Social Connections and Information Production: Evidence from Mutual Fund Portfolios and Performance

Yuyuan (Anthony) Zhu*

December 16, 2016

Abstract

While connected investors have access to information in their social network (information diffusion effect), social connections also reduce their incentives to acquire costly information, since they can free ride on connected peers ("free riding on friends" effect). In this paper, I find this negative "free riding on friends" effect of social connections dominates information diffusion effect in the mutual fund industry, using fund managers' connections built upon their prior career experiences. First, I find that connected funds are more likely to hold the same stocks and to trade in the same direction, relative to unconnected funds. Second, I find that funds with lower network centrality earn higher alphas, even after controlling for other fund and manager characteristics. A one-standard-deviation increase in eigenvector centrality predicts a decrease of 29-37 basis points in annualized fund alphas. Third, when I define a stock-level variable PMC (Peripheral minus Central) as the difference in average portfolio weights between peripheral funds and central funds, I find that stocks with higher PMC have significantly higher abnormal stock returns. A one-standard-deviation increase in PMC predicts an increase of 1.48%-1.52% in the next quarter risk-adjusted returns (annualized). Finally, I find that PMC predicts firms' future earnings surprises.

JEL Classification: G11, G23
Keywords: Social Connections; Mutual Funds; Endogenous Information Production

---

*PhD Candidate, Fulton Hall 337, Carroll School of Management, Boston College, Chestnut Hill, MA 02467. Phone: 617-552-2050. E-mail: zhuyh@bc.edu.

I thank my dissertation committee, Thomas Chemmanur (Chair), Leonard Kostovetsky, Alan Marcus, Jonathan Reuter, and Robert Taggart for their valuable comments, feedback, and encouragement. I also thank Ronnie Sadka, Darren Kisgen, Edie Hotchkiss, Ian Appel, David Hirshleifer, Brad Barber, Aymen Karoui, and seminar participants at Boston College and the 2016 Financial Management Meetings for their comments. I am solely responsible for any remaining errors or omissions.

1 Introduction

Communication based on social connections among investors is an important part of investment processes in financial markets. Casual observation suggests that investors frequently share and communicate their investment ideas and strategies, even among professional money managers who might be competitors for returns and flows.[1] Shiller and Pound (1989) present survey evidence that both institutional and individual investors may be influenced by peer communications. Due to similar "word-of-mouth" effects, geographically proximate investors are more likely to exhibit similar trading behaviors compared to geographically distant investors (e.g. Hong, Kubik, and Stein (2005); Pool, Stoffman, and Yonker (2015)).

Despite the growing evidence that investors invest similarly with their socially connected peers, there is relatively little analysis linking social connections to investors' investment performance. This paper aims to fill this gap in the literature by providing new evidence on the link between social connections and investment performance using data on mutual fund holdings and returns. The mutual fund industry is an ideal setting to study the social connections between investors due to the rich amount of background information available on mutual fund managers from regulatory filings.

Theoretically, the effect of social connections on investors' performance remains ambiguous. At first glance, better connected investors may have access to better and more precise information, since they have a higher chance of receiving more valuable signals (information diffusion effect). This is, however, not necessarily the case when information production by investors is endogenous. Han and Yang (2013) analyze a Grossman and Stiglitz (1980) style economy with the addition of a social network. Investors have three sources of information: the market price; costly information production; and communication with other traders through a social network. They argue that under endogenous information production, social connections may reduce investors ex-ante incentive to acquire costly signals, since they can free-ride on "connected" peers ("free riding on friends" effect). Due to this "free riding on friends" effect, better connected investors, in the aggregate, may hold less precise information, compared to less connected investors. Given the two opposing effects

[1] Stein (2008) rationalizes this phenomenon that the truthful information exchange among competitors exists because of the complementarity in their information structure. Another reason for information sharing is to attract additional arbitrage capital to successfully correct mispricing. For empirical evidence, see Gray, Crawford, and Kern (2012).

of social connections on the precision of information held by investors, whether better connected investors will have better or worse investment performance, is an empirical question.[2]

To answer this question, I begin by asking whether social connections have an effect on the portfolio holdings and trades of mutual funds. I use data on the career paths of managers within the mutual fund industry to construct my proxy for social connections between mutual funds. I identify two fund managers as "connected" today if they both work as portfolio managers in the same fund family at a particular time point in the past.[3] Since I conduct the empirical analysis at the fund level, I further define a pair of funds as "connected" if they have at least one pair of "connected" portfolio managers. I construct measures of pairwise overlap in holdings and trades for all fund pairs, and test whether the overlap is greater when a pair of funds is connected. Remarkably, the portfolio overlap for a pair of connected funds is 18% higher in my baseline model than that of a pair of unconnected funds, even after controlling for funds' geographical locations, family memberships, size, and investment styles. The effect is economically significant and of similar magnitude for overlap in stock purchases and sales.

While I use career experiences to proxy for the social connection between fund managers, this connection variable may be correlated with other "unobserved" manager characteristics (e.g. ethnicity, political affiliation). If these "unobserved" characteristics of fund managers drive both the formation of managers' social connections and their portfolio choices, then the similarity in portfolio choices between connected mutual funds is not driven by social interactions, but rather by these "unobserved" manager characteristics.[4] To rule out this alternative hypothesis, I build a "future" version of social connections between mutual funds based on the *future* connections

[2]Although I motivate this paper using the costly information production and the "free riding on friends" effect through social connections, there may be other reasons why social connections can have a negative effect on investment performance. For example, in the social psychology literature, social connections may induce "groupthink" phenomenon that individuals' "striving for unanimity override their motivation to realistically appraise alternative course of action" (Janis (1982)), and hence independent critical thinking will be replaced by "groupthink", resulting in irrational and inefficient decision-making. In the behavior economics literature, DeMarzo, Vayanos, and Zwiebel (2003) theoretically analyze that individuals are subject to persuasion bias, in a social network, that they fail to account for possible repetition in the information they receive through social connections.

[3]Empirically, various proxies for social connections has been studied in the literature. For education links, see Cohen, Frazzini, and Malloy (2008) and Shue (2013); for employment connections, see Gerritzen, Jackwerth, and Plazzi (2016), Spilker (2016) and Engelberg, Gao, and Parsons (2012); for geographical proximity, see Pool, Stoffman, and Yonker (2015), Hvide and Östberg (2015) and Ivković and Weisbenner (2007). In my study, I am focused on a particular dimension of social connections, past career experience in the mutual fund industry, as the "free riding on friends" incentive may be particularly strong when a pair of managers share the experience of managing money in the same fund family.

[4]Separating out this "correlated effects" from the "social effects" is empirically challenging, and has long been recognized as the "reflection problem" in the economics literature (e.g., Manski (1993)).

of *current* portfolio fund managers. Assuming these "unobserved" manager characteristics are persistent across time, then I should expect a pair of mutual funds exhibit similar portfolio choices even before they become connected, and the "future" connection variable be correlated with current overlap in portfolio holdings and trades. However, in this falsification test, I do not find that "future" social connections have a statistically significant effect on funds' portfolio holdings and trades.

Given the evidence that social connections influence mutual fund portfolio holdings and trades, I next study the effect of social connections on the investment performance of mutual funds. To quantitatively measure the connectedness of different funds, I adopt the network centrality measures developed in the social network analysis literature.[5] I find that eigenvector centrality *negatively* predicts future fund returns and alphas. A one-standard-deviation increase in fund eigenvector centrality predicts a decrease of 29-37 basis points in annualized fund returns. The predictive power of eigenvector centrality measure for fund returns holds before and after expenses, and is robust to controlling for a set of observable fund characteristics including fund size, management team size, family size, net flow, fund age, and fund turnover. Further, the predictive power of fund centrality measure survives controlling for manager characteristics measuring their ability (e.g. managers' undergraduate institution SAT scores and whether managers have an MBA degree, as studied in Chevalier and Ellison (1999)), suggesting that the above finding is not driven by less connected mutual funds hiring managers with better ability or education. Using family fixed effects model, I find both "within-family" estimator and "between-family" estimators are economically and statistically significant, suggesting that: 1) social connections affect average returns of mutual fund families, as fund families internalize social connections of their portfolio managers when making information production decisions; 2) managers' social connections affect individual fund returns even across funds with common family-level information production. Further, I find the relationship between centrality and fund returns is not driven by geographical locations of mutual funds. To summarize, these results suggest that while both effects of social connections, information diffusion effect and "free riding on friends" effect, are at play, "free riding on friends" effect plays the dominant role in this particular setting and in aggregate, more social connections lead to less information

[5]Specifically, I compute three measures of network centrality (degree, eigenvector, and closeness). I use eigenvector centrality primarily in my empirical analysis, and use other two measures of centrality as robustness checks.

production by fund managers and worse fund returns and alphas.

An alternative explanation is that the network centrality measures are correlated with past performance of its manager(s) through managerial turnovers, such that the finding of an inverse relationship between network centrality measures and fund future performance is driven by the persistence of bad performance of "Frequent Job Switchers". I address this concern using three different empirical tests. First, I control for manager tenure (in the fund family) and I do not find any impact on the predictive power of centrality measures for fund returns. Second, I construct an index variable, $MgrPerformanceHist$, to measure a fund manager's alpha generation during his entire career in the mutual fund industry. While I find $MgrPerformanceHist$ has significant predictive power for fund returns, the predictive power of centrality measures for fund returns is not affected. Third, I decompose fund centrality measures (degree centrality and eigenvector centrality) into an "In" and an "Out" component based on the direction of social connections. The direction of connection is determined by whether the manager joins a new fund family ("Out" connection) or whether another manager joins from a different fund family ("In" connection). "Frequent Job Switchers" are likely to have more "Out" connections than "In" connections, while managers with a long tenure in the family are likely to have more "In" connections than "Out" connections. Interestingly, centrality measures based on both "In" and "Out" connections exhibit predictive power for future fund returns, suggesting that the negative relationship between fund centrality measures and future fund performance is not solely driven by fund managers who are frequent job switchers.

Further, I investigate whether the information channel is driving the outperformance of less connected funds, compared to better connected funds. Using the eigenvector centrality measure defined above, I classify funds into *central* investors (those with above median eigenvector centrality) and *peripheral* investors (those with below median eigenvector centrality). To test whether *peripheral* investors hold an information advantage over *central* investors, I explore the information content contained in their portfolio holdings. Specifically, I construct a stock-level measure, PMC (Peripheral minus Central), defined as the difference in the average portfolio weights between *peripheral* investors and *central* investors. I find that PMC measure is a strong predictor for abnormal stock returns. A one-standard-deviation increase in PMC measure predicts an increase of 1.48%-1.52% (annualized) in next quarter risk-adjusted returns. The predictive power of the PMC measure persists up to three quarters after the focal date. Furthermore, I find that my PMC measure is a strong predictor for firm's earnings surprises. A one-standard-deviation increase in PMC measure predicts an increase of 20 basis points in SUE (Standardized Earnings Surprises) in quarter $t+1$ and 19 basis points in SUE in quarter $t+2$. The predictive power of PMC measure for both future abnormal stock returns and earnings surprises is consistent with that *peripheral* investors hold more precise information signals. The predictive power of PMC for earnings surprises also suggests that at least a portion of the information advantage enjoyed by *peripheral* investors is related to their ability to better forecast earnings over and above the market prevailing consensus.

Last, I also investigate whether social connections have an effect on the flow-performance relationship. I find that investors' response to lagged fund performance is much stronger for peripheral funds, compared to central funds. This result is robust to using lagged raw returns or lagged Fama-French-Carhart 4-factor alphas. This result also holds after controlling for the effect of fund age on the flow-performance relationship.[6] This finding is consistent with *peripheral* fund managers being more likely to produce independent information, and as a result, investors in mutual funds being better able to learn the stock-picking abilities of these managers, since past performance of *peripheral* fund managers is a stronger signal for their stock picking skills, compared to that of *central* fund managers.

The rest of the paper is organized as follows. Section 2 discusses the related literature. In Section 3, I describe the data and the construction of the mutual fund sample used in my empirical analysis. In Section 4, I study whether social connections have an effect on the portfolio holdings and trades for mutual funds. In Section 5, I make use of network centrality measures based on the social connections between mutual funds, and study the relationship between fund centrality measures and fund performance. In Section 6, I construct my PMC measure based on mutual fund holdings, and I study whether this PMC measure has predictive power for future abnormal stock returns and earnings surprises. Section 7 concludes.

[6] Chevalier and Ellison (1997) document that flows for younger funds are more sensitive to past performance than older funds.

