stock should have no impact on the aggregate trading by the mutual funds in general.

In specification 3, I test whether trading by mutual funds is driven by common reaction to a public information signal rather than by transmission of information or attention through a stock ownership network. As in other recent studies (see Kacpercyzk and Seru (2007) for example), I use the most recent analyst stock recommendation as a proxy for the set of available public information in the market about a stock. To control for correlated trading of mutual funds in response to public information, I include both changes and levels of stock recommendations as additional control variables in the regression in equation 2, along with other stock characteristics. A lower value of stock recommendation corresponds to a buy recommendation and a higher value corresponds to sell recommendations. As shown in column 3 in Table II, the coefficient corresponding to average trading by other funds in the network remains almost unchanged from the baseline case and highly significant. Therefore, correlated trading by a mutual fund with other funds in its stock ownership network cannot be explained by common reaction to public information. The coefficients corresponding to mean recommendation and changes in mean recommendation are negative as expected but are insignificant. Therefore, correlated trading by mutual funds cannot be explained by changes in public information as captured by mean analyst recommendations.

In regression specifications 4 and 5, I divide the sample into two groups to examine the variations on the intensive and extensive margins separately. In specification 4, I examine the intensive margin i.e. include only the observations where the fund already owns the stock. This specification examines the effect of the trading decisions of other funds in the network when the fund already owns the stock. In this much smaller subsample, the coefficient corresponding to other funds in the network is higher at 0.61 and is highly significant. The coefficients are higher as a fund that already owns a stock is more likely to collect and analyze information about it than one that does not.

In specification 5, to examine the extensive margin I use a probit regression instead of OLS. The left-hand variable takes a value of 1 if in quarter *t*, a fund initiates a position in a stock *i* that it did not own in quarter *t-1* and a value of 0 otherwise. I exclude all the observations where fund *j* already owned the stock *i* in quarter *t-1* ($h_{i,j,t-1} > 0$). This regression tests the effect of a fund's stock ownership network on its decision to initiate a new stock position. After controlling for the average trading by all mutual funds in the sample, the coefficient corresponding to average trading by funds in the stock-ownership network of a given fund is positive and highly significant. The baseline probability of a fund buying any stock it doesn't own is 0.47% for mean values of within network and aggregate mutual fund sector trades. The probability increases to 0.55% for a two standard deviation increase in the within network average holding. Therefore the increase in probability is around 17% from the baseline case, which is an economically meaningful increase. In specification 6, I show that the results in Table II are robust to the use of an alternative estimation methodology. Instead of using the Fama-MacBeth method, I estimate a pooled panel regression and cluster the standard errors by stock and quarter. The coefficient on average stock ownership network trades is 0.27, which is similar to the corresponding Fama-MacBeth coefficient in the first column. The significance of the coefficient measured by the t-statistic is also similar to that for the Fama-MacBeth regression.

B. Style Investing and Correlated Trading

The most plausible alternative explanation for the results in section III.A is that the correlated trading by fund managers in the same stock ownership network is driven by herding in trades of managers following similar investment styles. Although it is not possible to completely rule out a style-investing explanation because no list of investment style controls can be exhaustive, I provide evidence that the correlated trading by a fund with other funds in its stock ownership network cannot be explained entirely by known definitions of style-investing.

First, in specifications 7 and 8 of Table II, I estimate the same regression as the baseline case in column 1, but include additional controls for style-based trading and a given fund's preference for stocks from a particular industry. In calculating the style controls, I use the same methodology as Ivkovich and Weisbenner (2007). Each of the fund's stock trades is characterized by the stock's size, book-to-market ratio, past eleven month return (t-11 to t-1) and industry group. Each quarter, stocks are divided into quintiles of book-to-market ratio, past eleven month return and market cap based upon NYSE cutoffs. Stocks are also assigned to one of the 12 industry groups based on their SIC code[4]. Each stock buy or sell is associated with 27 dummy variables. The first five dummy variables are set to 1 if the stock belongs to a particular

[4] The definition of 12 industry groups based upon SIC codes is obtained from Kenneth French's website.

size quintile and 0 otherwise. Similarly, the next ten dummy variables are defined for book to market ratio and past returns. The last 12 dummy variables are set to 1 if the stock belongs to a particular industry group and 0 otherwise. The 27 controls are computed separately for buys and sells as weighted sums of the respective 27 dummy variables across all the buys or sells made by the mutual fund in the quarter and thus represent the share of quarterly buys or sells in each size, book-to-market, momentum quintile and industry group. The control variables for sells are assigned a negative sign to allow for estimation on a combined sample of buys and sells.

In specification 7 of Table II, I include 15 style controls and in specification 8, I include all 27 style and industry control variables. The coefficient corresponding to trades by funds in the stock ownership network remains highly significant after controlling for style-based trading. This shows that correlated trading with funds in the stock ownership network remains robust after including stock characteristics based style controls. A number of style control variables also come out significant in the regressions.

I also control for style investing using another methodology. For each fund, I sort the funds in its information network into funds following the same investment style and funds following a different investment style from the fund itself. Investment style is defined using two methods. In the first method, each quarter I calculate the aggregate quintile of book-to-market and size for every fund by weighting the book to market and size quintile of the stocks in its portfolio by the percentage holding of the stocks it owns. Then, each quarter, I sort funds into three independent groups along each of the dimensions of aggregate book-to-market and size of their portfolio. This divides funds into nine groups based on their aggregate size and book to market quintile score. Funds in the same size and book-to market group are assumed to be following the same investment style.

In the second method, funds with the same benchmark index are assumed to be following the same investment style. Most of the funds closely follow the investment style of their passive benchmark indexes as their performance is evaluated relative to the performance of their benchmark index. Therefore, the benchmark index of a fund should give a more accurate description of its investment style compared to a stock-characteristics based approach. As benchmarks, I include almost all the indexes used by mutual funds during my sample period. This gives a total of 19 indexes from three index families: S&P/Barra, Russell, and Wilshire. Then for each fund, following the methodology in Cremers and Petajisto (2009), the index with the lowest deviation of holdings from the fund's portfolio is assigned as its benchmark.

The results are reported in Table III. I calculate the correlation of a fund's trades with the average trades of funds in its network following the same investment style and with the average trades of funds in its network that follow a different investment style. Correlated trading within a stock ownership network is robust to an alternative style-investing based explanation if the coefficient corresponding to average trades by funds in the stock ownership network and following a different investment style remains significant after controlling for average trades of the remaining funds in the stock ownership network and the average trading of the entire mutual fund sector.

As shown in column 3 of Table III, a one percentage point increase in the average weight of a stock by the funds in the same stock ownership network and with the same benchmark index and by the funds in the network with different benchmarks leads to an increase in the weight of the stock in the fund's portfolio by around 0.06% and 0.09% respectively. Similarly, for the alternative definition of style investing based on stock-characteristics (specification 6), the corresponding coefficients for average trading by funds in the network and following same style and following different styles are 0.05% and 0.09% respectively. The strong effect of the average trades of the same-style funds on trading behavior of mutual funds is expected given the prior evidence on style investing by mutual funds (e.g., Chan, Chen and Lakonishok (2002)). But the correlated trading with funds with different investment styles remains positive and highly significant after controlling for the average trading of funds with similar investment style. The hypothesis of information diffusion through the stock ownership ties therefore remains robust after controlling for style investing.

C. Geographic Location and Stock Ownership Network

Prior research (Hong, Kubik and Stein (2005)) has shown that the trades and holdings of mutual fund managers are correlated with the trades and holdings of other fund managers in the same city. If the information diffusion between mutual fund managers is constrained within the same geographic location, then it is possible that the network effect due to funds sharing large position in same stocks may also be limited to funds in the same geographic location. On the other hand, if stock ownership network is a better representation of a fund's information network compared to its location, we should expect the effect of geographic location to be much weaker or to be subsumed within the effect of stock ownership network.

I examine the joint effect of geographic location and stock ownership network on the trading behavior of mutual fund managers. The sample for this analysis is from the second quarter of 2003 to the end of 2006: the availability of location and fund family data is limited to this time period. Individual fund location zip codes are obtained from CRSP mutual fund database. Using latitude and longitude data from *U.S. Census Bureau's Gazetteer Place and Zip code database*, I match the fund zip codes with the location's latitude and longitude. Next, the distance between two funds is calculated using the methodology in Coval and Moskowitz (1999). The distance between two funds i and j is given by:

$$d_{i,j} = \arccos\left\{\begin{array}{l}\cos(lat_i)\cos(lon_i)\cos(lat_j)\cos(lon_j) \\ +\cos(lat_i)\sin(lon_i)\cos(lat_j)\sin(lon_j)+\sin(lat_i)\sin(lat_j)\end{array}\right\}\frac{2\pi r}{360}, \tag{3}$$

where *lat* and *lon* are the latitudes and longitudes of the fund location (measured in degrees) and r is the radius of the earth (~3963 miles). Using this methodology, I calculate the distance of each fund from every other fund in the sample. All the funds within a 100 miles radius of a given fund's location are defined as belonging to the same geographic location as that fund. To avoid any bias, I exclude funds that belong to the same fund family as that fund from its stock ownership and location networks because trades within the same fund family may be correlated due to internal information transfer unrelated to location or direct information linkages.

The results are presented in Table IV. In column 1, I include average trades within ownership and location networks in the same regression to test the relative importance of the two networks. I find that the effect of the ownership network (t-stat of 18.69) is much stronger compared to the effect of the location network (t-stat of 3.21). In terms of economic significance, a one standard deviation increase in the average weight of a stock in a fund's stock-ownership network leads to 0.06 standard deviation increase in the portfolio weight of that stock in the fund's portfolio. Similarly, a one standard deviation increase in the average weight of a stock in the fund's location network leads to only a 0.01 standard deviation increase in the portfolio weight of the stock in the fund's portfolio. Therefore, the economic impact of funds in the stock ownership network of a mutual fund is six times that of the funds in its location network.