2 Relation to the Existing Literature and Contribution

The findings of this paper relate to several strands of literature. First, I contribute to the growing evidence that social connections among investors affect their portfolio choices. Hong, Kubik, and Stein (2005) show that mutual fund managers in a given city tend to have more similar trading behavior than those in different cities. Pool, Stoffman, and Yonker (2015) further show that fund managers reside in the same neighborhood exchange private information, and are more likely to hold similar stocks and make the same-direction trades. Gerritzen, Jackwerth, and Plazzi (2016) find that employment in the same industry or in the same firm, among hedge fund managers, lead to more similar investment behavior in terms of systematic risk and abnormal performance. Hvide and Östberg (2015), using Norwegian individual investors' data, find that stock investment decisions of individuals are positively correlated with those of coworkers. Ivković and Weisbenner (2007) find that households' stock purchase in an industry is correlated with neighbors' purchase of stocks from that industry, and they attribute that correlation partly to word-of-mouth communication.

Second, this paper contributes to studying the effect of social connections on the investment performance. Hvide and Östberg (2015) do not find that social connections improve individual investors' welfare, but instead find evidence of investment mistakes propagating through social connections. However, there are several papers that document a diffusion effect of private information through social connections, and show that it is positive for investment performance. Using account-level trade data from Istanbul Stock Exchange, Ozsoylev, Walden, Yavuz, and Bildik (2014) show that central investors earn higher returns and trade earlier during informational events than peripheral investors. In the mutual funds setting, Pool, Stoffman, and Yonker (2015) find valuable information is transmitted among fund managers living in the same neighborhood, and they show stocks purchased by neighboring managers outperform stocks sold by neighboring managers. In the hedge fund setting, Gerritzen, Jackwerth, and Plazzi (2016) find that more connected hedge funds perform better, and prior experience in pension funds and banks aids performance. A more recent paper by Rossi, Blake, Timmermann, Tonks, and Wermers (2016), using connections among managers in UK's defined-benefit pension fund market, show that managers with high centrality in the network have better risk-adjusted returns.[7] In this paper, I show that in addition to the

[7] Rossi, Blake, Timmermann, Tonks, and Wermers (2016) define connections among managers through their connections to the investment consultants hired by defined-benefit pensions funds in UK. They acknowledge that the

information diffusion effect, which is positive for investment performance, social connections can potentially have a negative effect on fund managers' incentives to produce independent information, and in my setting, this negative disincentive effect on information production dominates the positive information diffusion effect, which leads to worse returns for better connected mutual funds.

This paper is also broadly related to the study of social connections in other financial market settings. Social connections have been shown to be beneficial to firms and investors if they facilitate information sharing. Cohen, Frazzini, and Malloy (2008) show that mutual fund managers have education links with corporate board members gain significant information advantage. Engelberg, Gao, and Parsons (2012) find that firms that have social connections with their banks obtain loans with lower interest rates and fewer covenants. Hochberg, Ljungqvist, and Lu (2007) find that better-networked VC investors experience better fund performance. Cai and Sevilir (2012) find that social connections between board directors of target and acquirer firms lead to better merger performance. Huang, Jiang, Lie, and Yang (2014) find that acquirers with investment banker directors earn higher announcement returns, pay lower takeover premiums, and exhibit superior long-run performance. Stuart and Yim (2010) find that companies whose directors with private equity deal exposure (gained from interlocking directorships) are more likely to receive private equity offers. Engelberg, Gao, and Parsons (2013) find that CEOs with social connections to outsiders bring valuable information into the firm through these connections, and receive higher compensation. Bajo, Chemmanur, Simonyan, and Tehranian (2016) show that higher centrality of lead IPO underwriter in the underwriter network is associated with higher ability to induce a larger number of institutions to pay attention to the firm it takes public and to disseminate and extract information about the IPO firm from these institutions.

Finally, social connections have also been shown to have a potential negative effect on firms or investors. For example, Fracassi and Tate (2012) show that CEO-director connections weaken board monitoring and reduce firm value, particularly in the absence of other governance mechanisms to substitute for board oversight. Hwang and Kim (2009) find that board directors who are socially connected to the CEO are less efficient in monitoring and discipline the CEO. Ishii and Xuan (2014) find that social connections between target and acquirer firms lead to poorer decision making

positive relationship between connectedness and fund performance might be driven by that "investment consultants may choose particular fund managers because they like that manager's investment style and believe it fits well with a particular sponsor's overall set of managers", instead of an information diffusion effect.

and lower value creation for shareholders overall. Gompers, Mukharlyamov, and Xuan (2016) show that venture capitalists who share similar background are more likely to syndicate with each other and this homophily reduces the probability of investment success. Shue (2013) exploits the random assignment of MBA students to sessions at Harvard Business School and finds that executive compensation and acquisition strategy are significantly more similar among graduates from the same MBA session than among graduates from different sessions, and this may potentially lower firm productivity. Kuhnen (2009) finds that both "improved monitoring" and "increased potential for collusion" exist in the social connection between mutual fund advisors and boards. Duchin and Sosyura (2013) document that the social connections between CEOs and divisional managers increase (decrease) investment efficiency and firm value when information asymmetry is high (corporate governance is weak). In this paper, the negative effect of social connections arises not from weakened monitoring, but rather from weakened incentives for fund managers to produce independent information and inefficient contracting between fund managers and shareholders.

3 Data and Sample Construction

I obtain information on fund managers from Morningstar, who reports the name of each manager for a fund, their start and end dates with the fund, and information about the manager's educational background. I limit the sample to actively managed U.S. equity funds with Morningstar category in the 3 by 3 size/value grid (large growth, large blend, large value, medium growth, medium blend, medium value, small growth, small blend, small value). I remove index funds since their behavior is mechanically determined and is less likely to be influenced by information sharing through social connections.[8]

I obtain mutual fund monthly returns from the CRSP survivor-bias-free mutual fund database (matched using ticker symbol, cusip or fund name). I aggregate funds across fund classes into portfolios using Mutual Fund Links (MFLINKs) variable (WFICN). The number of funds in the sample grows from 1096 in January 1996 to 1709 in December 2010, with an average of 1824 funds per month. Additionally, I obtain holdings from Thomson Financial CDA/Spectrum Mutual Fund database, which contains the quarter-end holdings reported by US based mutual funds in mandatory

[8] I remove index funds by searching for the words "index", "idx", "S&P", "Dow Jones", and "NASDAQ" in the CRSP fund name.

SEC filings. I restrict holdings to common stocks traded in NYSE, NASDAQ or AMEX.

My goal is to identify pairs of managers who are connected socially and are more likely to engage in social communication regarding investment ideas. I do so by looking at their previous working experience in the mutual fund industry. I define indicator variable $Connected_{i,j,t}$, which equals to one if fund managers i and j worked in the same fund family as portfolio managers any time prior to the focal date. While I define here social connections using fund managers' prior working experience, I am aware there are alternative definitions of social connections in the literature (e.g. educational link in Cohen, Frazzini, and Malloy (2008), geographical proximity in Hong, Kubik, and Stein (2005) and Pool, Stoffman, and Yonker (2015)). Compared to other proxies of social connection using education background or geographical proximity, the experience of managing money in the same mutual fund family builds stronger social ties among fund managers, and increases probability of sharing and communicating investment ideas among themselves.

4 Social Connections and Mutual Fund Portfolios

4.1 Measuring overlap

Following Pool, Stoffman, and Yonker (2015), I measure the portfolio overlap in holdings between fund i and j during quarter t as

$$PortOverlap_{i,j,t} = \sum_{k \in H_t} \min\{w_{i,k,t}, w_{j,k,t}\} \tag{1}$$

where $w_{i,k,t}$ is fund i's portfolio weight in stock k at the end of calendar quarter t, and H_t is the set of all stocks held by funds i and j as reported at the end of calendar quarter t.

I also measure the overlap in stock purchases and sales between mutual funds. I define

$$BuyOverlap_{i,j,t} = \frac{\sum_{k \in T_t} \min\{I^+_{i,k,t}, I^+_{j,k,t}\}}{\min\{\sum_{k \in T_t} I^+_{i,k,t}, \sum_{k \in T_t} I^+_{j,k,t}\}} \tag{2}$$

$$SellOverlap_{i,j,t} = \frac{\sum_{k \in T_t} \min\{I^-_{i,k,t}, I^-_{j,k,t}\}}{\min\{\sum_{k \in T_t} I^-_{i,k,t}, \sum_{k \in T_t} I^-_{j,k,t}\}} \tag{3}$$

where $I^+_{i,k,t}$ is an indicator variable which equals to one if fund i increases its holding in stock k between quarter $t-1$ and t, and zero otherwise. $I^-_{i,k,t}$ equals to one if fund i decreases its holding in stock k between quarter $t-1$ and t, and zero otherwise. T_t is the union of all stock traded by funds i and j.

4.2 Summary Statistics of Fund Pairs

Table 1 presents the summary statistics of $Connected_{i,j,t}$, $PortOverlap_{i,j,t}$, $BuyOverlap_{i,j,t}$, $SellOverlap_{i,j,t}$, and other control variables used in my analysis. $SameCity_{i,j,t}$ is a dummy variable which equals to one if funds i and j are headquartered in the same city (using the mutual fund company address); $SameFamily_{i,j,t}$ equals to one if funds i and j are affiliated with the same mutual fund family; $CommonManager_{i,j,t}$ equals to one if funds i and j have at least one portfolio manager in common; $MngOtherFundTogether_{i,j,t}$ equals to one if at least one pair of portfolio managers from funds i and j managing at least one other fund together at quarter t. $SameMSGrid_{i,j,t}$ equals to one if both funds i and j belong to the same Morningstar size and value/growth grid. I also include as control variables a set of dummies that equal to one if funds i and j match on Morningstar size or value/growth categories (For example, $BothValue_{i,j,t}$ equals to one if both funds in the pair are classified as Value fund by Morningstar; $BothLargeCap_{i,j,t}$ equals to one if both funds in the pair are classified as Large-Cap fund by Morningstar). I also include the absolute value of the difference between the total net asset (TNA)-based quintiles of funds i and j ($TNAQuinDiff_{i,j,t}$) and the average TNA-based quintiles of funds i and j ($TNAQuinAvg_{i,j,t}$).

Table 1 tabulates the summary statistics for both connected fund pairs ($Connected_{i,j,t} = 1$) and unconnected fund pairs ($Connected_{i,j,t} = 0$). Unconditionally, I find that connected fund pairs have 2.02% higher overlap in portfolio holdings, 2.49% higher overlap in stock purchases, and 2.45% higher overlap in stock sales, compared to unconnected fund pairs. However, connected fund pairs are more likely to be located in the same city or be affiliated with the same fund family. Connected fund pairs are also more likely to have common fund manager or have the a pair of managers managing the same fund together. These confounding factors all contribute to the abnormal overlap in portfolio holdings and stock trades for connected fund pairs, hence I will carefully control for these variables in my multivariate analysis.

4.3 Overlap in Holdings and Trades

To test the hypothesis that connected mutual funds are more likely to make similar investments, I estimate the following regression

$$PortOverlap_{i,j,t} = \alpha + \beta Connected_{i,j,t} + \delta SameCity_{i,j,t} + \Gamma' Controls_{i,j,t} + \epsilon_{i,j,t} \tag{4}$$

My main variable of interest, $Connected_{i,j,t}$, is a dummy variable that equals to one if at least one pair of portfolio managers from funds i and j work as portfolio managers in the same fund at certain time point before quarter t. I conduct the analysis at the fund level (Pool, Stoffman, and Yonker (2015)) rather than at the stock level (Hong, Kubik, and Stein (2005)) as the latter approach involves billions of observations and the analysis is not computationally feasible. $Controls_{i,j,t}$ includes a list of controls discussed in the previous section.

Table 2 shows the coefficient estimates and t-statistics for various specifications of equation 4. Standard errors are two-way clustered at the fund level for each fund in the pair. The coefficient for $Connected_{i,j,t}$ is 1.03 in model (1) after controlling for a list of fund characteristics, implying additional 1.03% portfolio overlap for connected fund pairs, compared to unconnected fund pairs. To put this number into perspective, the same-city effect documented in Hong, Kubik, and Stein (2005) is estimated to be 54 basis points (coefficient for $SameCity_{i,j,t}$). In model (2), I exclude fund pairs which have at least one common portfolio manager; in model (4), I exclude fund pairs from the same family. The coefficient for $Connected_{i,j,t}$ is of similar statistical significance and economic magnitude in both cases. In column (6), I include only fund pairs when both the funds have only one portfolio manager. Compared to the team-managed mutual funds, single-manager funds are more likely to be influence by the social network of its sole portfolio manager. The empirical results in column (6) confirm my hypothesis. The coefficient for $Connected_{i,j,t}$ is 1.94, about 90% higher than the case where both types of funds are included in the sample.

In addition, I find that a pair of funds from the same family tends to hold similar stocks (documented in Elton, Gruber, and Green (2007)), and this effect is estimated to be 1.59% in my base model (1). In model (3), I limit the sample to pairs of funds from different families, and I find that the coefficient for $CommonManager_{i,j,t}$ is 9.69, reflecting the effect of a sub-advisor relationship on portfolio holdings. Specifically, a fund, sub-advised by a fund manager from a different family, is likely to have 9.69% more overlap in the portfolio holdings with another fund managed by the same manager than otherwise. Not surprisingly, the variables matching on the Morningstar size and value/growth categories have significant power for explaining the commonality between mutual fund holdings. Meanwhile, funds similar in size ($TNAQuinDiff_{i,j,t}$) and large-sized funds ($TNAQuinAvg_{i,j,t}$) tend to have more common holdings.

I next investigate whether fund pairs managed by socially connected portfolio managers are more likely to make similar trades than those managed by portfolio managers not socially connected. I use the $BuyOverlap_{i,j,t}$ and $SellOverlap_{i,j,t}$ measure defined earlier as the dependent variables and re-estimate the regression in equation 4. In Table 3, I estimate the regressions using three different specifications for both purchases and sales: the sample excluding fund pairs with common managers, the sample excluding fund pairs within the same fund family, and fund pairs where both funds are managed by a single manager.

The results are similar to the case of overlap in portfolio holdings. Socially connected mutual funds are more likely to make purchases and sales simultaneously (within the same quarter). In my baseline model (1) and (4), a pair of connected funds have 1.47% more overlap in purchases and 1.47% more overlap in stock sales than otherwise. In model (3) and (6), I found the effect of social connections is higher for stock sales than stock purchase. One possible explanation is that a negative signal shared by other fund managers may be more credible and the fund manager is more likely to trade on this negative signal within the same quarter.

4.4 Alternative Hypothesis: Manager Preferences

It is possible that some of the correlation I uncover between my portfolio overlap measures and the social connections between portfolio managers may be driven by unobserved characteristics of these managers(e.g. ethnicity, political affiliation), rather than by social connections. Indeed, the formation of social connections, as well as portfolio choices, could both be driven by a common set of unobserved manager characteristics.

I test this alternative hypothesis by exploiting the dynamics of social network of the portfolio managers over time. In this analysis, I limit the sample to fund pairs when both funds are managed by a single manager. I construct a new connection variable, $MgrConnectedFuture_{i,j,t}$, which equals to one if portfolio managers from fund i and j established a connection in the future. Assuming managers' preferences are stable over time, I expect $MgrConnectedFuture_{i,j,t}$ to be correlated with portfolio overlap between mutual funds, since the underlying unobserved manager characteristics drive both $MgrConnectedFuture_{i,j,t}$ and portfolio overlap measures. I put both $Connected_{i,j,t}$ and $MgrConnectedFuture_{i,j,t}$ as independent variables in the regression and run a horse-race test.

Table 4 presents the results for this test. In columns (1), the dependent variable is the overlap in holdings; in columns (2) and (3), the the dependent variable is the overlap in purchases and sales, respectively. Overall speaking, $Connected_{i,j,t}$ retains its explanatory power, in terms of the economic magnitude and statistical significance of the coefficient estimates, for various specifications of the overlap measure. Meanwhile, the coefficient estimate for $MgrConnectedFuture_{i,j,t}$ is small and insignificant when dependent variable is overlap in stock purchases, and the coefficient estimate for $MgrConnectedFuture_{i,j,t}$ is negative when the dependent variable is the overlap in stock holdings or stock sells. In conclusion, the results reported in Table 4 provide evidence against the hypothesis that the abnormal overlap in portfolio and trades between connected mutual funds is driven by unobserved managers' preferences.

5 Social Connections and Mutual Fund Performance

Chen, Jegadeesh, and Wermers (2000) show that active mutual fund managers possess superior private information regarding the stock they buy and sell. Kacperczyk and Seru (2007) find that highly skilled managers rely less on public information in their portfolio allocations. The origin of private information is multi-fold[9]. In the previous section, I show that social connections have an effect on the portfolio holdings and trades of mutual funds. However, it remains unclear ex-ante whether better connected mutual funds will have better or worse returns. On the one hand, better connected funds will have access to more signals, including their own signal and signals shared by their connected peers, and hold more precise information in aggregate. On the other hand, information production is costly, social connections may reduce funds' ex-ante incentives to devote more resources to produce more precise signals, since they can instead free-ride on their connected peers. In this section, I study whether social connections have an effect on mutual fund

[9] Coval and Moskowitz (1999) show that U.S. investment managers exhibit a strong preference for local firms. Cohen, Frazzini, and Malloy (2008) show that mutual fund managers place larger bets on connected firms and perform significantly better on these holdings relative to their nonconnected holdings.

performance, and specifically whether that effect is positive or negative.

5.1 Mutual Fund Centrality

To quantify each fund's social connections, I make use of the centrality measures first developed in social network analysis.[10] I compute common measures of centrality, including degree, eigenvector and closeness centrality, for my sample funds in monthly frequency. The degree centrality is defined as the number of links incident upon a node. Eigenvector centrality is a measure of the influence of a node in a network. It assigns relative scores to all nodes in the network based on the concept that connections to high-scoring nodes contribute more to the score of the node in question than equal connections to low-scoring nodes; The farness of a node is defined as the sum of its distances from all other nodes, and the closeness centrality is defined as the reciprocal of the farness. Empirically, all three measures of centrality are highly correlated.

I now discuss the potential concerns related to the definition of social connection I have chosen. First, I am aware that social connections based on prior careers may only constitute a subset of the entire space of social connections between fund managers. However, focusing on this particular type of social connections biases my tests against finding a significant relationship between the fund centrality measures and fund performance. Second, using prior career experiences does not necessarily mean that I completely ignore other forms of social connections. In fact, it is likely that a pair of connected fund managers (through their prior careers in the same fund family) are also likely to establish other forms of social connections (e.g. being a neighbor) and therefore my measures of centrality might have captured these other types of social connections.

Table 5 presents the summary statistics of these measures of centrality for the sample funds, as well other fund characteristics and manager characteristics, in monthly frequency. TNA_t is the total net assets of the fund (in millions). $FamilySize_t$ is the total net assets of the fund family (in millions). $NetFlow_t$ is defined as

$$NetFlow_t = \frac{TNA_t - TNA_{t-1}(1 + R_t)}{TNA_{t-1}} \tag{5}$$

where R_t is the net raw return of the fund. $TurnoverRatio_t$ is the minimum of aggregated sales

[10] Ozsoylev, Walden, Yavuz, and Bildik (2014) use centrality measures in studying the trading profits of all investors in Istanbul Stock Exchange in 2005.

or aggregated purchases of securities, divided by the average 12-month TNA of the fund. Age_t is the age of the fund since inception. $ManagerSAT_t$ is the median SAT of matriculants at the manager's undergraduate institution. $ManagerMBA_t$ is a dummy variable which equals to one if the manager has an MBA degree and zero otherwise. $ManagerTenure_t$ is the number of years that the manager has been managing the fund. $ManagerAge_t$ is the age of the manager. If the fund is managed by multiple managers, $ManagerSAT_t$, $ManaagerMBA_t$, $ManagerTenure_t$, and $ManagerAge_t$ are averaged at the fund level.

5.2 Determinants of Mutual Fund Centrality

In this section, I study the determinants of fund centrality using pooled panel regressions. Specifically, I regress measures of fund centrality ($EigenvectorCentrality_t$, $DegreeCentrality_t$, and $ClosenessCentrality_t$) on a list of fund characteristics ($Log(FundSize)_t$, $Log(FamilySize)_t$, $NumMgrs_t$, and $Log(Age+1)_t$) and manager characteristics ($ManagerSAT_t$, $ManaagerMBA_t$). Time fixed effects are included. Standard errors are clustered at the fund family level.

Table 6 presents the regression results for each of the three centrality measures. Importantly, $FamilySize$ and $NumMgrs$ are the two most significant determinants of the fund centrality measures. Large families tend to have more funds and hire more fund managers, which establishes more social connections according to my definition. Funds managed by more managers tend to have more connections with other funds. In addition, funds with managers from higher SAT undergraduate school and managers with MBA degree are more likely to have higher centrality measures.

Next, I study the cross-sectional difference in manager behavior and its link to fund centrality. I run a monthly rolling regression of fund gross return on Fama-French-Carhart four factors (MKT, SMB, HML, and UMD) using a 24-month lookback window. I obtain the estimates of the factor loadings including β_{MKT}, β_{SMB}, β_{HML}, and β_{UMD}. I regress these beta estimates on the eigenvector centrality measure. In addition, I also regress $Turnover_t$ and $ExpenseRatio_t$ on the eigenvector centrality measure.

Table 7 presents the regression results. I find funds with higher eigenvector centrality measure have higher loadings on market (MKT) and momentum factors (UMD). I also find funds with higher eigenvector centrality measure tend to hold more large-cap stocks and growth stocks. On the other hand, I do not find a significant relationship between eigenvector centrality measure and

$Turnover_t$ or $ExpenseRatio_t$.

5.3 Predictability of Centrality for Fund Performance

In this section, I test whether the centrality measures are able to predict fund performance adjusting for risk factors. I estimate the following regression

$$r_{i,t+1} = \alpha + \beta Centrality_{i,t} + \gamma X_{i,t} + \epsilon_{i,t+1} \tag{6}$$

where the dependent variable $r_{i,t+1}$ is fund i's monthly gross return or Fama-French-Carhart 4-factor alpha for month $t+1$. As in Fama and French (2010) and Cohen, Coval, and Pastor (2005), I use pre-expense returns to best capture fund manager's stock picking skills. Thus, I add 1/12-th of the annual expense ratio to the net returns reported in CRSP.[11] Fama-French-Carhart four-factor alpha is calculated with respect to the market, size, value and momentum factors following Carhart (1997). The factor loadings are estimated with a 24-month look-back period and I require at least 12 monthly returns. In the regression, I control for fund characteristics $X_{k,t}$, including $Log(FundSize)_t$ (fund size), $Log(FamilySize)_t$ (family size), $NumMgrs_t$ (team size), $NetFlow_t$, $NetFlow_t^2$ (liquidity cost), $TurnonverRatio_t$ (fund turnover), $Log(1 + Age)_t$ (age of the fund), and factor loadings β_{MKT}, β_{SMB}, β_{HML}, and β_{UMD}. I also control for manager characteristics variables including $ManagerSAT_t$ and $ManagerMBA_t$.

Table 8 summarizes the results of the Fama-MacBeth (1973) regressions with Newey-West(1987) adjusted (12 lags) standard errors. The eigenvector centrality is used throughout this section.[12] In model (1)-(5), the dependent variable is the gross return of the fund. While I do not find a statistically significant univariate relationship between the eigenvector centrality and fund gross return in model (1), the coefficient for $EigenvectorCentrality_t$ is negative and highly statistically significant in model (2), indicating that $EigenvectorCentrality_t$ has significant predictive power for fund performance, after controlling for fund's exposure to systematic risk factors and other fund characteristics. The coefficient for $EigenvectorCentrality_t$ is -1.424 in model (2), implying a one-standard-deviation increase in $EigenvectorCentrality_t$ predicts a decrease of 2.4 basis

[11] My results are robust if I use net returns instead of gross returns.

[12] While I primarily use eigenvector centrality in the empirical analysis, I also use degree and closeness centrality measures to verify the the results still hold.

points in monthly fund gross returns and 29.0 basis points in annualized fund gross returns. In model (3), I include additional control variables of manager characteristics. Consistent with findings in Chevalier and Ellison (1999), I find that both $ManagerSAT_t$ and $ManagerMBA_t$ have positive and statistically significant predictive power for fund returns. In addition, I find that $EigenvectorCentrality_t$ retains its predictive power and is of similar economic magnitude to that in model (3), suggesting that it is not the selection of manager quality that is driving my results. In model (4)-(6), I use Fama-French-Carhart 4-factor alpha as the dependent variable and find even stronger results. The coefficient for $EigenvectorCentrality_t$ is -1.772 in model (5), implying a one-standard-deviation increase in $EigenvectorCentrality_t$ predicts a decrease of 3.0 basis points in monthly fund Fama-French-Carhart 4-factor alpha and 36.1 basis points in annualized fund Fama-French-Carhart 4-factor alpha.