In regression specifications 2 and 3, I divide the sample into two groups to examine the variations on the intensive and extensive margins separately. In specification 2, I examine the intensive margin i.e. include only the observations where the fund already owns the stock. Similar to the results for the entire sample in column 1, the effect of the ownership network is much stronger compared to the effect of the location network also for the trades where the fund already owns the stock. In specification 3, to examine the extensive margin I use a probit regression instead of OLS. The left-hand variable takes a value of 1 if in quarter *t*, a fund initiates a position in a stock *i* that it did not own in quarter *t-1* and a value of 0 otherwise. I exclude all the observations where fund *j* already owned the stock *i* in quarter *t-1* ($h_{i,j,t-1} > 0$). This regression tests the effect of a fund's stock ownership and location network on its decision to initiate a new stock position. After controlling for the average trading by all mutual funds in the sample, the coefficient corresponding to average trading by funds in both the stock-ownership network and in the location network of a given fund are positive and highly significant but the coefficient corresponding to average trading within the stock ownership network has a stronger economic and statistical significance. For a two standard deviation increase in ownership network average holdings and location network average holdings, the probability of a fund buying any stock it doesn't own increases by 12% and 9% respectively from the baseline case of mean values for all the independent variables.

In column 4, I divide the funds in a given fund's ownership network into two groups: funds that also belong to the same location (OWN_LOCAL) and funds that don't belong to the same location (OWN_NONLOCAL) as the fund itself. The results show that the effect of the ownership network is very strong both inside and outside the location network.

Similarly, in regression specification 5, I divide the funds in a given fund's location network into two groups: funds that also belong to the same stock ownership network (LOC_OWNERSHIP) and funds that don't belong to the same stock ownership network (LOC_NONOWNERSHIP) as the fund. The results show that the effect of location network on the trading behavior of mutual funds is subsumed by the effect of the stock ownership network. The coefficient corresponding to the funds that belong to both ownership and location network is 0.04% and is highly significant (t-stat=9.81) whereas the coefficient corresponding to funds that belong to location but not to ownership network is statistically insignificant. In other words, the positive correlation between trades of mutual funds with other funds in the same location is limited to the funds with which they also share a large position in at least one stock. This result also provides evidence against the alternative style-investing based explanation. As shown in column 5, the effect of location is subsumed by the effect of stock ownership network. This empirical observation combined with the result in the previous literature that correlated trading within the same geographic location is due to information diffusion shows that at least part of the correlated trading within the stock ownership based information network is due to information diffusion. In specifications 6 and 7, I report the results corresponding to the intensive and extensive margin. The results are similar to the results for the entire sample. The effect of location network is subsumed by the effect of ownership network in trades where the fund already owns the stock and in the trades where the fund has to make a decision on initiating a new stock position.

IV. Gradual Information Diffusion and the Cross-Section of Stock Returns

A. Stock Return Momentum

Having provided evidence that the flow of private information can be captured using a stock ownership network, I now use my network methodology to test empirical predictions from theoretical models of information diffusion. First, I test the main cross-sectional prediction in Hong and Stein (1999), namely that the momentum effect should be stronger for stocks with slower speed of information diffusion among investors, after controlling for firm size and other factors. I use residual network density (the residual after adjusting network density for market cap) as a measure of the speed or efficiency of information diffusion. Network density is defined as the number of ties in a network divided by the maximum number of possible ties. The construction of residual network density measure is discussed in detail in Section I.B.

In Table V, I present the results from portfolio strategies based on past returns and residual density of the stock ownership network. I form momentum portfolios using a six-month ranking period and six-month and twelve-month holding periods. Specifically, at the end of each quarter I form portfolios of stocks ranked on the basis of their returns over the previous six months, and compute the average returns to such portfolios over the next six or twelve months. I leave a gap of one month between the ranking and holding period to account for any microstructure issues. As shown in Panel A of Table V, the monthly equal-weighted long-short raw return for an unconditional momentum strategy is 0.85% for a holding period of six months, consistent with the return on momentum strategy for large cap stocks (Jegadeesh and Titman (2001)).

To examine the effect of network density on momentum, at the beginning of each quarter I first sort stocks into quintiles based on past 6 month returns and then independently sort the stocks into three groups based on lagged residual network density measured at the beginning of the previous quarter. Panel A and Panel B of Table V presents the returns for each of the 15 portfolios measured over the holding period of next 6 months. An equal-weighted long-short momentum strategy with a holding period of 6 months earns a monthly raw return of 1.01% and 3-factor alpha of 1.17% for the bottom residual network density group and a raw return of 0.63% and alpha of 0.83% for the top network density groups. The difference in momentum returns between the top and bottom residual network density groups is negative and significant. The results are similar for a trading strategy with a 12 months holding period.

In Table VI, I present regression evidence on the effect of network density on future stock momentum returns. I use the Fama and MacBeth (1973) methodology and estimate predictive cross-sectional regressions of next 6 month or 12 month returns on past network density, past returns and other stock characteristics. The results are consistent with the portfolio results. The coefficient on the interaction term between momentum and logarithm of network density is negative and highly significant for all the return regression specifications (columns 1 to 3) after controlling for other stock characteristics. This confirms that momentum increases with decreasing network density. The insignificant coefficients on the interaction term between momentum and analyst coverage and on the interaction term between momentum and turnover show that the effect of residual network density on momentum is much stronger in my sample compared to the previously documented effects of turnover ((Lee and Swaminathan (1999)) and analyst coverage (Hong, Lim and Stein (2000)). I also include the interaction between logarithm of market capitalization and past returns in the regressions to control for the mechanical relationship of analyst coverage and network density with stock's market capitalization.

Following Hong, Lim and Stein (2000), I also follow an alternative regression approach in specifications 4 and 5 and include the serial correlation in stock returns measured over next five years as a dependent variable. The more positive the serial correlation in stock returns, stronger is the momentum effect. For each stock *i* in the sample at the beginning of a quarter, I estimate the serial correlation of its six-month excess returns (relative to T-bills), using 49 overlapping observations over the next 5 year or 20 quarter period from t to t + 20**,** and call this variable RHO_{it}. Next, I estimate cross-sections regressions with $RHO_{i,t}$ as a dependent variable and network density, analyst coverage, market capitalization and other stock characteristics as explanatory variables. All the right-hand-side variables are lagged by one quarter and are measured at the beginning of quarter t-1. We find that coefficient corresponding to network density is negative and highly significant both in columns 4 and 5. In column 5, one standard deviation decrease in network density leads to an increase in RHO_{it} by 0.02 standard deviations. This confirms that stock return momentum increases with decreasing network density or with decreasing speed of information diffusion. The coefficients corresponding to institutional ownership and analyst coverage are also negative and highly significant which confirms the previous finding that serial correlation in returns is stronger for stocks with low institutional ownership and analyst coverage.

B. Price Delay

Using network density as a measure for the speed of information flow, I next examine the effect of gradual information diffusion on the delayed reaction to market-wide information. Following Hou and Moskowitz (2005) and Hou (2007), I measure market-wide information by past market returns or average past returns of the stocks in the same industry group. If gradual information diffusion among investors causes price delays, we should expect this price delay or lead-lag effects in returns to increase with decreasing network density after controlling for other stock characteristics known to affect price delay.

To measure price delay, I employ measures used in the previous literature (see Hou and Moskowitz (2005)). The market return or the industry return is employed as the news to which stocks respond. At the end of December of each calendar year, I run a regression of each stock's weekly return on contemporaneous and four weeks of lagged returns of either the market portfolio or the equally weighted industry portfolio over the prior one year. The industry portfolio of a stock is defined as all other stocks in CRSP with the same Fama and French 12-industry classification as the given stock at the end of that year.

$$r_{i,t} = \alpha_i + \beta_i R_{m,t} + \sum_{n=1}^{4} \delta_i^{(-n)} R_{m,t-n} + \varepsilon_{i,t} \qquad (4)$$

Where $r_{i,t}$ is the return on stock i and $R_{m,t}$ is the return on the CRSP value-weighted market index or the average return of the stocks in the same industry portfolio as stock i. If the stock responds immediately to market news, then β_i will be significantly different from zero but none of $\delta_i^{(-n)}$ will differ from zero. If however stock i's price responds with a lag, then some of the $\delta_i^{(-n)}$ will differ significantly from zero. The measure of delay that I use is the fraction of variation of contemporaneous individual stock returns explained by lagged market or industry returns. The delay measure is thus defined as one minus the ratio of the R^2 from regression (4) restricting $\delta_i^{(-n)}=0$, for all n ε [1,4], over the R^2 from regression (4) with no restrictions.

$$Delay = 1 - \frac{R^2_{\delta_i^{(-n)}=0, \forall n \in [1,4]}}{R^2} \qquad (5)$$

To examine the effect of network density, I follow the Fama-MacBeth (1973) methodology and each year I regress the price delay measure on lagged network density and other stock characteristics known to affect price delay. The delay measure is bounded between 0 and 1. Therefore for the regressions to be well specified I include the logit transformation of delay measure as a dependent variable. The independent variables are lagged and include network density and the following stock-level variables known from previous literature to affect price delay: market capitalization, Amihud Illiquidity measure, Average daily turnover, Analyst Coverage and Price. The results are presented in Table VII. The coefficient corresponding to network density is negative and highly significant in all the regressions specifications. This shows that price delay increases with decreasing network density or decreasing speed of information diffusion. For example, one standard deviation decrease in network density leads to 0.02 standard deviations increase in price delay. The results also provide evidence in favor of an investor attention based explanation for price delay as opposed to a liquidity based explanation as shown by negative and significant coefficients on proxies for attention (Analyst Coverage, Institutional Ownership and past 12-month returns) and insignificant coefficients corresponding to all the liquidity variables (Price, Turnover and Amihud Illiquidity measure).

V. Information Networks and Stock Volatility

Campbell, Lettau, Malkiel and Xu (CLMX, 2001) document that the average idiosyncratic volatility of individual stocks has increased steadily over the period from 1962 to 1997, whereas the aggregate market volatility has remained almost constant over the same period. This steady increase in idiosyncratic risk has been a puzzling fact for researchers. Recent papers have followed different approaches to try to explain this time trend. Proposed explanations include increase in institutional ownership (Xu and Malkiel (2002); Bennett, Sias and Starks (2003)), product markets becoming more competitive (Irvine and Pontiff (2008)), speculative retail trading (Brandt, Brav, Graham and Kumar (2008)), increased propensity of firms to issue public equity at an earlier age (Fink, Fink, Grullon and Weston (2005)) and firm fundamentals becoming more volatile (Wei and Zhang (2006)).