In Table 8, consistent with the findings in the literature, the control variables also show the right direction of predictability for fund performance. I find a negative and significant relationship between fund size and fund returns, and a positive and significant relationship between family size and fund returns. This result is consistent with the findings in Chen, Hong, Huang, and Kubik (2004), who argue that fund size erodes performance due to liquidity reasons and, controlling for fund size, belonging to a large family is beneficial for the fund return because of the economy of scale. I also find a significant and positive relationship between $NetFlow$ and fund return, which is consistent with the "smart-money" effect documented in Zheng (1999). Meanwhile, there exists a significant and negative relationship between $NetFlow^2$ and fund return, which is likely due to liquidity costs associated with flow. On the other hand, fund age and turnover play a secondary role in predicting fund returns.[13]

Next, I examine the predictability of eigenvector centrality measure for the fund's future alpha using the portfolio sort approach. For each calendar month, I sort funds into decile portfolios based on the eigenvector centrality measure unconditionally. Next, I calculate the equal-weighted Fama-French-Carhart 4-factor alpha over the next one month, three months, six months, and twelve

[13] In the Online Appendix, I run pooled regression with month fixed effects and fund fixed effects, and I find similar predictive power of eigenvector centrality for fund returns. I also show that superior performance of less connected funds is not driven by those managers taking a large "Active Share" (as documented in Cremers and Petajisto (2009)). In addition, I split the sample periods into pre-Reg FD period and post-Reg FD periods, and find no difference in predictive power of centrality for fund performance. Finally, I also scale the centrality measure by management team size and find equally strong results.

months after the portfolio formation date. The average returns of these portfolios are presented in Table 10. The 1-10 decile spread is the zero-investment long-short portfolio that is long on decile one and short on decile ten. I find that the eigenvector centrality measure is a reliable predictor of its future alpha with a 1-10 decile spread of 5.2 basis points for the 1-month horizon. The spread is 16.9 basis points at 3-month horizon, 35.4 basis points at 6-month horizon, and 73.1 basis points at 12-month horizon.[14]

In conclusion, in this section I show that there is a negative relationship between fund centrality and fund returns, i.e. better connected funds have less alphas compared to less connected funds. It suggests that the "free riding on friends" effect dominates the information diffusion effect (through social connections) in the information production decision of individual mutual funds. Fund managers in better connected funds devote less resources or efforts into information production, compared to fund mangers in less connected funds, and the extra signals they receive from their social connections are not sufficient to compensate for the loss of precision in the signals produced on their own. As a result, social connections demonstrate a negative effect on mutual fund performance.

5.4 Fixed Effects

Large fund families, e.g. Fidelity Investments, hire a large number of research analysts and support staffs to build up in-house information production capacity for all affiliated funds. In this section, first, I study whether the relationship between centrality and fund returns I uncovered in the previous section is only driven by the differences in performance between fund families. Specifically, I study whether the relationship between centrality and fund returns holds, even within the same family. Second, I want to study whether fund families internalize their managers' external social connections, as a source of information, when allocating resources into internal research. Specifically, I study whether the relationship between centrality and fund returns holds across different families.

Empirically, I add family fixed effects to the regression model, and specifically I estimate the

[14]In the Online Appendix, I additionally make sure there is no overlap in returns between different period for the same portfolio. For instance, I re-balance the portfolio every quarter if the portfolio return is calculated over a 3-month horizon. I show that the 1-10 decile spreads have similar point estimates, and are statistically significant at the 10% level.

following "within family" and "between family" predictive power of the eigenvector centrality measure,

$$r_{i,t+1} - \bar{r}_{j,t+1} = \alpha + \beta(EigenvectorCentrality_{i,t} - \overline{EigenvectorCentrality}_{j,t}) + \gamma(X_{i,t} - \bar{X}_{j,t}) + \epsilon_{i,t+1} \tag{7}$$

$$\bar{r}_{j,t+1} = \alpha + \beta\overline{EigenvectorCentrality}_{j,t} + \gamma\bar{X}_{j,t} + \epsilon_{j,t+1} \tag{8}$$

where $\bar{r}_{j,t+1}$ represents the cross-section average of $r_{i,t+1}$ (fund i is affiliated with family j) for family j during period $t+1$. Meanwhile, $\overline{EigenvectorCentrality}_{j,t}$ and $\bar{X}_{j,t}$ also represent the family average of their corresponding variable during period t. Table 9 presents the results for both "within family" and "between family" in columns (1) and (2). The coefficient estimate is -0.877 for the "within family" estimator and -2.240 for the "between family" estimator, and both coefficient estimates are statistically significant at 5% level. This suggests that the predictive power of eigenvector centrality for future fund returns exists both within family and between families, and is stronger between families. The result of the "within family" estimator in column (1) implies that managers' social connections matter for their own information production, and consequently their fund performance, even within the same mutual fund family. The result of the "between family" estimator in column (2) suggests that fund families internalize managers' external social connections, as a source of information, when deciding how much to invest in internal research.

It is also interesting to study whether the relationship between centrality and fund returns is driven by the differences in performance between mutual funds located in different geographical locations. In columns (3) and (4) of Table 9, I similarly study the "within city" and "between city" predictive power of the eigenvector centrality measure. The coefficient estimate is -1.970 for the "within city" estimator and -1.546 for the "between city" estimator, and both coefficient estimates are statistically significant at 5% level. The "within city" estimator in column (3) suggests that the relationship between centrality and fund returns holds, even within the same city. The "within city" estimator is even larger than the "between city" estimator in terms of both economic magnitude and statistical significance. The predictive power of the eigenvector centrality measure exists both within city and between cities.

5.5 Alternative Hypothesis: Frequent Job Switchers

It is documented in the literature that there is an inverse relationship between fund manager turnover and lagged fund performance (e.g., Kostovetsky and Warner (2015)). Hence, managers' centrality may be endogenous to their stock-picking skills through turnovers. More specifically, managers with low stock picking skills are more likely to be fired and switch jobs across fund families, and thereby establish more "connections" in the fund industry. Hence, fund centrality could be correlated with the past performance of its manager(s), and the finding of an inverse relationship between mutual fund centrality and future fund performance in section 5.3 could be driven by the persistence of bad performance of "Frequent Job Switchers".

To address this endogeneity concern, I adopt three empirical tests. First, I include management tenure, $Log(ManagerTenure+1)$, as an additional control variable in my baseline regression. The result is presented in columns (1) and (3) of Table 11. The coefficient for $Log(ManagerTenure+1)$ is not statistically significant and does not affect the predictive power of $EigenvectorCentrality$ for fund performance. The weakness of this test is that management tenure only reflects the length of current employment relationship and does not fully capture the historical performance of the fund manger being considered.

In the second test, I directly measure the historical performance of each fund manager. I rank Fama-French-Carhart 4-factor alpha of every fund in monthly frequency and assign a percentile value (higher percentile, better performance). I average the percentile ranking value for all funds managed by every manager in my sample. I construct a new variable $MgrPerformanceHist$ as the cumulative average of manager's past 4-factor percentile rankings. If a fund has multiple managers, I average $MgrPerformanceHist$ equally across managers in the fund level. Lagged $MgrPerformanceHist$ is negatively correlated with fund centrality measures, confirming my conjecture that managers with bad performance are more likely to be fired, switch jobs, and therefore establish more connections. If my main results are driven by these "Frequent Job Switchers", I will expect centrality has no predictive power for fund performance after controlling for $MgrPerformanceHist$. The empirical result is presented in columns (2) and (4) of Table 11. The coefficient for $MgrPerformanceHist$ is positive and highly statistically significant, suggesting past performance of managers predicts future returns of the fund. Rejecting the "Frequent Job

Switchers" hypothesis, I find that eigenvector centrality retains its predictive power even with the presence of $MgrPerformanceHist$ as a control variable. The results hold when I use either gross fund return or Fama-French-Carhart 4-factor alpha as the dependent variable.

In the third test, I decompose the fund centrality measure into "In" and "Out" components based on the direction of social connections. The direction of connection is determined by whether the manager joins a new fund family ("Out" connection) or whether the other party joins from another family ("In" connection). In these new centrality measures, "Frequent Job Switchers" are likely to have many "Out" connections and little "In" connections, while managers with long tenure in the family are likely to have many "In" connections and little "Out" connections. Based on the "In" and "Out" connections, I calculate two sets of eigenvector and degree centrality measures, and label them as $EigenvectorCentrality(In)$ and $DegreeCentrality(In)$ (I refer to them as "in centrality"), and $EigenvectorCentrality(Out)$ and $DegreeCentrality(Out)$ (I refer to them as "out centrality").

If the inverse relationship between fund centrality measures and future performance is entirely driven by the persistence of bad performance of "Frequent Job Switchers", there should exist an inverse relationship between "out centrality" measures and future fund performance, and simultaneously no relationship between "in centrality" measures and future fund performance. The empirical results are presented in Table 12. In columns (1)-(4), I find that both "in centrality" measures and "out centrality" measures negatively predict future fund alpha performance and the coefficient estimates are statistically significant and are of similar economic magnitude as my baseline results. I do find, however, the coefficient estimates for "out centrality" measures are weaker in terms of economic magnitude. This indicates that while my results are not fully explained by the "Frequent Job Switchers", the presence of "Frequent Job Switchers" does contribute to the worse performance of funds they are managing.

5.6 Fund Flows and Centrality

Previous studies document that outsider investors chase past fund performance when allocating their wealth (e.g. Chevalier and Ellison (1997)). The response of flow to performance indicates that investors learn from past returns about managers' stock picking abilities (Berk and Green (2004)). In this section, I study whether mutual fund centrality directly affects flows of money into the funds, and also whether mutual fund centrality affects the flow-performance relationship.

To examine the two effects empirically, I estimate the following panel regression:

$$\begin{aligned} NetFlow_{i,t} &= \alpha + \beta_0 EigenvectorCentrality_{i,t-1} + \beta_1 Return_{t-1} \\ &+ \beta_2 Return_{t-1} \times EigenvectorCentrality_{i,t-1} + \gamma X_{i,t-1} + \epsilon_{i,t} \end{aligned} \tag{9}$$

For the lagged return performance measure, I use both raw returns R_{t-1} and Fama-French-Carhart 4-factor alpha α^{4f}_{t-1}. Following the existing literature, I control for fund-specific characteristics such as log of fund size, family size, log of fund age, expenses ratio, and turnover. I estimate this panel regression using pooled regressions with month fixed effects. Robust standard errors, reported in parentheses, are two-way clustered in family and month levels.

I report the empirical results in Table 13. In columns (1) and (4), I reproduce results documented in the literature: Fund flows from outside investors chase past performance, and the flow-performance relationship is robust using both raw returns and Fama-French-Carhart 4-factor alphas. The significant negative coefficient on the standard deviation of lagged fund performance ($ReturnVol_{t-1}$) suggests that investors care about risk. In columns (2) and (5), I find that eigenvector centrality measure is negatively correlated with future fund flows. In addition, the interaction term between eigenvector centrality measure and past performance ($\alpha^{4f}_{t-1} \times EigenvectorCentrality_{t-1}$ or $R_{t-1} \times EigenvectorCentrality_{t-1}$) is negative and statistically significant, indicating that the flow-performance relationship is stronger for less connected, comparing to better connected funds. The effect of centrality on the flow-performance relationship is also economically significant. In column (2), the coefficient for $\alpha^{4f}_{t-1} \times EigenvectorCentrality_{t-1}$ is -1.251, implying that a two-standard-deviation difference in eigenvector centrality corresponds to a difference of 4.3% in the flow-performance relationship, which is 18.9%[15] of the unconditional flow-performance relationship.

Chevalier and Ellison (1997) document that flows for younger funds are more sensitive to past performance than older funds. To control for the effect of fund age on the flow-performance relationship, I further add an interaction term, $\alpha^{4f}_{t-1} \times Log(Age+1)_{t-1}$ and $R_{t-1} \times Log(Age+1)_{t-1}$, in columns (3) and (6) respectively. I find the coefficients for $\alpha^{4f}_{t-1} \times EigenvectorCentrality_{t-1}$ and

[15] Calculated as follows: 4.3%/22.6% = 18.9%.

$R_{t-1} \times EigenvectorCentrality_{t-1}$ remain statistically significant and retain similar economic magnitude, suggesting that the effect of centrality on the flow-performance relationship is not driven by fund age.

Taken together, the results in this section show that mutual funds with lower centrality are able to attract larger money inflows. In addition, investors' flow seems to be more responsive to the past performance of mutual funds with lower centrality. This is consistent with the results in section 5.3 where I find "free riding on friends" effect of social connection dominates the information diffusion effect. In this case, mutual funds with lower centrality produce more precise signals, and past returns of these mutual funds are stronger signal about the stock picking abilities of their managers, compared to mutual funds with higher centrality.

6 Mechanism: Stock-level Evidence

In the previous section, I show that a higher centrality for mutual funds predicts worse future fund alphas. I interpret the findings as that managers from less connected funds devote more efforts into producing more precise information, and this overcomes the disadvantages that they do not receive as much information from social connections as managers from better connected funds. If this is true, less connected funds should aggregately have more precise information than better connected funds. In this section, I explore the information content of stock holdings of mutual fund investors.[16] More specifically, I study whether the holdings of less connected funds are more informed about stocks' future abnormal returns and earnings-related fundamentals, compared to those of better connected funds.

6.1 Central and Peripheral Funds

In each quarter t, I classify fund i with above median eigenvector centrality as *central* fund and below median eigenvector centrality as *peripheral* fund. The average portfolio weights for *central*

[16]Following modern portfolio theory, a mutual fund manager's portfolio holdings are the outcome of an optimization based on his specific beliefs about stock expected returns and the covariance structure of these returns. Shumway, Szefler, and Yuan (2011) propose a method to extract the information embedded in the cross-sectional portfolio holdings for fund managers' beliefs. Other papers investigating the information revealed by portfolio holdings of mutual funds include Chen, Jegadeesh, and Wermers (2000), Cohen, Coval, and Pastor (2005), Kacperczyk and Seru (2007), Cremers and Petajisto (2009), and Jiang, Verbeek, and Wang (2014).

funds and *peripheral* funds are represented as $CTR_{k,t}$ and $PER_{k,t}$, respectively.[17] I construct a PMC measure, which is defined as the difference in average portfolio weights between *peripheral* funds and *central* funds,

$$PMC_{k,t} = \frac{PER_{k,t} - CTR_{k,t}}{2} \tag{10}$$

I also use variable $ALL_{k,t}$ to represent the average portfolio weights of stock k for all funds in the sample.

Table 14 presents summary statistics for the variables used in my analysis. $\Delta BREADTH_t$ is the change in breadth of ownership from the end of quarter $t-1$ to quarter t[18]. ΔIO_t is the change in fraction of shares outstanding of a stock held by 13F institutions from the end of quarter $t-1$ to quarter t. $LOGSIZE_t$ is the natural logarithm of market capitalization at the end of quarter t. BK/MKT_t is the most recently available observation of book-to-market ratio at the end of quarter t. $MOM12_t$ is the raw stock return for the last 12 months excluding the recent one month. XTR_t is the quarterly share turnover (volume normalized by number of shares outstanding) adjusted for the average share turnover of the firm's exchange.

Panel A of Table 14 shows the summary statistics for each size quintiles (size quintiles are determined using NYSE breakpoints), as well as the total. Size quintile 1 includes the smallest cap stocks and size quintile 5 has the largest cap stocks. The average portfolio weight for a stock is 41 basis points (of the fund's total net assets). On average, large-cap stocks have larger average portfolio weights across mutual funds, compared to small-cap stocks. Interestingly, PMC_t is positive across each size quintile, which suggests that *peripheral* funds hold larger, and more concentrated position in a typical stock, compared to *central* funds. It reflects the superior stock picking skills of *peripheral* funds, and their information advantage in a particular stock they invest in. The alternative theory is that *central* funds are typically large funds and they are refrained from taking a large position in a particular stock due to liquidity constraints and price impact (Chen, Hong, Huang, and Kubik (2004)). However, if this "liquidity hypothesis" is true, PMC_t is ought to be more positive in small-cap stocks where liquidity constraint is more close to be binding,

[17] I use the average portfolio weight of mutual funds instead of the fractional holdings (as a percentage of total shares outstanding) to rule out the possibility that a few large funds are driving the results.

[18] I follow Lehavy and Sloan (2008) to construct $\Delta BREADTH_t$ using 13F data.

compared to large-cap stocks. In fact, there is no monotonic relationship between PMC and size.

Panel B of Table 14 show the contemporaneous correlations between these variables. $MOM12_t$ is highly positively correlated with ALL_t, suggesting that average mutual funds tend to hold and purchase past winners (as documented in Wermers (1999)). $MOM12_t$ is also highly positively correlated with $\Delta BREADTH_t$ and ΔIO_t, suggesting average 13F institutions are also engaged in momentum trading strategies. PER_t is highly correlated with CTR_t with average correlation about 53%. Therefore, I am primarily focused on the PMC_t variable in studying the relative information advantage held by *peripheral* funds over *central* funds. PMC_t is weakly correlated with ALL_t. Also, PMC_t, is only weakly correlated with the other control variables.

6.2 Forecast Stock Returns

In the baseline test, I estimate the following two equivalent regression models,

$$r_{k,t+j-1,t+j} = \alpha + \beta_1^1 PER_{k,t} + \beta_2^1 CTR_{k,t} + \gamma X_{k,t} + \epsilon_{k,t,t+1} \tag{11}$$

$$r_{k,t+j-1,t+j} = \alpha + \beta_1^2 PMC_{k,t} + \beta_2^2 ALL_{k,t} + \gamma X_{k,t} + \epsilon_{k,t,t+1} \tag{12}$$

where the independent variable $r_{k,t+j-1,t+j}$, $j = 1, 2, 3, 4$, is stock i's cumulative returns (raw returns or risk-adjusted returns) from the end of quarter $t+j-1$ to the end of quarter $t+j$. My main variable of interest, $PMC_{k,t}$, reflects the private information advantage of *peripheral* funds over *central* funds regarding the future stock return. The control variables $X_{k,t}$ represent public available information including $\Delta IO_{k,t}$, $\Delta BREADTH_{k,t}$, and $XTR_{k,t}$, which are known in the literature to have predictive power for stock returns in cross-section.

In Table 15, I present the results of a series of Fama-MacBeth (1973) regressions forecasting stock returns over the first, second, third and fourth quarter following the formation date. I run cross-sectional regression every quarter and report the mean coefficients across different specifications. The standard errors are adjusted for serial correlation and heteroscedasticity following Newey-West (1987) with four lags. There are three groups of regressions in Table 15. The first group corresponds to forecasting raw cumulative returns. The second quarter uses Fama-French-Carhart four-factor alpha over the same horizon as the dependent variable. The third group uses DGTW-adjusted returns as the dependent variable instead.[19] In each group, two sets of regression models are used: one uses $PER_{k,t}$ and $CTR_{k,t}$, and the other one uses $PMC_{k,t}$ and $ALL_{k,t}$.

In Panel A of Table 15, the coefficient for PER is positive and significant, while the coefficient for CTR is negative and significant. The results imply that the average portfolio weight by *peripheral* funds is a positive predictor for stock returns and the average portfolio weight by *central* funds is a negative predictor for stock returns. Since PER and CTR is high correlated, I focus my discussion around my main variable of interest, PMC. The coefficient for PMC is positive and highly statistically significant across three different specifications of cumulative raw and abnormal return measures. To get a sense of the economic magnitude, the coefficient estimate 2.126 for PMC in model (2) implies that a one-standard-deviation increase in PMC predicts an increase of 37 basis points in the next quarter cumulative return (1.48% on an annualized basis). Similarly, the coefficient estimate 2.216 for PMC in model (4) implies that a one-standard-deviation increase in PMC predicts an increase of 38 basis points in the next quarter cumulative Fama-French-Carhart 4-factor alpha (1.52% on an annualized basis). Interestingly, the coefficient for ALL is small and not statistically significant across all three return specifications, suggesting that the average portfolio weight for all funds does not contain incremental information for predicting stock returns.[20]

The coefficient for PMC is also positive and statistically significant in Panel B and Panel C of Table 15, suggesting that the predictive power of PMC for cumulative stock returns persists until the second and the third quarter after the formation date. However, the predictive power of PMC disappears when forecasting cumulative stock returns for the fourth quarter after the formation date (as seen in Panel D of Table 15).

In conclusion, the results in Table 15 suggest that the average portfolio weight of *peripheral* funds have superior forecasting power than that of *central* funds. I show that this forecasting power is statistically and economically significant, and it lasts up to three quarters after the formation date. In addition, there is no reversal of the relationship between PMC and stock abnormal returns after the third quarter, suggesting that PMC proxies for an information advantage by *peripheral* over

[19] I create portfolio benchmarks using a characteristics-based procedure similar to Daniel, Grinblatt, Titman, and Wermers (1997). The DGTW benchmarks are available via http://www.smith.umd.edu/faculty/rwermers/ftpsite/Dgtw/coverpage.htm

[20] In the Online Appendix, I also show that the predictive power of my PMC measure for abnormal stock returns holds after excluding small stocks (lowest NYSE quntile), or funds' local holdings (the firm is within 50 miles from the fund family's headquarter).

central funds, and that information is gradually impounded into stock prices through the portfolio rebalancing by these funds.

6.3 Forecast Earnings Surprises

Baker, Litov, Wachter, and Wurgler (2010) find that mutual fund trades forecast earnings surprises and they conclude that mutual fund managers are able to trade profitably in part because they are able to forecast earnings-related fundamentals. Given the evidence of superior forecasting power of the PMC measure in forecasting future stock abnormal returns, it is natural to turn to the question whether it is due to an ability to forecast fundamental news not yet release into the public market or, say, proprietary technical signals. In this section, I will test whether the holdings of *peripheral* funds are able to predict earnings surprises better, compared to those of *central* funds.

Similar to the previous section, I estimate the following two equivalent regression models,

$$SUE_{k,t+j} = \alpha + \beta_1^1 PER_{k,t} + \beta_2^1 CTR_{k,t} + \gamma X_{k,t} + \epsilon_{k,t,t+1} \tag{13}$$

$$SUE_{k,t+j} = \alpha + \beta_1^2 PMC_{k,t} + \beta_2^2 ALL_{k,t} + \gamma X_{k,t} + \epsilon_{k,t,t+1} \tag{14}$$

where $SUE_{k,t+j}, j = 1, 2$ is the earnings announcement surprise of earnings announced between the end of quarter $t + j - 1$ and the end of quarter $t + j$. I define the SUE (standardized unexpected earning) as follows,

$$SUE_{k,t+j} = \frac{EPS^A_{k,t+j} - EPS^E_{k,t}}{P_t} \tag{15}$$

where $EPS^A_{k,t+j}$ is actual announced earnings during quarter $t+j$ for stock k. EPS^E_t is the median of I/B/E/S analysts forecasts for stock k at the end quarter t for the earnings to be announced in the future quarter $t + j$.

Table 16 presents the results of Fama-MacBeth (1973) regressions of model 13. The coefficient for PER is positive and significant, while the coefficient for CTR is negative and significant. My main variable of interest, PMC, is positive and statistically significant with p-value less than 0.001 for the standard earnings surprises based on the first quarter and second quarter earnings announcement after the formation date. In terms of economic magnitude, a one-standard-deviation increase in PMC predicts an increase of approximately 20 basis points in SUE for the first quarter after the formation date, and 19 basis points for the second quarter after the formation date. The findings in this section complements the evidence in Baker, Litov, Wachter, and Wurgler (2010), where they shown mutual fund managers as a group have forecasting abilities for earnings-related fundamentals.

In conclusion, I am able to show that *peripheral* funds have information advantage over *central* funds in terms of forecasting the earnings announcement surprises for the first and second quarter after the formation date. However, keep in mind this test only partially explores the source of information advantage of *peripheral* funds. As argued by Baker, Litov, Wachter, and Wurgler (2010), this approach is complementary to tests using long-horizon returns.

7 Conclusion

In this paper, I build a proxy for social connections between mutual funds through career experiences of their fund managers in the mutual fund industry. I find that connected funds are more likely to hold similar stocks and make same-direction trades, compared to unconnected funds. This result confirms the findings in the literature that the portfolio choices of institutional investors are affected by the social connections among their managers.