Brandt et al (2008) show that the trend in average idiosyncratic risk documented by CLMX reversed itself by 2007. Therefore any plausible explanation for the variation in average firm-level idiosyncratic risk in economy over time should be able to explain these episodes of earlier increase and recent decline in idiosyncratic risk. Based on the structure of individual stock and economy-wide information networks, I test a new explanation for the cross-sectional differences in firm-level idiosyncratic risk and the time-series variation in average stock idiosyncratic risk in the economy.

A number of recent theoretical papers study the impact of information networks on asset volatility. Cao and Xia (2006) generalize Kyle (1985)'s noisy rational expectations model by incorporating direct information transmission between agents. They study properties of prices in four different network structures and show that trading volume and price volatility are higher in all the models with information linkages compared to the benchmark Kyle model. Ozsoylev (2007) proposes a static rational expectations model in which agents learn from each other through a social network. In this paper, I test a specific prediction of the model that relates asset volatility to network centralization. Given two different networks N_1 and N_2, the model predicts that the information driven volatility component is greater in information network N_1 than that in N_2 if on average agents in N_1 and N_2 have the same number of information sources and network N_1 is more centralized than network N_2. The intuition behind this prediction is as follows. In a network with high degree of centralization, changes in price are significantly dependent on changes in central nodes' signals. Whereas in a network with lower centralization, changes in prices are relatively equally dependent on changes in each mutual fund's private signal, and given the independence between the error terms of their private signals most of these changes wash out each other. Therefore, the information-driven volatility component increases with increase in network centralization. The coefficient of variation (the ratio of the standard deviation of number of connection to mean number of connections) accurately captures the network measure implied in this proposition. It measures the centralization of networks after controlling for the average number of information sources. This allows for comparison across networks with different average numbers of connections.

For individual stocks, I measure the centralization of their information network as the coefficient of variation of the number of connections of the mutual funds present in the ownership network of that stock at the end of each quarter. The definition of stock ownership network is provided in the network methodology section (section II). Similarly, to estimate the economy-wide centralization measure, I calculate the coefficient of variation of the number of connections of all the connected mutual funds in the economy at any given point of time. I use these network measures to forecast the average stock volatility in the next quarter.

Similar to Xu and Malkiel (2002), I calculate stock idiosyncratic volatility as the standard deviation of residuals from a Fama French 3-Factor model. At the end of each quarter, I run the following three-factor regression separately for every stock using daily return data from that quarter:

$$R_{i,t} = \alpha_i + \beta_i^{MKT} R_t^{MKT} + \beta_i^{SMB} R_t^{SMB} + \beta_i^{HML} R_t^{HML} + \varepsilon_{i,t} \qquad (6)$$

I use a quarterly estimation window as the mutual fund holdings-based independent variables used to explain stock volatility are calculated at a quarterly frequency. Idiosyncratic volatility for a given stock is calculated as the standard deviation of residuals from regression (6) for that stock. To estimate the cross-sectional effect of network structure on stock volatility, I regress stock idiosyncratic volatility on the network centralization measure and on other control variables from the end of previous quarter. I pool the firm-time observations together and estimate the following panel regression:

$$\log(\sigma i_{i,t}^{IDIO}) = \alpha + \beta \log(CENTRALIZATION_{i,t-1}) + \delta_j Control\_Variable_{j,t-1} + \varepsilon_{i,t} \quad (7)$$

Following Petersen (2009), I cluster standard errors along time and firm dimensions to control for correlation among observations over time and across firms. I use control variables which have been shown in the literature to predict future idiosyncratic risk: market cap, turnover, book to market ratio, institutional ownership, lagged idiosyncratic volatility, past 12 month return and stock price.

The results are presented in Table VIII. All the variables including the dependent variable are standardized to allow for comparison between coefficients. The results strongly confirm the hypothesis that stock volatility increases with network centralization. In the basic specification in column 1, I control only for market capitalization of the stock. A one standard deviation increase in the coefficient of variation (CVRATIO) leads to a 0.051 standard deviation increase in next quarter's idiosyncratic risk. The coefficient corresponding to the network centralization measure is significant and economically meaningful in all the regression specifications. The economic significance of the centralization coefficient is of the same order as for the other coefficients except for the lagged idiosyncratic risk which is a stronger predictor. To ensure that the results are not driven by a small part of the sample, for instance by a few firms from the dot-com bubble period, I divide the sample into two parts and estimate the regression in equation (7) separately for 1980 to 1995 and for 1996 to 2006. As shown in specifications 5 and 6, network centralization is a strong predictor of idiosyncratic volatility in both subsamples. In the appendix, I show that these cross-sectional volatility results and the stock momentum results (reported in section IV) are robust to an alternative network definition based on large active positions in the same stocks.

Next, I test whether network centralization can explain the variation in aggregate economy wide idiosyncratic risk. Theory predicts that aggregate volatility should increase in network centralization. Initial evidence in favor of the hypothesis is presented in Figure 4. As shown in the figure, the network centralization measures are positively correlated with next period average idiosyncratic risk except during the crash of 1987. For example, the correlation between the aggregate coefficient of variation and the next period average idiosyncratic risk is 0.33. To provide formal evidence in support of this hypothesis, I estimate the following time series regression:

$$\sigma_t^{AGGREGATE} = \alpha + \beta t + \delta AGG\_CVRATIO_{t-1} + \gamma_j ControlVariable_{j,t-1} + \varepsilon_t \qquad (8)$$

The regression also includes a time trend variable. The standard errors are corrected for autocorrelation using a Newey-West correction with four lags. The results are presented in Panel A of Table IX. The first column gives the results for the basic specification which includes a time trend variable and the aggregate coefficient of variation as independent variables. The dependent variable is next quarter's average idiosyncratic risk of connected stocks in the sample. The time trend coefficient is insignificant and the coefficient corresponding to the aggregate network centralization measure is positive and highly significant. This is consistent with the hypothesis that volatility increases with network centralization. The results remain robust after including average institutional ownership (IO) of the stocks in the sample as a control variable (column 2). In specification 3, I also include lagged average idiosyncratic risk as an independent variable. The significance of the aggregate coefficient of variation drops as the average idiosyncratic risk series is highly autocorrelated, but it is still significant at the 90% level (t-stat of 1.70). The time trend variable is excluded from specifications 2 and 3 as it is highly correlated with the average institutional ownership. In specifications 4 and 5, I use the S&P500 index volatility as another measure for aggregate volatility. The results are similar. The R-squared for the basic specification (column 1), including only network centralization and a time trend as independent variables is 11.5% and increases to 76% after including lagged average idiosyncratic volatility. Therefore network centralization explains a substantial proportion of the variation in aggregate idiosyncratic risk over time.

An alternative explanation for the results in Table IX, Panel A is that aggregate volatility may be causing networks of investors to become more centralized rather than the other way around. This can happen if during volatile periods, investor herd more and thus hold positions similar to a group of central or leader mutual funds causing centralization to increase. I investigate the direction of causality in the Granger Causality sense between my aggregate centralization measure (AGG_CVRATIO) and average idiosyncratic risk (IDIORISK). I report the results in Panel B of Table IX. The first column shows that lagged average idiosyncratic risk has no predictive power for the aggregate centralization measure (AGG_CVRATIO). The null hypothesis that idiosyncratic risk does not Granger cause the aggregate coefficient of variation cannot be rejected (Wald chi Square statistic of 0.61 and p-value of 0.44). Therefore, idiosyncratic risk does not Granger cause network centralization. The result for opposite direction is presented in second column. The null hypothesis that aggregate network centralization does not Granger cause idiosyncratic volatility can be rejected at 10% level (Wald statistics of 3.41 and p-value of 0.07). Therefore, the evidence is consistent with the hypothesis that the centralization of aggregate network of mutual funds explains the variation in idiosyncratic risk over time.

VI. Conclusion

I introduce a new methodology to identify information linkages between mutual fund managers. I define information linkages or stock ownership ties between mutual funds by common large positions in the same stocks. The existence of information flow through these stock ownership ties is confirmed by the strong correlation of the trades of fund managers with other managers in their stock ownership networks after controlling for the overall trading behavior of the mutual fund sector. The effect is robust and cannot be explained by style investing or geographic location.

Having established that the holding-based networks I define can capture the flow of private information, I then use these networks to test some empirical predictions from recent theoretical models of information diffusion. Next, I examine the effect of information diffusion through stock ownership ties on stock returns. Using network density as a measure for the efficiency of information flow, I find that stocks with lower network density demonstrate stronger momentum at short to medium horizons. The evidence is consistent with the gradual information diffusion model of Hong and Stein (1999). I also find that stocks with lower network density show a delayed response to the market-wide information measured by lagged market return or lagged industry returns. This result provides further support to the hypothesis that information diffuses gradually through the holdings-based networks.

Finally, I examine the implications of the structure of stock-ownership based information networks on stock price volatility. I find that centralized information networks lead to higher volatility of individual stocks in cross-section and also explain the variation in average stock idiosyncratic volatility over time. This provides evidence supporting the predictions of recent theoretical models (for example, Ozsolylev (2007)) that study the effect of information networks on stock prices.

References

Bennett, J. A., R. W. Sias, and L. T. Starks, 2003, "Greener Pastures and the Impact of Dynamic Institutional Preferences," *Review of Financial Studies* 16, 1203–1238.

Brandt, M. W., A. Brav, J. Graham, and A. Kumar, 2008, "The Idiosyncratic Volatility Puzzle: Time Trend or Speculative Episodes?" working paper, Duke University.

Bushee, B., and T. Goodman, 2007, "Which Institutional Investors Trade Based on Private Information about Earnings and Returns?" *Journal of Accounting Research* 45, 289-321.

Buskens, V. and K. Yamaguchi, 1999, "A New Model for Information Diffusion in Heterogeneous Social Networks," *Sociological Methodology* 29, 281-325.