My paper takes a step further by showing that social connections among investment managers dampen their incentives to produce independent signals, and thereby managers of peripheral funds collectively hold more precise signals than managers of central funds. I show that funds with higher centrality earn less returns/alphas. Further, I empirically construct a PMC variable that approximates the relative information advantage of peripheral funds over central funds, and I find PMC has significant predictive power for future stock abnormal returns and earnings surprises. My results contrast with the findings in the literature that information diffusion through social connections is beneficial for the information precision of managers since they have access to more signals and hold more precise information collectively. This could be reconciled under the theoretical framework of Han and Yang (2013) where they discuss two opposite effects of social connections, i.e. the information diffusion effect and "free riding on friends" effect. My empirical study identifies a setting (using a sample of mutual fund managers and their career experience as proxy for social connections) where the "free riding on friends" effect dominates the information diffusion effect.[21] The implication for investors in mutual fund is that controlling for fund characteristics and manager characteristics, fund managers' social connections carry additional information that is relevant for the future performance of the fund. And my test regarding the flow-performance relationship suggests that investors, rationally, are more responsive to the past performance of less connected fund or fund managers.

Notably, the finding that funds with lower network centrality have better returns/alphas is not a direct implication from the model of Han and Yang (2013). In fact, under their rational expectations equilibrium framework, mutual funds should earn equal investment returns after the cost of information production. There are two possible explanations. First, managers in funds with lower centrality devote extra effort producing more precise signals and incur higher information production costs. The true information production cost is unobserved, hence the difference in alphas, between funds with high and low centrality, simply reflects the difference in true information production costs; Or, there are certain forms of inefficiencies associated with the incentive contracts of fund managers (e.g. career risk from taking unique risky investment positions) that refrain managers from devoting optimal efforts into information production. However, exactly identifying these inefficiencies (or agency issues) in the mutual fund industry falls beyond the scope of this paper, and may be of interest to the readers for future research.

[21] It is possible that in the cases of social connections based on education or geographical proximity, the sharing of investment ideas between fund managers is more likely to be sporadic, and managers may not internalize the effect of social connections when making information production decisions. On the other hand, in the case when two fund managers previously work together in the same fund family, they may share and communicate investment ideas or strategies systematically, and consequently it is more likely that they internalize the effect of social connections, as a source of information, when making information production decisions.

References

Bajo, Emanuele, Thomas J. Chemmanur, Karen Simonyan, and Hassan Tehranian, 2016, Underwriter networks, investor attention, and initial public offerings, *Journal of Financial Economics* 122, 376–408.

Baker, Malcolm, Lubomir Litov, Jessica A. Wachter, and Jeffrey Wurgler, 2010, Can Mutual Fund Managers Pick Stocks? Evidence from Their Trades Prior to Earnings Announcements, *Journal of Financial and Quantitative Analysis* 45, 1111–1131.

Berk, Jonathan B., and Richard C. Green, 2004, Mutual Fund Flows and Performance in Rational Markets, *Journal of Political Economy* 112, 1269–1295.

Cai, Ye, and Merih Sevilir, 2012, Board connections and M&A transactions, *Journal of Financial Economics* 103, 327–349.

Carhart, Mark M., 1997, On Persistence in Mutual Fund Performance, *The Journal of Finance* 52, 57–82.

Chen, Hsiu-Lang, Narasimhan Jegadeesh, and Russ Wermers, 2000, The Value of Active Mutual Fund Management: An Examination of the Stockholdings and Trades of Fund Managers, *Journal of Financial and Quantitative Analysis* 35, 343–368.

Chen, Joseph, Harrison Hong, Ming Huang, and Jeffrey D. Kubik, 2004, Does fund size erode mutual fund performance? the role of liquidity and organization, *American Economic Review* 94, 1276–1302.

Chevalier, Judith, and Glenn Ellison, 1997, Risk Taking by Mutual Funds as a Response to Incentives, *Journal of Political Economy* 105, 1167–1200.

——— , 1999, Are Some Mutual Fund Managers Better Than Others? Cross-Sectional Patterns in Behavior and Performance, *The Journal of Finance* 54, 875–899.

Cohen, Lauren, Andrea Frazzini, and Christopher Malloy, 2008, The Small World of Investing: Board Connections and Mutual Fund Returns, *Journal of Political Economy* 116, 951–979.

Cohen, Randolph B., Joshua D. Coval, and Lubos Pastor, 2005, Judging Fund Managers by the Company They Keep, *The Journal of Finance* 60, 1057–1096.

Coval, Joshua D., and Tobias J. Moskowitz, 1999, Home Bias at Home: Local Equity Preference in Domestic Portfolios, *The Journal of Finance* 54, 2045–2073.

Cremers, K. J. Martijn, and Antti Petajisto, 2009, How Active Is Your Fund Manager? A New Measure That Predicts Performance, *Review of Financial Studies* 22, 3329–3365.

Daniel, Kent, Mark Grinblatt, Sheridan Titman, and Russ Wermers, 1997, Measuring Mutual Fund Performance with Characteristic-Based Benchmarks, *The Journal of Finance* 52, 1035–1058.

DeMarzo, Peter M., Dimitri Vayanos, and Jeffrey Zwiebel, 2003, Persuasion Bias, Social Influence, and Unidimensional Opinions, *The Quarterly Journal of Economics* 118, 909–968.

Duchin, Ran, and Denis Sosyura, 2013, Divisional Managers and Internal Capital Markets, *The Journal of Finance* 68, 387–429.

Elton, Edwin J., Martin J. Gruber, and T. Clifton Green, 2007, The Impact of Mutual Fund Family Membership on Investor Risk, *Journal of Financial and Quantitative Analysis* 42, 257–277.

Engelberg, Joseph, Pengjie Gao, and Christopher A. Parsons, 2012, Friends with money, *Journal of Financial Economics* 103, 169–188.

——— , 2013, The Price of a CEO's Rolodex, *Review of Financial Studies* 26, 79–114.

Fama, Eugene F., and Kenneth R. French, 2010, Luck versus Skill in the Cross-Section of Mutual Fund Returns, *The Journal of Finance* 65, 1915–1947.

Fracassi, Cesare, and Geoffrey Tate, 2012, External Networking and Internal Firm Governance, *The Journal of Finance* 67, 153–194.

Gerritzen, Marc, Jens Carsten Jackwerth, and Alberto Plazzi, 2016, Birds of a Feather - Do Hedge Fund Managers Flock Together?, SSRN Scholarly Paper ID 2729993 Social Science Research Network Rochester, NY.

Gompers, Paul A., Vladimir Mukharlyamov, and Yuhai Xuan, 2016, The cost of friendship, *Journal of Financial Economics*.

Gray, Wesley R., Steve Crawford, and Andrew E. Kern, 2012, Do Fund Managers Identify and Share Profitable Ideas?, SSRN Scholarly Paper Social Science Research Network Rochester, NY.

Grossman, Sanford J., and Joseph E. Stiglitz, 1980, On the impossibility of informationally efficient markets, *American Economic Review* 70, 393–408.

Han, Bing, and Liyan Yang, 2013, Social Networks, Information Acquisition, and Asset Prices, *Management Science* 59, 1444–1457.

Hochberg, Yael V., Alexander Ljungqvist, and Yang Lu, 2007, Whom You Know Matters: Venture Capital Networks and Investment Performance, *The Journal of Finance* 62, 251–301.

Hong, Harrison, Jeffrey D. Kubik, and Jeremy C. Stein, 2005, Thy Neighbor's Portfolio: Word-of-Mouth Effects in the Holdings and Trades of Money Managers, *The Journal of Finance* 60, 2801–2824.

Huang, Qianqian, Feng Jiang, Erik Lie, and Ke Yang, 2014, The role of investment banker directors in M&A, *Journal of Financial Economics* 112, 269–286.

Hvide, Hans K., and Per Östberg, 2015, Social interaction at work, *Journal of Financial Economics* 117, 628–652.

Hwang, Byoung-Hyoun, and Seoyoung Kim, 2009, It pays to have friends, *Journal of Financial Economics* 93, 138–158.

Ishii, Joy, and Yuhai Xuan, 2014, Acquirer-target social ties and merger outcomes, *Journal of Financial Economics* 112, 344–363.

Ivković, Zoran, and Scott Weisbenner, 2007, Information Diffusion Effects in Individual Investors' Common Stock Purchases: Covet Thy Neighbors' Investment Choices, *Review of Financial Studies* 20, 1327–1357.

Janis, Irving L., 1982, *Groupthink: Psychological Studies of Policy Decisions and Fiascos* (Houghton Mifflin Company: Boston,MA).

Jiang, Hao, Marno Verbeek, and Yu Wang, 2014, Information Content When Mutual Funds Deviate from Benchmarks, *Management Science* 60, 2038–2053.

Kacperczyk, Marcin, and Amit Seru, 2007, Fund Manager Use of Public Information: New Evidence on Managerial Skills, *The Journal of Finance* 62, 485–528.

Kostovetsky, Leonard, and Jerold B. Warner, 2015, Youre Fired! New Evidence on Portfolio Manager Turnover and Performance, *Journal of Financial and Quantitative Analysis* 50, 729–755.

Kuhnen, Camelia M., 2009, Business Networks, Corporate Governance, and Contracting in the Mutual Fund Industry, *The Journal of Finance* 64, 2185–2220.

Lehavy, Reuven, and Richard G. Sloan, 2008, Investor recognition and stock returns, *Review of Accounting Studies* 13, 327–361.

Manski, Charles F., 1993, Identification of Endogenous Social Effects: The Reflection Problem, *The Review of Economic Studies* 60, 531–542.

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

Pool, Veronika K., Noah Stoffman, and Scott E. Yonker, 2015, The People in Your Neighborhood: Social Interactions and Mutual Fund Portfolios, *The Journal of Finance* 70, 2679–2732.

Rossi, Alberto G., David P. Blake, Allan G. Timmermann, Ian Tonks, and Russ Wermers, 2016, Network Centrality and Delegated Investment Performance, SSRN Scholarly Paper ID 2743187 Social Science Research Network Rochester, NY.

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

Shue, Kelly, 2013, Executive Networks and Firm Policies: Evidence from the Random Assignment of MBA Peers, *Rev. Financ. Stud.* 26, 1401–1442.

Shumway, Tyler, Maciej Szefler, and Kathy Yuan, 2011, The Information Content of Revealed Beliefs in Portfolio Holdings, *Working paper.*

Spilker, Harold, 2016, Hedge fund family ties, *Working paper.*

Stein, Jeremy C., 2008, Conversations among competitors, *American Economic Review* 98, 2150–62.

Stuart, Toby E., and Soojin Yim, 2010, Board interlocks and the propensity to be targeted in private equity transactions, *Journal of Financial Economics* 97, 174–189.

Wermers, Russ, 1999, Mutual Fund Herding and the Impact on Stock Prices, *The Journal of Finance* 54, 581–622.

Zheng, Lu, 1999, Is Money Smart? A Study of Mutual Fund Investors' Fund Selection Ability, *The Journal of Finance* 54, 901–933.

Table 1: **Summary Statistics of Sample Fund Pairs**

The sample includes actively managed U.S. equity mutual funds between 1996 and 2010 (I restrict the samples to those with Morningstar category in the 3 by 3 size/value grid). $Connected_{i,j,t}$ equals to one if fund managers i and j worked in the same fund family as portfolio managers any time prior to quarter t. $PortOverlap_{i,j,t}$ measures the portfolio overlap in holdings (in percentage) between funds i and j during quarter t. $BuyOverlap_{i,j,t}$ measures the overlap in stock purchases (in percentage) between fund i and j during quarter t. $SellOverlap_{i,j,t}$ measures the overlap in stock sales (in percentage) between fund i and j during quarter t. $SameCity_{i,j,t}$ is a dummy variable which equals to one if funds i and j are headquartered in the same city (using the mutual fund company address); $SameFamily_{i,j,t}$ equals to one if funds i and j are affiliated with the same mutual fund family; $CommonManager_{i,j,t}$ equals one if funds i and j have at least one portfolio manager in common; $MngOtherFundTogether_{i,j,t}$ equals to one if at least one pair of portfolio managers from funds i and j managing at least one other fund together at quarter t. $SameMSGrid_{i,j,t}$ equals to one if both funds i and j belong to the same Morningstar size and value/growth grid. In addition, We include as control variables a set of dummies that equal to one if funds i and j match on Morningstar size or value/growth categories (For example, $BothValue_{i,j,t}$ equals to one if both funds in the pair are classified as Value funds by Morningstar; $BothLargeCap_{i,j,t}$ equals to one if both funds in the pair are classified as Large-Cap funds by Morningstar). We also include the absolute value of the difference between the total net asset (TNA)-based quintiles of funds i and j ($TNAQuinDiff_{i,j,t}$) and the average TNA-based quintiles of funds i and j ($TNAQuinAvg_{i,j,t}$).

	$Connected_{i,j,t} = 0$			$Connected_{i,j,t} = 1$			Total		
	Mean	Std.	N (thousands)	Mean	Std.	N (thousands)	Mean	Std.	N (thousands)
$PortOverlap_{i,j,t}(\%)$	7.39	8.45	56,117	9.41	10.26	5,871	7.58	8.66	61,988
$BuyOverlap_{i,j,t}(\%)$	7.99	12.93	56,117	10.48	14.66	5,871	8.23	13.12	61,988
$SellOverlap_{i,j,t}(\%)$	7.50	12.59	56,117	9.95	14.05	5,871	7.73	12.76	61,988
$Connected_{i,j,t}$	0.000	0.000	56,117	1.000	0.000	5,871	0.095	0.293	61,988
$SameCity_{i,j,t}$	0.051	0.221	56,117	0.112	0.315	5,871	0.057	0.232	61,988
$SameFamily_{i,j,t}$	0.001	0.038	56,117	0.076	0.264	5,871	0.008	0.092	61,988
$CommonManager_{i,j,t}$	0.001	0.035	56,117	0.018	0.134	5,871	0.003	0.053	61,988
$MngOtherFundTogether_{i,j,t}$	0.000	0.012	56,117	0.045	0.207	5,871	0.004	0.066	61,988
$SameMSGrid_{i,j,t}$	0.203	0.402	56,117	0.212	0.409	5,871	0.204	0.403	61,988
$BothBlend_{i,j,t}$	0.092	0.289	56,117	0.076	0.265	5,871	0.091	0.287	61,988
$BothValue_{i,j,t}$	0.057	0.232	56,117	0.078	0.268	5,871	0.059	0.236	61,988
$BothGrowth_{i,j,t}$	0.228	0.420	56,117	0.223	0.416	5,871	0.228	0.420	61,988
$BothLargeCap_{i,j,t}$	0.442	0.497	56,117	0.490	0.500	5,871	0.447	0.497	61,988
$BothMidCap_{i,j,t}$	0.049	0.215	56,117	0.034	0.181	5,871	0.047	0.212	61,988
$BothSmallCap_{i,j,t}$	0.068	0.253	56,117	0.066	0.248	5,871	0.068	0.252	61,988
$TNAQuinDiff_{i,j,t}$	1.61	1.21	56,117	1.51	1.17	5,871	1.60	1.21	61,988
$TNAQuinAvg_{i,j,t}$	2.03	1.00	56,117	2.29	0.98	5,871	2.05	1.00	61,988

Table 2: **Social Connections and Overlap in Mutual Fund Portfolio Holdings**

This table presents the OLS regression analysis of the the effect of social connections on mutual fund portfolio holdings. The dependent variable is $PortOverlap_{i,j,t}$, which measures the portfolio overlap in holdings (in percentage) between funds i and j during quarter t. The sample includes 62 million mutual fund pairs between 1996 and 2010. In column (1), the sample excludes fund pairs with common portfolio managers during quarter t. In column (2), the sample is limited to fund pairs from different mutual fund families. In column (3), the sample is restricted to fund pairs where both funds have only a single portfolio manager. $Connected_{i,j,t}$ equals to one if fund managers i and j worked in the same fund family as portfolio managers any time prior to quarter t. $SameCity_{i,j,t}$ is a dummy variable which equals to one if funds i and j are headquartered in the same city (using the mutual fund company address); $SameFamily_{i,j,t}$ equals to one if funds i and j are affiliated with the same mutual fund family; $CommonManager_{i,j,t}$ equals one if funds i and j have at least one portfolio manager in common; $MngOtherFundTogether_{i,j,t}$ equals to one if at least one pair of portfolio managers from funds i and j managing at least one other fund together at quarter t. $SameMSGrid_{i,j,t}$ equals to one if both funds i and j belong to the same Morningstar size and value/growth grid. In addition, We include as control variables a set of dummies that equal to one if funds i and j match on Morningstar size or value/growth categories (For example, $BothValue_{i,j,t}$ equals to one if both funds in the pair are classified as Value funds by Morningstar; $BothLargeCap_{i,j,t}$ equals to one if both funds in the pair are classified as Large-Cap funds by Morningstar). We also include the absolute value of the difference between the total net asset (TNA)-based quintiles of funds i and j ($TNAQuinDiff_{i,j,t}$) and the average TNA-based quintiles of funds i and j ($TNAQuinAvg_{i,j,t}$). Standard errors are two-way clustered by each fund in the pair. t-statistics are in parentheses. ***, **, and * indicate significance at the 1, 5, and 10 percent levels, respectively.

	Dependent Variable: $PortOverlap_{i,j,t}(\%)$				
	Full Sample	Full Sample	No Common Managers	Different Families	Funds with Single Manager
	(1)	(2)	(3)	(4)	(5)
$Connected_{i,j,t}$	1.03***		1.07***	1.08***	1.94***
	(7.68)		(7.96)	(8.04)	(7.28)
$SameCity_{i,j,t}$	0.54***	0.56***	0.49***	0.43***	0.30
	(4.17)	(4.32)	(3.79)	(3.28)	(1.59)
$SameFamily_{i,j,t}$	1.59***	2.29***	1.06***		3.01***
	(7.54)	(12.03)	(5.27)		(7.28)
$CommonManager_{i,j,t}$	11.70***	11.70***		9.69***	9.61***
	(16.11)	(16.26)		(12.62)	(9.74)
$MngOtherFundTogether_{i,j,t}$	1.03***	1.83***	0.25	0.94***	0.83*
	(3.69)	(6.65)	(1.12)	(3.42)	(1.72)
$SameMSGrid_{i,j,t}$	2.62***	2.62***	2.57***	2.60***	2.39***

	(23.49)	(23.47)	(23.13)	(23.35)	(14.34)
$BothValue_{i,j,t}$	0.69***	0.72***	0.69***	0.69***	0.04
	(2.97)	(3.07)	(2.97)	(2.98)	(0.10)
$BothGrowth_{i,j,t}$	1.20***	1.20***	1.20***	1.19***	1.05***
	(11.99)	(11.97)	(11.97)	(11.88)	(6.90)
$BothBlend_{i,j,t}$	0.68***	0.66***	0.69***	0.67***	0.67**
	(3.15)	(3.08)	(3.20)	(3.13)	(2.18)
$BothLargeCap_{i,j,t}$	9.32***	9.33***	9.32***	9.30***	8.84***
	(48.40)	(48.46)	(48.41)	(48.31)	(31.39)
$BothMidCap_{i,j,t}$	0.76***	0.75***	0.73***	0.74***	0.15
	(5.14)	(5.01)	(4.99)	(4.99)	(0.79)
$BothSmallCap_{i,j,t}$	0.03	0.03	0.00	0.01	-0.49***
	(0.20)	(0.23)	(0.04)	(0.06)	(2.70)
$TNAQuinDiff_{i,j,t}$	-0.16***	-0.17***	-0.16***	-0.16***	-0.20***
	(6.93)	(7.10)	(6.94)	(6.96)	(5.50)
$TNAQuinAvg_{i,j,t}$	0.48***	0.50***	0.48***	0.48***	0.59***
	(7.59)	(7.87)	(7.50)	(7.54)	(6.23)
Adjusted R^2	0.366	0.365	0.364	0.364	0.335
N(thousands)	61,988	61,988	61,812	61,462	10,878

Table 3: **Social Connections and Overlap in Mutual Fund Trades**

This table presents the OLS regressional analysis of the effect of social connections on mutual fund trades (stock purchases and sales). The sample includes 62 million mutual fund pairs between 1996 and 2010. In columns (1) (2) and (3), the dependent variable is $BuyOverlap_{i,j,t}$, which measures the overlap in stock purchases (in percentage) between funds i and j during quarter t. In columns (4) (5) and (6), the dependent variable is $SellOverlap_{i,j,t}$, which measures the overlap in stock sales (in percentage) between funds i and j during quarter t. In column (1) and (4), the sample excludes fund pairs with common portfolio managers during quarter t. In columns (2) and (5), the sample is limited to fund pairs from different mutual fund families. In columns (3) and (6), the sample is restricted to fund pairs where both funds have only a single portfolio manager. $Connected_{i,j,t}$ equals to one if fund managers i and j worked in the same fund family as portfolio managers any time prior to quarter t. $SameCity_{i,j,t}$ is a dummy variable which equals to one if funds i and j are headquartered in the same city (using the mutual fund company address); $SameFamily_{i,j,t}$ equals to one if funds i and j are affiliated with the same mutual fund family; $CommonManager_{i,j,t}$ equals one if funds i and j have at least one portfolio manager in common; $MngOtherFundTogether_{i,j,t}$ equals to one if at least one pair of portfolio managers from funds i and j managing at least one other fund together at quarter t. $SameMSGrid_{i,j,t}$ equals to one if both funds i and j belong to the same Morningstar size and value/growth grid. In addition, We include as control variables a set of dummies that equal to one if funds i and j match on Morningstar size or value/growth categories (For example, $BothValue_{i,j,t}$ equals to one if both funds in the pair are classified as Value funds by Morningstar; $BothLargeCap_{i,j,t}$ equals to one if both funds in the pair are classified as Large-Cap funds by Morningstar). We also include the absolute value of the difference between the total net asset (TNA)-based quintiles of funds i and j ($TNAQuinDiff_{i,j,t}$) and the average TNA-based quintiles of funds i and j ($TNAQuinAvg_{i,j,t}$). Standard errors are two-way clustered by each fund in the pair. t-statistics are in parentheses. ***, **, and * indicate significance at the 1, 5, and 10 percent levels, respectively.

	$BuyOverlap_{i,j,t}$(%)			$SellOverlap_{i,j,t}$(%)		
	No Common Managers (1)	Different Families (2)	Funds with Single Manager (3)	No Common Managers (4)	Different Families (5)	Funds with Single Manager (6)
$Connected_{i,j,t}$	1.47***	1.50***	1.29***	1.47***	1.50***	1.59***
	(8.23)	(8.35)	(4.42)	(8.71)	(8.93)	(6.01)
$SameCity_{i,j,t}$	-0.00	-0.11	0.08	0.55***	0.49***	0.67***
	(0.02)	(0.64)	(0.30)	(3.46)	(3.00)	(3.82)
$CommonManager_{i,j,t}$		10.45***	8.55***		9.60***	8.61***
		(10.32)	(7.38)		(11.91)	(9.13)
$SameFamily_{i,j,t}$	0.71***		4.26***	0.66***		3.17***
	(2.73)		(9.24)	(2.66)		(8.51)
$MngOtherFundTogether_{i,j,t}$	2.42***	3.60***	0.81	1.05***	2.02***	0.46
	(4.68)	(6.58)	(1.32)	(3.08)	(5.18)	(1.02)
$BothValue_{i,j,t}$	0.63**	0.64**	0.03	-0.33	-0.32	-0.11
	(2.47)	(2.49)	(0.09)	(1.43)	(1.40)	(0.35)

$BothGrowth_{i,j,t}$	0.80***	0.78***	0.64***	1.18***	1.16***	1.04***
	(5.48)	(5.34)	(2.92)	(9.09)	(8.98)	(6.43)
$BothBlend_{i,j,t}$	1.41***	1.40***	1.41***	0.98***	0.97***	0.61**
	(5.56)	(5.51)	(3.61)	(4.65)	(4.60)	(2.37)
$BothLargeCap_{i,j,t}$	7.53***	7.51***	6.84***	6.97***	6.94***	6.20***
	(37.32)	(37.13)	(22.74)	(37.55)	(37.41)	(24.83)
$BothMidCap_{i,j,t}$	0.87***	0.89***	0.60***	0.60***	0.61***	0.39**
	(5.27)	(5.32)	(2.76)	(4.03)	(4.07)	(2.19)
$BothSmallCap_{i,j,t}$	1.40***	1.42***	0.62	0.80***	0.81***	-0.32
	(4.82)	(4.85)	(1.43)	(3.75)	(3.78)	(1.38)
$SameMSGrid_{i,j,t}$	1.60***	1.63***	1.52***	1.42***	1.45***	1.26***
	(17.08)	(17.33)	(10.52)	(16.37)	(16.63)	(10.11)
$TNAQuinDiff_{i,j,t}$	0.07**	0.07**	0.05	-0.01	-0.01	-0.04
	(2.10)	(2.11)	(1.10)	(0.35)	(0.31)	(1.05)
$TNAQuinAvg_{i,j,t}$	0.86***	0.87***	0.96***	1.10***	1.10***	1.12***
	(9.20)	(9.27)	(7.01)	(12.66)	(12.74)	(11.68)
Adjusted R^2	0.102	0.104	0.094	0.097	0.098	0.097
N(thousands)	61,812	61,462	10,878	61,812	61,462	10,878