Campbell, J. Y., M. Lettau, B. G. Malkiel, and Y. Xu, 2001, "Have Individual Stocks Become More Volatile? An Empirical Exploration of Idiosyncratic Risk," *Journal of Finance* 56, 1–43.

Cao, R., C. Xia, 2006, "Information Networks in Financial Markets," working paper, University of Minnesota.

Chan, L.K.C., H.-L. Chen, J. Lakonishok, 2002, "On mutual fund investment styles," *Review of Financial Studies* 15 1407–1437.

Cohen, L.H., A. Frazzini and C. J. Malloy, 2007, "The Small World of Investing: Board Connections and Mutual Fund Returns," *Journal of Political Economy*, Forthcoming.

Cohen, R., C. Polk and B. Silli, 2010, "Best Ideas," working paper, London School of Economics.

Colla, P., and A. Mele, 2010, "Information Linkages and Correlated Trading," *Review of Financial Studies* 23, 203-246.

Coval, J. D., and T. J. Moskowitz, 1999, "Home bias at home: Local equity preference in domestic portfolios," *Journal of Finance* 54, 145-166.

Cremers, M., and A. Petajisto, 2009, "How Active Is Your Fund Manager? A New Measure that Predicts Performance," *Review of Financial Studies* 22, 3329-3365.

Fama, E., and J. MacBeth, 1973, "Risk, return and equilibrium: Empirical tests," *Journal of Political Economy*, 607–636.

Fink, J., K. W. Fink, G. Grullon, and J. P. Weston, 2005, "IPO Vintage and the Rise of Idiosyncratic Risk," working paper (February), Rice University.

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

Hong, H., T. Lim, and J. C. Stein, 2000, "Bad news travels slowly: Size, analyst coverage, and the profitability of momentum strategies," *Journal of Finance* 55, 265– 295.

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

Hou, K.W., 2006, Industry information diffusion and the lead-lag effect in stock returns, *Review of Financial Studies* 20, 1113–1138.

Hou, K.W., Moskowitz, T., 2005. Market frictions, price delay, and the cross-section of expected returns. *Review of Financial Studies* 18, 981–1020.

Irvine, P. J., and J. Pontiff, 2008, "Idiosyncratic Return Volatility, Cash Flows, and Product Market Competition," *Review of Financial Studies*, Forthcoming.

Ivkovich, Z. and S. Weisbenner, 2007, "Information Diffusion Effects in Individual Investors' Common Stock Purchases: Covet Thy Neighbors' Investment Choices," *Review of Financial Studies* 20, 1327-1357.

Jegadeesh, N., and S. Titman, 1993, "Returns to buying winners and selling losers: implications for stock market efficiency," *Journal of Finance* 48, 65-91.

Jegadeesh, N., and S. Titman, 2001, "Profitability of Momentum Strategies: An Evaluation of Alternative Explanations," *Journal of Finance* 56, 699-720.

Kacperczyk, M., and A. Seru, 2007, "Fund manager use of public information: New evidence on managerial skills," *Journal of Finance* 62, 485-528.

Kyle, A., 1985, "Continuous Auctions and Insider Trading," *Econometrica* 53, 1315-1335.

Nagel, S., 2005, "Short sales, institutional investors and the cross-section of stock returns," *Journal of Financial Economics* 78, 277-309.

Ozsoylev, H., 2007, "Asset Pricing Implications of Social Networks," working paper, Said Business School, University of Oxford.

Ozsoylev, H., and J. Walden, 2010 "Asset Pricing in Large Information Networks," working paper, Said Business School, University of Oxford.

Petersen, M., 2009, “Estimating Standard Errors in Finance Panel Data Sets: Comparing Approaches,” *Review of Financial Studies* 22, 435-480.

Shiller, R. and J. Pound, 1986, "Survey Evidence of Diffusion of Interest Among Institutional Investors," NBER Working Paper 1851.

Wei, S. X., and C. Zhang, 2006, “Why Did Individual Stocks Become More Volatile?,” *Journal of Business*, 79, 259–292.

Xu, Y., and B. G. Malkiel, 2003, “Investigating the Behavior of Idiosyncratic Volatility,” *Journal of Business*, 76, 613–644.

Yamaguchi, K., 1994, “The Flow of Information through Social Networks: Diagonal-Free Measures of Inefficiency and the Structural Determinants of Inefficiency," *Social Networks* 16, 57-86.

Appendix

Robustness Checks: Alternative Network Specification

The evidence in this appendix shows that the results presented in the paper are robust to an alternative definition of information networks. In the alternative network specification, I hypothesize that funds with large active positions (measured by deviation from their benchmark index) in the same stock are likely to be informationally connected to each other[5]. First, I calculate the deviation of stock holdings of each fund from the holdings of its benchmark index by calculating the absolute value of the difference between the weight of a stock in a fund's portfolio and its corresponding weight in that fund's benchmark index. Each quarter, I then divide the holdings pooled across all the funds and stocks into 20 equal groups based on the absolute value of the deviation from the benchmark.

In this network specification, two funds are defined as connected to each other if their deviation from the benchmark index for the same stock is in the top 5 percentile of the pooled deviations across all funds and stocks in the given quarter. Using the deviation from a benchmark instead of large portfolio holdings is expected to be a more accurate measure for an information network as it eliminates the effect of mutual funds' passive indexing strategies. On the other hand, it has a disadvantage that the holdings data is not readily available for all the benchmark indexes. As expected, the alternative benchmark-based definition is highly correlated to the raw holdings based definition used in the paper as the benchmark indexes are highly diversified and therefore the active portfolio weights are very similar to raw portfolio weights for large holdings.

[5] As benchmarks, I include almost all the indexes used by mutual funds during my sample period. This gives a total of 19 indexes from three index families: S&P/Barra, Russell, and Wilshire. Following the methodology in Cremers and Petajisto (2008), the index with the lowest deviation of holdings from the fund's portfolio is assigned as its benchmark.

Next, I test the robustness of momentum, and volatility results (presented in sections IV and V) to the alternative network specification based on the deviation of a fund's holdings from its benchmark index. The results are reported in Panels A and B of Table X. In Panel A, I present the results for a portfolio strategy based on lagged residual network density and past 6 month returns. The difference in momentum returns between the top and bottom residual network density groups is negative and significant for the holding period of next 6 months or 12 months. In columns 1 and 2 of Panel B, I present regression evidence on the effect of network density on future stock momentum returns. I use the Fama and MacBeth (1973) methodology and estimate predictive cross-sectional regressions of next 6 month or 12 month returns on past network density, past returns and other stock characteristics. The regression coefficient for the interaction term between the past six-month returns and log of network density (measured for the new network definition) is negative and highly significant. These results are consistent with the portfolio results in Panel A and the results in section V of the paper and confirm that the returns from a momentum strategy are decreasing in network density.

In specifications 3 and 4, the dependent variable is the log of the next quarter idiosyncratic risk. The coefficient corresponding to the measure of network centralization (CVRATIO) for the new network specification is positive and highly significant. This confirms that stock volatility increases in network centralization. These results show that the empirical findings in this paper are robust to the alternative network definition based on active deviations from the benchmark.

Table I

Summary Statistics

Panel A reports the summary statistics for the sample of stocks and funds used in this paper. First part of the panel presents the time series averages of the number and percentage coverage of stocks and funds in my sample. The bottom part of the panel presents the statistics for the pooled time-stock or time-fund data in this paper. The sample period is from 1980 to 2006. Panel B presents the rank correlations of the following stock-level network variables: network density (NETDENSITY), residual network density (RESDENSITY), coefficient of variation (CVRATIO) and network size (NSTOCKCONN) with other stock characteristics. The definitions of these network variables are discussed in detail in section I of the text. Correlations significant at the 1% level are denoted in bold. Panel C reports the results examining the persistence or stability of Fund and Stock Information Networks over time.

Panel A

	Mean	Stdev	Median	25th Percentile	75th Percentile
	Time-Series (108 quarterly observations)				
Number of Stocks	692	181	651	544	836
% of CRSP Stocks	18.5	2.8	18.6	16.6	20.1
% of CRSP MCap	69.6	9.5	70.1	59.3	78.7
Number of Funds	616	316	549	312	880
% of Funds	70.2	11.7	71.8	59.6	77.4
	Pooled Stock-Time/Fund-Time Observations				
Avg Daily Turnover (%)	0.59	0.71	0.37	0.21	0.69
Mom12 (%)	29.0	67.4	17.9	-2.5	43.9
Bmratio	0.5	1.4	0.4	0.2	0.6
Market Cap ($ Million)	5894.4	19611.9	1349.8	490.1	4039.8
Idiorisk (%)	2.03	1.13	1.76	1.29	2.46
InstOwn (%)	56.1	20.5	57.3	41.8	71.2
Network Density	0.44	0.18	0.41	0.31	0.55
Residual Density	0.00	0.76	-0.12	-0.48	0.36
CV ratio	0.50	0.14	0.50	0.42	0.58
Network Size (Stock)	120.9	153.4	62.0	19.0	154.0
Stddegree	22.3	29.4	10.6	3.6	30.3
Rho	-0.16	0.26	-0.17	-0.33	0.01
Mkt_Delay	0.36	0.28	0.28	0.13	0.54
Ind_Delay	0.33	0.28	0.24	0.11	0.50
Illiquidity	0.06	0.25	0.01	0.00	0.03
NumAnalyst	6.96	6.30	6.00	2.00	10.00
Network Size (Fund)	116.7	99.1	54.0	16.0	131.0

Panel B

Variable	NETDENSITY	RESDENSITY	CVRATIO	NETWORK SIZE
NETDENSITY	**1.00**			
RESDENSITY	**0.82**	**1.00**		
CVRATIO	**-0.66**	**-0.70**	**1.00**	
STDDEGREE	**-0.24**	-0.03	**0.14**	**0.98**
NETWORK SIZE	**-0.30**	**-0.09**	**0.14**	**1.00**
MCAP	**-0.30**	0.05	**0.16**	**0.62**
BMRATIO	**0.11**	**0.08**	**-0.07**	**-0.13**
TURNOVER	-0.04	**-0.11**	**0.06**	-0.02
INSTOWN	**-0.12**	**-0.10**	**0.11**	**0.19**

Panel C

	Fund Network				Stock Ownership Network		
		Observed		Random Graph		Observed	
	Year +1	Year + 3	Year +5		Year+1	Year+3	Year +5
Mean	54	33	27	12	77	58	54
Median	55	33	25	11	77	58	53
Stdev	9	9	9	7	4	6	7
Min	39	16	13	4	69	46	42
Max	70	50	41	32	87	77	68

Table II

Information Network Effects in Mutual Fund Trades

This table presents the results for the OLS regression corresponding to equation (2) of the text. The dependent variable is the next quarter change in the percentage holding of a stock by a fund. Table shows the regression coefficients for following independent variables: change in average portfolio weight of the stock within the information network of a fund excluding the fund itself (Dw_NETWORK), change in average portfolio weight of the stock for all mutual funds in the sample except for the given fund (Dw_ALL) and the dummy variable denoting whether the fund belongs to the information network of the given stock or not (IN). Stock's book to market ratio (BMRATIO), past twelve month return (MOM12), market capitalization (MCAP), average daily turnover (TURNOVER), mean analyst recommendation (MEANREC) and change in mean analyst recommendation (ChangeREC) are included as control variables. Style controls include 15 fund level variables that are calculated by weighting stock book to market, size and past 6 month returns quintile dummies by that fund's buys or sells. Fund-level industry control variables are calculated similarly for 12 industry groups. The regression includes mutual fund holdings data from March 1980 to December 2006. Following the Fama-MacBeth (1973) approach, all the regression specifications except in columns 5 and 6 are performed quarterly and time series averages of the quarterly coefficients are reported. Regression specification 5 is a Probit regression with dependent variable taking two values: 0 if the fund doesn't hold a stock in both quarters t-1 and t, 1 if fund that doesn't hold a stock in quarter t-1, buys it in quarter t. In regression specification 6, I estimate a pooled panel regression and cluster standard errors by quarter and stock. Coefficients significant at the 5% level are denoted in bold and t-statistics are given in parentheses. Standard-errors are shown in italics.

					PROBIT	CLUSTERED		
				INT	EXT			
	1	2	3	4	5	6	7	8
Dw_NETWORK	**0.2206**	**0.0758**	**0.2559**	**0.6112**	**41.585**	**0.2655**	**0.1493**	**0.1496**
	(12.48)	**(9.10)**	**(8.83)**	**(18.58)**	***0.9550***	**(8.65)**	**(8.66)**	**(8.69)**
Dw_ALL	**0.3850**	**0.3779**	**0.4524**	**2.2635**	**322.600**	**0.4544**	**0.3649**	**0.3647**
	(11.40)	**(15.38)**	**(13.08)**	**(11.94)**	***2.6970***	**(12.38)**	**(11.32)**	**(11.32)**
Dw_NETWORK*IN		**0.2352**						
		(12.80)						
Dw_ALL*IN		-0.0115						
		(-0.26)						
IN		**-0.0002**						
		(-9.20)						
log(MCAP)			**0.000**					
			(-3.39)					
BMRATIO			**0.000**					
			(-2.53)					
TURNOVER			0.000					
			(-0.92)					
MOM12			0.000					
			(1.37)					
MEANREC			**0.000**					
			(-1.22)					
ChangeREC			0.000					
			(-1.31)					
Style Controls	No	No	No	No	No	No	Yes	Yes
Industry Controls	No	No	No	No	No	No	Yes	No
R-Square (%)	0.71	0.89	0.90	2.36		0.93	14.66	14.58
N	108	108	53	108	43554090	43554090	108	108

Table III

Correlated Trading and Style Investing

This table presents the results of OLS regressions with change in the percentage holding of a stock by a fund as a dependent variable. The table reports regression coefficients for following independent variables: change in average portfolio weight of the stock for other funds within the stock ownership network of a fund and following the same investment style as the fund itself or following different styles. Investment style is defined either by the fund's benchmark index or by sorting funds into groups by the weighted book to market ratio and size quintiles for their portfolio. Coefficient corresponding to change in average portfolio weight of the stock for all the mutual funds in the sample except for the fund itself is also reported. Regressions are performed quarterly and time series averages of the quarterly coefficients are reported. Coefficients significant at the 5% level are denoted in bold and t-statistics are given in parentheses.

	STYLE: SAME BENCHMARK			STYLE: SIZE AND BK/MKT		
	1	2	3	4	5	6
Ownership Network-same style	**0.0596**		**0.0551**	**0.0546**		**0.0518**
	(11.25)		**(12.86)**	**(8.09)**		**(9.01)**
Ownership Network-different style		**0.0963**	**0.0915**		**0.0941**	**0.0903**
		(8.27)	**(8.69)**		**(11.48)**	**(13.07)**
All Mutual Funds	**0.5925**	**0.5360**	**0.4790**	**0.5929**	**0.5427**	**0.4844**
	(12.20)	**(13.46)**	**(12.52)**	**(13.84)**	**(12.64)**	**(12.50)**
Average R-Square (%)	0.533	0.583	0.707	0.460	0.480	0.612
Number of Quarters	108	108	108	108	108	108

Table IV

Correlated Trading: Interaction between Fund Location and Stock Ownership Network

This table presents the results for OLS regression corresponding to equation (2) of the text. The dependent variable is the change in the percentage holding of a stock by a fund. LOCAL is calculated as change in average portfolio weight of a stock by funds in the local network of a given fund, defined by all other funds within 100 miles radius of that fund's location; OWNERSHIP is the change in average portfolio weight of a stock by all funds in the stock ownership network of a given fund, defined by all other funds that invest 5% in at least one stock in which that fund also invests 5% of its portfolio. LOC_OWNERSHIP is the change in average holding by set of funds in a given fund's local network which also belong to its stock ownership network. The changes in average holdings of the remaining local funds that don't belong to a fund's stock ownership network are given by LOC_NONOWNERSHIP. The average change in holding of a given stock by entire mutual fund sector excluding the given fund itself is given by "All Mutual Funds". OWN_LOCAL is the change in average holding by set of funds in the given fund's stock ownership network also located within 100 miles radius of that fund. The changes in average holdings of the remaining funds in the stock ownership network that are nonlocal are given by OWN_NONLOCAL. The regression includes mutual fund holding data from September 2003 to December 2006. Following the Fama-MacBeth (1973) approach, the regressions are performed quarterly and time series averages of the quarterly coefficients are reported. Regression specification 3 and 6 are Probit regressions with dependent variable taking two values: 0 if the fund doesn't hold a stock in both quarters t-1 and t, 1 if fund that doesn't hold a stock in quarter t-1, buys it in quarter t. Coefficients significant at 5% level are denoted in bold and t-statistics are given in parentheses. Standard-errors are shown in italics.

		INT	EXT		INT	EXT	
	1	2	3	4	5	6	7
OWNERSHIP	**0.256**	**0.646**	**37.383**				
	(18.69)	**(18.93)**	***2.627***				
LOCATION	**0.077**	**0.330**	**42.185**				
	(3.21)	**(3.16)**	***4.444***				
OWN_LOCAL							**0.0364**
							(9.74)
OWN_NONLOCAL							**0.1685**
							(10.08)
LOC_OWNERSHIP				**0.038**	**0.127**	**7.054**	
				(9.81)	**(7.11)**	***0.972***	
LOC_NON OWNERSHIP				0.006	0.049	0.716	
				(0.38)	(0.49)	*4.378*	
All Mutual Funds	**0.549**	**2.896**	**382.000**	**0.918**	**4.326**	**465.200**	**0.7038**
	(17.22)	**(7.85)**	***10.184***	**(20.62)**	**(9.12)**	***9.712***	**(23.61)**
Average R-Square (%)	0.660	2.33		0.37	1.60		0.56
N	14	14	7507652	14	14	6735542	14

Table V

Residual Network Density and Momentum Returns

This table presents the results for the effect of residual network density of the information network of a stock on future momentum profits. Network density for a stock is measured as the ratio of the number of linkages in its information network to the maximum number of possible links. Residual network density is the residual obtained by regressing the logit transformation of network density on the log of market capitalization. The regression to calculate residual network density is estimated each quarter on the cross-section of all stocks in the sample. Stocks are then sorted into three equal portfolios based on residual network density and then further independently sorted into five equal portfolios based on past 6-month returns. Panel A presents the equal-weighted average monthly raw returns for the residual density-momentum portfolios measured over the next 6 to 12 months. Panel B presents the Fama-French 3-factor alphas for the same portfolios. All the returns are in monthly percentage. 5% significance level is denoted in bold and t-statistics are given in parentheses.

Panel A

	6 Months					1 Year				
	Raw Returns					Raw Returns				
	Residual Network Density					Residual Network Density				
Momentum	Uncond	RD1	RD2	RD3	RD3-RD1	Uncond	RD1	RD2	RD3	RD3-RD1
R1	**0.92**	**0.80**	**0.87**	**1.10**	**0.30**	**0.98**	**0.82**	**0.99**	**1.11**	**0.28**
	(2.67)	**(2.24)**	**(2.50)**	**(3.14)**	**(2.42)**	**(3.00)**	**(2.46)**	**(3.06)**	**(3.34)**	**(3.03)**
R2	**1.25**	**1.20**	**1.21**	**1.33**	0.13	**1.18**	**1.16**	**1.16**	**1.23**	0.07
	(4.74)	**(4.23)**	**(4.75)**	**(5.00)**	(1.45)	**(4.61)**	**(4.28)**	**(4.64)**	**(4.77)**	(0.94)
R3	**1.31**	**1.45**	**1.24**	**1.24**	**-0.20**	**1.25**	**1.32**	**1.18**	**1.24**	-0.08
	(5.26)	**(5.33)**	**(5.07)**	**(5.07)**	**(-2.29)**	**(5.03)**	**(4.92)**	**(4.86)**	**(5.12)**	(-0.93)
R4	**1.37**	**1.43**	**1.31**	**1.36**	-0.07	**1.26**	**1.29**	**1.26**	**1.23**	-0.06
	(5.34)	**(5.18)**	**(5.09)**	**(5.42)**	(-0.78)	**(4.97)**	**(4.80)**	**(4.96)**	**(4.96)**	(-0.82)
R5	**1.77**	**1.81**	**1.74**	**1.73**	-0.08	**1.46**	**1.48**	**1.49**	**1.41**	-0.06
	(5.25)	**(5.08)**	**(5.12)**	**(5.19)**	(-0.67)	**(4.42)**	**(4.26)**	**(4.57)**	**(4.30)**	(-0.62)
R5-R1	**0.85**	**1.01**	**0.87**	**0.63**	**-0.38**	**0.48**	**0.65**	**0.50**	0.31	**-0.34**
	(3.54)	**(4.05)**	**(3.40)**	**(2.47)**	**(-2.59)**	**(2.47)**	**(3.22)**	**(2.44)**	(1.50)	**(-2.96)**

Panel B

	6 Months Equal Weighted 3-Factor Alpha Residual Network Density					1 Year Equal Weighted 3-Factor Alpha Residual Network Density				
Momentum	Uncond	RD1	RD2	RD3	RD3-RD1	Uncond	RD1	RD2	RD3	RD3-RD1
R1	**-0.48**	**-0.59**	**-0.53**	-0.32	**0.27**	**-0.38**	**-0.52**	**-0.35**	-0.28	**0.24**
	(-2.78)	**(-3.15)**	**(-2.97)**	(-1.68)	**(2.15)**	**(-2.51)**	**(-3.22)**	**(-2.34)**	(-1.72)	**(2.52)**
R2	-0.08	-0.18	-0.06	-0.01	0.17	-0.09	-0.14	-0.08	-0.05	0.09
	(-0.85)	(-1.52)	(-0.56)	(-0.05)	(1.92)	(-1.01)	(-1.26)	(-0.88)	(-0.48)	(1.16)
R3	0.01	0.13	-0.05	-0.05	**-0.18**	0.01	0.07	-0.04	0.01	-0.06
	(0.14)	(1.13)	(-0.53)	(-0.53)	**(-2.03)**	(0.12)	(0.60)	(-0.49)	(0.08)	(-0.73)
R4	0.12	0.14	0.11	0.10	-0.04	0.09	0.08	0.12	0.04	-0.04
	(1.50)	(1.27)	(1.24)	(1.10)	(-0.38)	(1.18)	(0.91)	(1.48)	(0.55)	(-0.56)
R5	**0.54**	**0.57**	**0.51**	**0.51**	-0.06	**0.32**	**0.34**	**0.37**	**0.27**	-0.07
	(4.44)	**(3.88)**	**(3.58)**	**(3.90)**	(-0.49)	**(3.32)**	**(2.77)**	**(3.44)**	**(2.56)**	(-0.67)
R5-R1	**1.02**	**1.17**	**1.04**	**0.83**	**-0.34**	**0.70**	**0.85**	**0.72**	**0.54**	**-0.31**
	(4.23)	**(4.57)**	**(4.01)**	**(3.22)**	**(-2.21)**	**(3.70)**	**(4.27)**	**(3.60)**	**(2.72)**	**(-2.60)**

Table VI

Information Network and Momentum Returns: Fama MacBeth Regressions

This table presents regression evidence on the effect of network density on future stock momentum returns. In specifications 1 to 3, dependent variable is future twelve month stock returns (RET12MONTH) or future six month stock returns (RET6MONTH). In specifications 4 and 5, the serial correlation of six month excess returns (Rho) measured over the next five years is included as a dependent variable. Independent variables are lagged by one quarter and include network density (NETDENSITY) and the following firm characteristics: book to market ratio (BK/MKT), size (MCAP), past 6 months returns (MOM6), past quarter average daily turnover (TURNOVER), stock price (PRC), Analyst Coverage (NUMANALYST) and institutional ownership (IO). Interactions between past 6 month returns (MOM6), Network Density (NETDENSITY) and firm characteristics are also included as independent variables. 5% significance level is denoted in bold and t-statistics are given in parentheses.

Independent Variables	1 RET12MONTH	2 RET12MONTH	3 RET6MONTH	4 Rho	5 Rho
LOG(IO)					**-0.025**
					(-3.51)
LOG(BK/MKT)	**0.027**	**0.026**	**0.013**		0.007
	(2.85)	**(2.77)**	**(2.48)**		(1.24)
LOG(MCAP)	-0.005	-0.007	-0.004	-0.001	**0.009**
	(-0.78)	(-0.98)	(-1.71)	(-0.35)	**(2.14)**
LOG(TURNOVER)	-0.006	-0.006	-0.005		**0.008**
	(-0.44)	(-0.41)	(-0.69)		**(2.34)**
MOM6	**0.156**	0.223	**0.238**		**0.026**
	(2.20)	(1.85)	**(4.22)**		**(2.44)**
LOG(NETDENSITY)	-0.003	-0.001	-0.001	**-0.012**	**-0.015**
	(-0.37)	(-0.09)	(-0.12)	**(-2.58)**	**(-2.71)**
LOG(1+NUMANALYST)		0.007	0.006		**-0.016**
		(0.86)	(1.20)		**(-3.57)**
MOM6*LOG(NETDENSITY)	**-0.120**	**-0.100**	**-0.031**		
	(-3.44)	**(-3.11)**	**(-2.06)**		
MOM6*LOG(MCAP)	-0.020	**-0.027**	**-0.015**		
	(-1.92)	**(-2.12)**	**(-2.16)**		
MOM6*LOG(TURNOVER)		0.012	-0.005		
		(0.65)	(-0.69)		
MOM6*LOG(1+NUMANALYST)		0.028	0.011		
		(1.58)	(0.88)		
Average R-Square	9.07	10.31	10.36	0.91	3.55
Number of Quarters	108	108	108	104	104

Table VII

Network Density and Price Delay

This table presents the results of annual Fama-MacBeth regressions of future market or industry price delay measures on past network density (NETDENSITY) and the following firm characteristics: market capitalization (MCAP), past 12 months returns (MOM12), average daily turnover (TURNOVER), stock price (Price), Analyst Coverage (NUMANALYST), Amihud Illiquidity Measure (Illiquidity) and institutional ownership (Inst_Own). Independent variables are lagged and are measured at the end of the year prior to the calculation of the Price delay measures. The logit transformation of price delay is included as dependent variable because the delay variable is bounded between 0 and 1. The measures of delay are calculated according to equation 5 of the text and are defined as the fraction of variation of contemporaneous individual stock returns explained by lagged market or industry returns. 5% significance level is denoted in bold and t-statistics are given in parentheses.

	Dependent Variable			
	Logit(Mkt_Delay)	Logit(Mkt_Delay)	Logit(Ind_Delay)	Logit(Ind_Delay)
Independent Variables	1	2	3	4
Log(MCAP)	**-0.276**	**-0.260**	**-0.413**	**-0.391**
	(-9.86)	**(-9.15)**	**(-13.36)**	**(-11.62)**
NetDensity	**-0.222**	**-0.205**	**-0.136**	**-0.115**
	(-4.64)	**(-4.38)**	**(-3.46)**	**(-3.08)**
Log(Price)	-0.057	-0.054	0.029	0.035
	(-1.06)	(-1.01)	(0.52)	(0.64)
MOM12	**-0.163**	**-0.179**	**-0.157**	**-0.178**
	(-2.62)	**(-2.78)**	**(-3.03)**	**(-3.52)**
Log(Turnover)	-0.059	-0.036	-0.049	-0.020
	(-1.04)	(-0.62)	(-1.09)	(-0.43)
Log(Inst_Own)	**-0.328**	**-0.309**	**-0.203**	**-0.177**
	(-6.17)	**(-5.50)**	**(-3.38)**	**(-2.74)**
Log(1+Num_Analyst)	-0.023	-0.023	**-0.127**	**-0.127**
	(-0.65)	(-0.66)	**(-4.08)**	**(-4.11)**
Illiquidity		1.391		1.526
		(1.54)		(1.78)
Average R-Square (%)	11.78	12.06	16.75	17.16
Number of Years	26	26	26	26

Table VIII

Network Centralization and Stock Volatility: Cross-sectional evidence

This table reports the coefficients corresponding to the following panel regression estimated using the data from March 1980 to December 2006:

$$\log(\sigma_{i,t}^{STOCK}) = \alpha + \beta \log(CENTRALIZATION_{i,t-1}) + \delta_j Control\_Variable_{j,t-1} + \varepsilon_{i,t}$$

The dependent variable is the natural logarithm of the stock idiosyncratic risk (measured as the standard deviation of residuals from a 3-factor Fama and French model). Idiosyncratic risk at the end of a given quarter is estimated using daily returns data from that quarter. The independent variables are from the end of previous quarter and include: standard deviation (STDDEGREE) and Coefficient of Variation (CVRATIO) of number of connections of mutual funds in a given stock's ownership network with other funds in the network, market capitalization (MCAP), book to market ratio (BMRATIO), average daily turnover (TURNOVER), institutional ownership (IO), stock price (PRC), past 12 month return (MOM12) and lagged idiosyncratic risk (LAGIDIORISK). The independent and dependent variables are standardized to allow for comparison across variables and across specifications. The t-statistics (reported in parentheses) are based on standard errors clustered along stock and quarter dimensions. The coefficients significant at the 5% level are denoted in bold.

					1980-1995	1996-2006
	1	2	3	4	5	6
Independent Variable	Dependent Variable: log(IDIORISK)					
log(CVRATIO)	**0.051**	**0.032**	**0.013**		**0.012**	**0.012**
	(10.20)	**(6.93)**	**(3.88)**		**(2.69)**	**(2.73)**
log(STDDEGREE)				**0.014**		
				(2.47)		
log(MCAP)	**-0.179**	**-0.138**	**-0.041**	**-0.048**	**-0.053**	**-0.034**
	(-13.35)	**(-8.11)**	**(-5.08)**	**(-5.27)**	**(-4.44)**	**(-3.66)**
log(PCTIO)		0.021	**0.066**	**0.067**	**0.058**	**0.073**
		(1.09)	**(4.38)**	**(4.44)**	**(4.93)**	**(2.96)**
log(BMRATIO)		**-0.089**	**-0.037**	**-0.037**	**-0.029**	**-0.038**
		(-9.35)	**(-6.88)**	**(-6.93)**	**(-4.07)**	**(-6.01)**
log(PRC)		**-0.138**	**-0.056**	**-0.057**	**-0.093**	**-0.034**
		(-3.98)	**(-3.82)**	**(-3.80)**	**(-5.29)**	**(-2.63)**
MOM12		**0.051**	**0.031**	**0.031**	**0.035**	**0.030**
		(3.92)	**(3.98)**	**(4.01)**	**(3.18)**	**(3.13)**
log(TURNOVER)		**0.164**	**-0.025**	**-0.025**	-0.006	**-0.042**
		(12.38)	**(-2.12)**	**(-2.11)**	(-0.38)	**(-2.47)**
log(LAGIDIORISK)			**0.375**	**0.376**	**0.322**	**0.398**
			(23.08)	**(23.06)**	**(9.82)**	**(24.56)**
R-Square (%)	3.51	10.95	19.38	19.37	16.01	21.62
Clustered(Firm,Qtr)	Yes	Yes	Yes	Yes	Yes	Yes
N	62430	62430	62430	62430	31153	31277

Table IX

Network Centralization and Aggregate Volatility: Time Series Regressions

Panel A reports the coefficients corresponding to the following time series regression estimated using quarterly data from March 1980 to December 2006:

$$\sigma_t^{AGGREGATE} = \alpha + \beta t + \delta AGG\_CVRATIO_{t-1} + \gamma_j ControlVariable_{j,t-1} + \varepsilon_t$$

The dependent variable is a measure of aggregate stock volatility: IDIORISK denotes average idiosyncratic volatility of stocks in the sample calculated using a Fama French 3-factor model estimated over a period of one quarter using daily returns data, S&P VOL is the standard deviation of S&P returns. Volatility measures at the end of a given quarter are calculated using daily returns data from that quarter. Independent variables are from the end of previous quarter and include: coefficient of variation (AGG_CVRATIO) of the number of linkages of all connected mutual funds in the economy and average institutional ownership (PCTIO) for the stocks in the sample. The lags of average idiosyncratic risk (LAGIDIORISK) and S&P500 volatility (LAGS&PVOL) are also included as controls. (The t-statistics (reported in parentheses) are based on Newey and West (1987) standard errors with four lags. Panel B shows the results of the Granger Causality test between average idiosyncratic risk (IDIORISK) and aggregate coefficient of variation (AGG_CVRATIO). The coefficients significant at 5% level are denoted in bold and t-statistics are given in parentheses.

Panel A

	Dependent Variable				
Independent Variable	IDIORISK	IDIORISK	IDIORISK	S&PVOL	S&PVOL
Intercept	0.0030	0.0025	**-0.0006**	0.0011	-0.0004
	(0.42)	(0.37)	**(-0.29)**	(0.23)	(-0.13)
AGG_CVRATIO	**0.0186**	**0.0207**	0.0044	0.0100	0.0065
	(2.05)	**(2.30)**	(1.70)	(1.68)	(1.70)
LAGIDIORISK			**0.8657**		
			(15.40)		
LAGS&PVOL					**0.4559**
					(3.43)
IO		0.0000	**0.0000**		
		(-0.20)	**(-0.85)**		
Time	0.0000			0.0000	0.0000
	(0.28)			(-0.21)	(-0.28)
R-Square	11.48	11.32	76.35	3.00	23.45
N	108	108	108	108	108

Panel B

GRANGER-CAUSALITY TEST		
	IDIORISK	AGGCVRATIO
LAG_IDIORISK	**0.84**	0.581
	(16.37)	(0.78)
LAG_AGGCVRATIO	0.0056	**0.8755**
	(1.85)	**(19.84)**
Constant	-0.0018	**0.098**
	(-0.20)	**(2.69)**
R-Square	0.754	0.8095
N	107	107
Wald Chi-Square test	3.4089	0.6062
p-value	0.065	0.436

Table X

Robustness Checks: Alternative Network Specification

Panel A presents the results for a momentum portfolio strategy based on lagged residual network density and past 6 month returns. The returns are calculated over a holding period of next 6 months. Panel B reports the regression results for the effect of the measures of network structure on momentum strategies and the cross-section of idiosyncratic volatility. The network measures are calculated using the definition of information networks based on deviations of the holdings from the benchmark index. In specifications 1 and 2, the dependent variable is the next six-month or twelve-month stock returns. The coefficients and t-statistics are calculating using a Fama-MacBeth methodology. The dependent variable in specification 3 and 4 is the natural logarithm of stock idiosyncratic risk (measured as the standard deviation of residuals from a 3-factor Fama and French model). Idiosyncratic risk at the end of a given quarter is estimated using daily returns data from that quarter. The independent variables are from the end of previous quarter and include: Coefficient of Variation (CVRATIO) of number of connections of mutual funds in the given stock's ownership network with other funds in the network, market capitalization (MCAP), book to market ratio (BMRATIO), turnover (TURNOVER), institutional ownership (IO), stock price (PRC), past 6 month return (MOM6), lagged idiosyncratic risk (LAGIDIORISK), network density (DENSITY) and analyst coverage (NUMANALYST). The independent and dependent variables are standardized in the volatility regressions (Specifications 3 and 4) to allow for comparison across variables and across specifications. In specifications 3 and 4, the t-statistics (reported in parentheses) are based on standard errors clustered along stock and quarter dimensions (Petersen (2009)). The coefficients significant at the 5% level are denoted in bold.

Panel A

		6 Months Equal Weighted 3-Factor Alpha Residual Network Density			
Momentum	Uncond	RD1	RD2	RD3	RD3-RD1
R1	**-0.48**	**-0.63**	**-0.57**	-0.29	**0.34**
	(-2.39)	**(-2.93)**	**(-2.83)**	(-1.31)	**(2.67)**
R2	-0.07	-0.01	-0.19	0.02	0.04
	(-0.67)	(-0.11)	(-1.66)	(0.19)	(0.32)
R3	-0.05	-0.02	-0.03	-0.08	-0.06
	(-0.51)	(-0.16)	(-0.31)	(-0.80)	(-0.62)
R4	0.09	0.16	0.00	0.09	-0.07
	(1.03)	(1.43)	(0.05)	(0.91)	(-0.66)
R5	**0.51**	**0.69**	**0.45**	**0.31**	**-0.39**
	(3.72)	**(4.20)**	**(3.17)**	**(2.02)**	**(-2.91)**
R5-R1	**0.99**	**1.32**	**1.02**	**0.60**	**-0.73**
	(3.65)	**(4.56)**	**(3.69)**	**(2.00)**	**(-4.04)**

Panel B

	MOMENTUM		VOLATILITY	
	Reg1	Reg2	Reg3	Reg4
Dependent Variable	RET6MONTH	RET12MONTH	LOG(IDIORISK)	LOG(IDIORISK)
LOG(CVRATIO)			**0.097**	**0.026**
			(10.82)	**(5.97)**
LOG(BK/MKT)	0.008	0.015		**-0.073**
	(1.40)	(1.42)		**(-10.20)**
LOG(MCAP)	-0.001	0.002	**-0.372**	**-0.088**
	(-0.20)	(0.33)	**(-17.16)**	**(-7.43)**
LOG(TURNOVER)	0.002	0.007		**0.036**
	(0.20)	(0.45)		**(3.07)**
LOG(IO)				0.010
				(1.37)
LOG(PRICE)				**-0.059**
				(-3.51)
MOM6	**0.226**	0.120		**0.085**
	(4.42)	(1.27)		**(3.07)**
LOG(1+NUMANALYST)	0.006	0.006		
	(0.82)	(0.62)		
LOG(LAGIDIORISK)				
MOM6*LOG(DENSITY)	**-0.043**	**-0.104**		
	(-2.43)	**(-3.00)**		
MOM6*LOG(TURNOVER)	0.019	0.010		
	(1.55)	(0.51)		
MOM6*LOG(MCAP)	-0.017	-0.015		
	(-1.69)	(-1.11)		
MOM6*LOG(1+NUMANALYST)	0.011	0.016		
	(0.54)	(0.57)		
Avg NObs	822	798		
R-Square	9.01	8.82	14.92	63.14
N	88	88	62239	62239

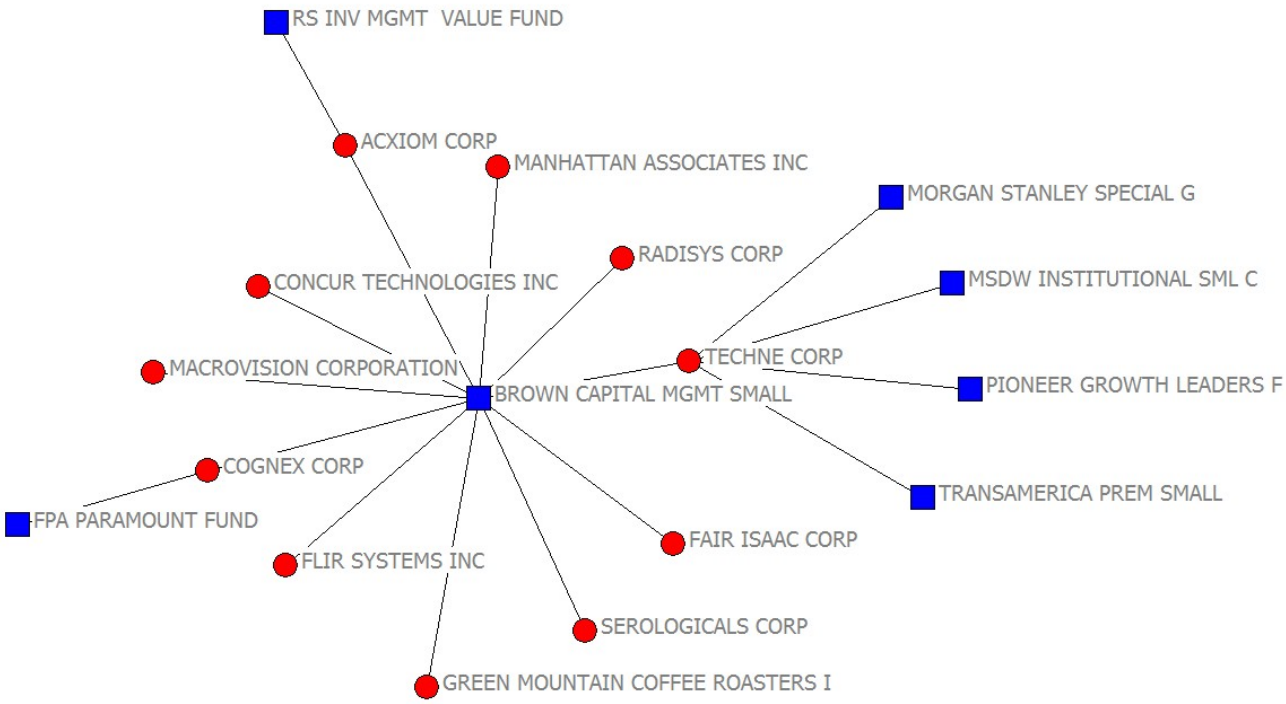

Figure 1 Information Network of a Fund: An Example

This figure shows the information network of "Brown Capital Management Small Company" fund as an example of an information or stock ownership network of a mutual fund. Links represented by solid lines are determined by investment of 5% or more of the fund portfolio in the given stock. Stocks are represented by solid circles and funds by solid squares.

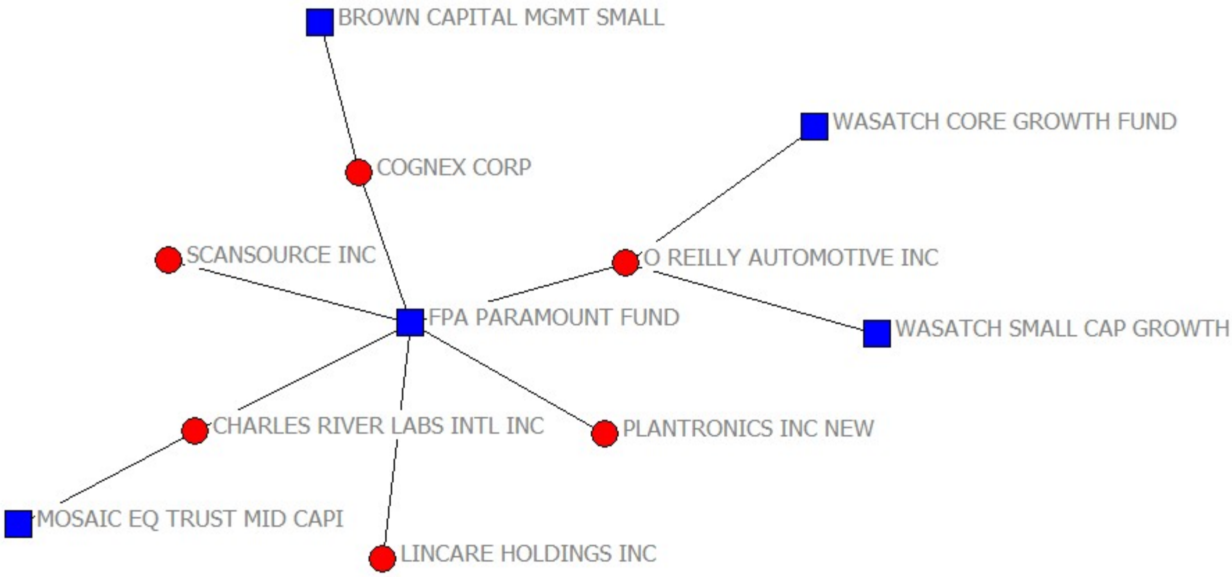

Figure 2 Information Network of a Fund: Another Example

This figure shows the information network of "FPA Paramount" fund as another example of an information or stock ownership network of a mutual fund. Links represented by solid lines are determined by investment of 5% or more of the fund portfolio in the given stock. Stocks are represented by solid circles and funds by solid squares.

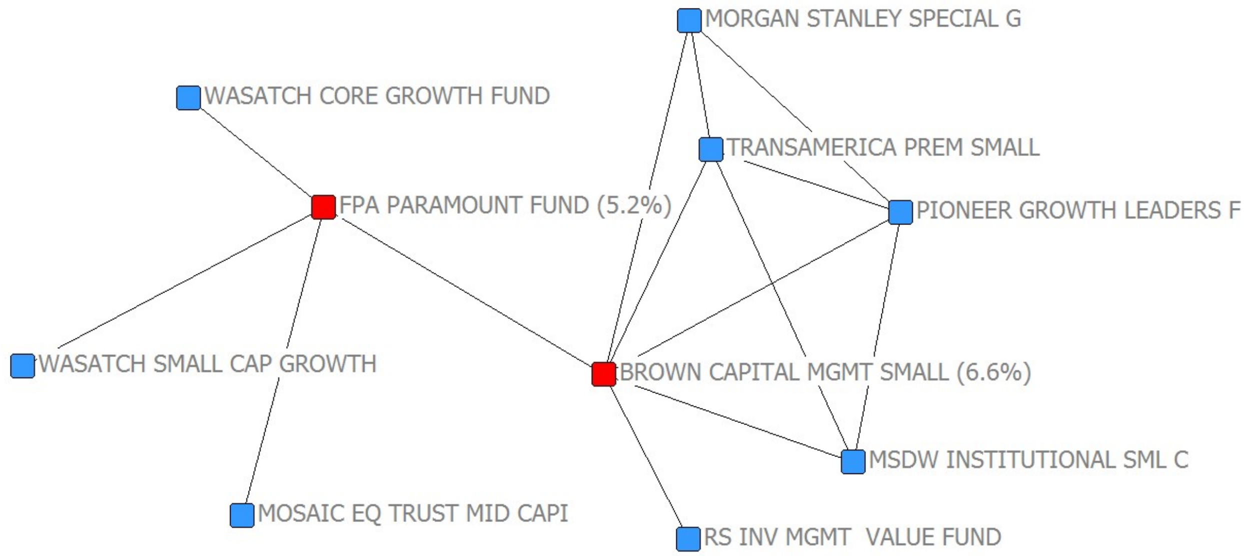

Figure 3 Ownership Network of a Stock: An Example

This figure shows the stock ownership network of Cognex Corporation (CGNX) in December 2005 as an illustrative example of a stock's information network. "Brown Capital Management Small Company Fund" and "FPA Paramount Fund" are the two mutual funds that invest more than 5% of their portfolio in CGNX. The mutual funds are at the nodes represented by solid circles. The stock ownership linkages represented by solid lines are determined by investment of 5% or more of the fund portfolio in the same stock.

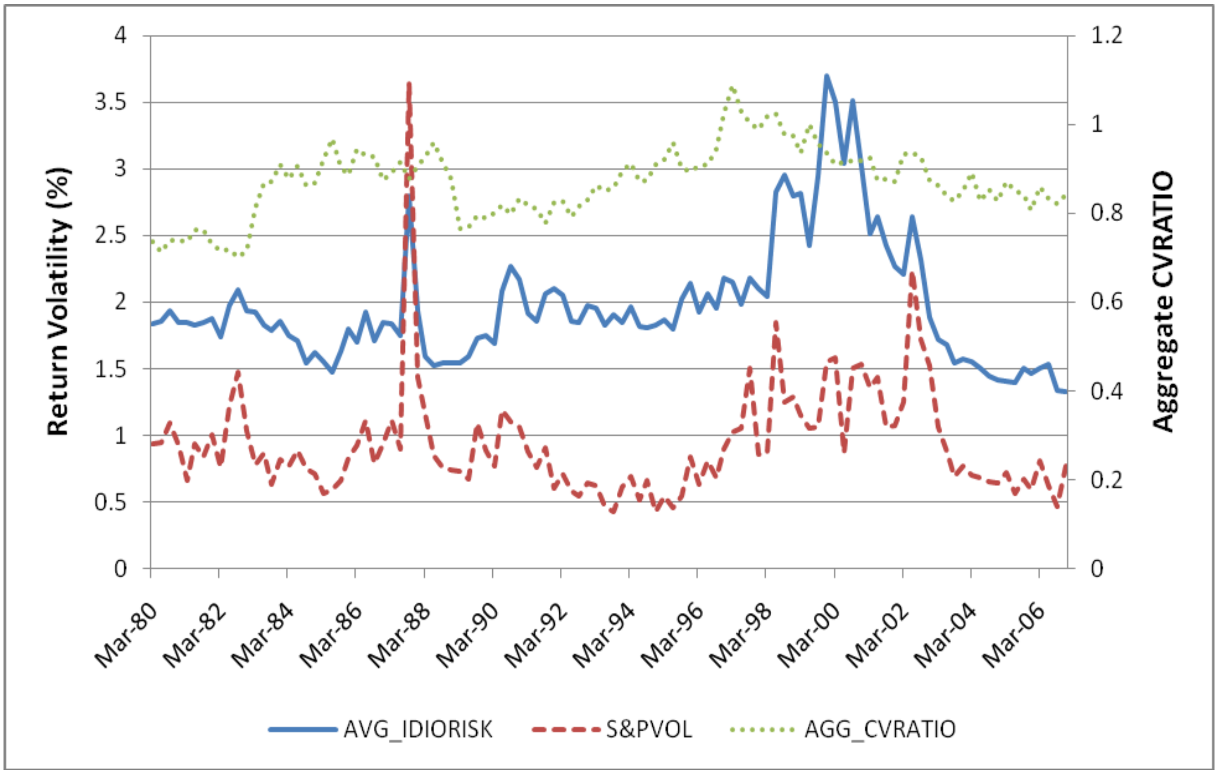

Figure 4 Aggregate Network Centralization Measure and Average Idiosyncratic Volatility: 1980-2006

This figure shows the time-series relationship between the average idiosyncratic risk in the economy (AVG_IDIORISK) and a centralization measure of the aggregate economy-wide information network. The aggregate centralization measure is lagged by one quarter and is measured as the coefficient of variation (AGG_CVRATIO) of the number of connections of all the connected mutual funds in the economy at any given time.