Table 4: **Social Connections and Overlap in Mutual Fund Portfolios: Falsification Test**

This table presents the OLS regressioni analysis of the effect of social connections on mutual fund portfolios (holdings, purchases and sales). The sample includes 10 million mutual fund pairs between 1996 and 2010 that both funds are managed by a single fund manager. $Connected_{i,j,t}$ equals to one if fund managers i and j worked in the same fund family as portfolio managers any time prior to quarter t. $MgrConnectedFuture_{i,j,t}$ equals to one if fund managers from fund i and j are connected at least four quarters after the focal quarter t. $PortOverlap_{i,j,t}$ measures the portfolio overlap in holdings (in percentage) between funds i and j during quarter t. $BuyOverlap_{i,j,t}$ measures the overlap in stock purchases (in percentage) between funds i and j during quarter t. $SellOverlap_{i,j,t}$ measures the overlap in stock sales (in percentage) between fund i and j during quarter t. $SameCity_{i,j,t}$ is a dummy variable which equals to one if funds i and j are headquartered in the same city (using the mutual fund company address); $SameFamily_{i,j,t}$ equals to one if funds i and j are affiliated with the same mutual fund family; $CommonManager_{i,j,t}$ equals one if funds i and j have at least one portfolio manager in common; $MngOtherFundTogether_{i,j,t}$ equals to one if at least one pair of portfolio managers from funds i and j managing at least one other fund together at quarter t. $SameMSGrid_{i,j,t}$ equals to one if both funds i and j belong to the same Morningstar size and value/growth grid. In addition, We include as control variables a set of dummies that equal to one if funds i and j match on Morningstar size or value/growth categories (For example, $BothValue_{i,j,t}$ equals to one if both funds in the pair are classified as Value funds by Morningstar; $BothLargeCap_{i,j,t}$ equals to one if both funds in the pair are classified as Large-Cap funds by Morningstar). We also include the absolute value of the difference between the total net asset (TNA)-based quintiles of funds i and j ($TNAQuinDiff_{i,j,t}$) and the average TNA-based quintiles of funds i and j ($TNAQuinAvg_{i,j,t}$). Standard errors are two-way clustered by each fund in the pair. t-statistics are in parentheses. ***, **, and * indicate significance at the 1, 5, and 10 percent levels, respectively.

	$PortOverlap_{i,j,t}$(%)	$BuyOverlap_{i,j,t}$(%)	$SellOverlap_{i,j,t}$(%)
	(1)	(2)	(3)
$Connected_{i,j,t}$	2.12***	1.17***	1.73***
	(8.43)	(4.26)	(6.78)
$MgrConnectedFuture_{i,j,t}$	-0.40**	0.28	-0.32*
	(2.29)	(1.25)	(1.87)
$SameCity_{i,j,t}$	0.30	0.07	0.67***
	(1.60)	(0.30)	(3.82)
$SameFamily_{i,j,t}$	3.01***	4.26***	3.17***
	(7.28)	(9.23)	(8.51)
$CommonManager_{i,j,t}$	9.60***	8.56***	8.60***
	(9.72)	(7.38)	(9.11)
$MngOtherFundTogether_{i,j,t}$	0.87*	0.78	0.49
	(1.78)	(1.26)	(1.09)
$BothValue_{i,j,t}$	0.04	0.03	-0.11
	(0.11)	(0.08)	(0.34)

$BothGrowth_{i,j,t}$	1.06***	0.64***	1.05***
	(6.91)	(2.92)	(6.43)
$BothBlend_{i,j,t}$	0.67**	1.41***	0.61**
	(2.17)	(3.61)	(2.36)
$BothLargeCap_{i,j,t}$	8.84***	6.83***	6.20***
	(31.38)	(22.70)	(24.83)
$BothMidCap_{i,j,t}$	0.15	0.61***	0.39**
	(0.77)	(2.77)	(2.18)
$BothSmallCap_{i,j,t}$	-0.48***	0.62	-0.31
	(2.69)	(1.43)	(1.37)
$SameMSGrid_{i,j,t}$	2.39***	1.52***	1.26***
	(14.33)	(10.52)	(10.10)
$TNAQuinDiff_{i,j,t}$	-0.20***	0.05	-0.04
	(5.50)	(1.11)	(1.05)
$TNAQuinAvg_{i,j,t}$	0.60***	0.96***	1.12***
	(6.26)	(6.99)	(11.72)
Adjusted R^2	0.335	0.094	0.097
N(thousands)	10,878	10,878	10,878

Table 5: **Summary Statistics**

The sample includes actively managed U.S. equity mutual funds between 1996 and 2010 (I restrict the sample to those with Morningstar category in the 3 by 3 size/value grid and having monthly return information in CRSP survivor-bias-free mutual fund database). The network centrality measures, $EigenvectorCentrality_t$, $DegreeCentrality_t$, and $ClosenessCentrality_t$, are calculated each month between 1996 and 2010 for all the fund samples in that month. $FundSize_t$ is the total net assets of the fund (in millions). $FamilySize_t$ is the total net assets of the all active equity funds in the fund family (in millions). $NetFlow_t$ is calculated as $NetFlow_t = \frac{TNA_t - TNA_{t-1}(1+R_t)}{TNA_{t-1}}$, where R_t is the net raw return of the fund during month t. $NumMgrs_t$ is the number of managers managing the fund during month t. $FundAge_t$ is the age of the fund since inception. $TurnoverRatio_t$ is the minimum of aggregated sales or aggregated purchases of securities, divided by the average 12-month Total Net Assets of the fund. $ManagerSAT_t$ is the median SAT of matriculants at the manager's undergraduate institution. $ManagerMBA_t$ is a dummy variable which equals to one if the manager has an MBA degree and zero otherwise. $ManagerTenure_t$ is the number of years that the manager has been managing the fund. $ManagerAge_t$ is the age of the manager. If the fund is managed by multiple managers, $ManagerSAT_t$, $ManagerMBA_t$, $ManagerTenure_t$, and $ManagerAge_t$ are averaged at the fund level. R_t is the net return of the fund. α_t^{4f} is the Fama-French-Carhart four-factor alpha (factor loadings are calculated using monthly fund returns of prior 36 months). β_{MKT}, β_{SMB}, β_{HML}, and β_{UMD} are estimates from monthly rolling regressions of gross fund returns on Fama-French-Carhart four factors (MKT, SMB, HML, and UMD) using a 36-month window.

	Mean	Std.	Median	10th	90th
$EigenvectorCentrality_t$	0.016	0.017	0.015	0.000	0.036
$DegreeCentrality_t$	0.096	0.138	0.062	0.000	0.222
$ClosenessCentrality_t$	0.402	0.199	0.478	0.000	0.548
TNA_t	798	3593	92	5	1415
$FamilySize_t$	18473	55079	2892	45	39846
$NumMgrs_t$	2	2	2	1	4
$FundAge_t$	11	12	7	1	23
$NetFlow_t$	0.01	0.09	-0.00	-0.04	0.06
$TurnoverRatio_t$	0.94	1.20	0.68	0.19	1.83
$ExpenseRatio_t$	0.013	0.016	0.012	0.008	0.018
$ManagerSAT_t$	1242	121	1240	1086	1410
$ManagerMBA_t$	0.50	0.41	0.50	0.00	1.00
$ManagerTenure_t$	4.43	4.55	3.22	0.67	9.50
$ManagerAge_t$	48	9	47	38	60
$R_t(GrossReturn, \%)$	0.70	5.77	1.13	-6.41	7.10
$\alpha_t^{4f}(GrossReturn, \%)$	0.02	2.32	0.00	-2.25	2.30
$R_t(NetReturn, \%)$	0.60	5.77	1.04	-6.52	7.00
$\alpha_t^{4f}(NetReturn, \%)$	-0.08	2.32	-0.10	-2.36	2.19
$\beta_{MKT,t}$	1.00	0.21	0.99	0.78	1.22
$\beta_{SMB,t}$	0.23	0.39	0.13	-0.19	0.79
$\beta_{HML,t}$	0.03	0.38	0.04	-0.44	0.48
$\beta_{UMD,t}$	0.03	0.21	0.01	-0.19	0.27

Table 6: **Determinants of Mutual Fund Centrality**

This table presents the results for the pooled panel regression for the determinants of mutual fund centrality measures. The dependent variables are monthly fund network centrality measures including $EigenvectorCentrality_t$, $DegreeCentrality_t$, and $ClosenessCentrality_t$. $Log(FundSize)_t$ is the natural logarithm of total net assets of the fund. $Log(FamilySize)_t$ is the natural logarithm of total net assets of all active equity funds in the fund family. $NumMgrs_t$ is the number of managers managing the fund during month t. Age_t is the age of the fund since inception. $ManagerSAT_t$ is the median SAT of matriculants at the manager's undergraduate institution. $ManagerMBA_t$ is a dummy variable which equals to one if the manager has an MBA degree and zero otherwise. $ManagerTenure_t$ is the number of years that the manager has been managing the fund. $ManagerAge_t$ is the age of the manager. If the fund is managed by multiple managers, $ManagerSAT_t$, $ManagerMBA_t$, $ManagerTenure_t$, and $ManagerAge_t$ are averaged at the fund level. Month fixed effects are included. Standard errors are clustered at the fund family level. t-statistics are reported in parentheses. ***, **, and * indicate significance at the 1, 5, and 10 percent levels, respectively.

	(1) $EigenvectorCentrality_t$	(2) $DegreeCentrality_t$	(3) $ClosenessCentrality_t$
$Log(TNA)_t$	0.000*	0.003**	-0.006***
	(1.67)	(2.43)	(3.04)
$Log(FamilySize)_t$	0.002***	0.007***	0.033***
	(7.00)	(5.25)	(13.01)
$NumMgrs_t$	0.002***	0.016***	0.014***
	(12.25)	(13.78)	(7.76)
$Log(Age+1)_t$	-0.001***	-0.009***	0.003
	(3.08)	(3.41)	(0.91)
$ManagerSAT_t$	0.001***	0.005***	0.007*
	(2.71)	(2.98)	(1.91)
$ManagerMBA_t$	0.003***	0.022***	0.052***
	(5.01)	(5.07)	(5.35)
$Log(ManagerTenure+1)_t$	-0.001***	-0.007**	-0.029***
	(2.81)	(2.30)	(4.89)
$Log(ManagerAge)_t$	-0.000	0.004	-0.059***
	(0.05)	(0.36)	(2.65)
Month Fixed Effects	Yes	Yes	Yes
Adjusted R^2	0.215	0.264	0.357
No. of observations	259,903	259,903	259,903

Table 7: **Fund Characteristics and Fund Centrality**

This table presents the results for the pooled panel regression for the relationship between fund characteristics and the eigenvector centrality measure. The dependent variables are β_{MKT}, β_{SMB}, β_{HML}, β_{UMD}, $Turnover_t$, and $ExpenseRatio_t$. $Log(FundSize)_t$ is the natural logarithm of total net assets of the fund. $Log(FamilySize)_t$ is the natural logarithm of total net assets of all active equity funds in the fund family. $NumMgrs_t$ is the number of managers managing the fund during month t. Age_t is the age of the fund since inception. $ManagerSAT_t$ is the median SAT of matriculants at the manager's undergraduate institution. $ManagerMBA_t$ is a dummy variable which equals to one if the manager has an MBA degree and zero otherwise. $ManagerTenure_t$ is the number of years that the manager has been managing the fund. $ManagerAge_t$ is the age of the manager. If the fund is managed by multiple managers, $ManagerSAT_t$, $ManagerMBA_t$, $ManagerTenure_t$, and $ManagerAge_t$ are averaged at the fund level. Month fixed effects are included. Standard errors are clustered at the fund family level. t-statistics are reported in parentheses. ***, **, and * indicate significance at the 1, 5, and 10 percent levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
	$\beta_{MKT,t}$	$\beta_{SMB,t}$	$\beta_{HML,t}$	$\beta_{UMD,t}$	$TurnoverRatio_t$	$ExpenseRatio_t$
$EigenvectorCentrality_t$	1.138***	-0.751*	-0.761*	0.582***	0.358	-0.003
	(6.05)	(1.81)	(1.70)	(2.67)	(0.21)	(0.43)
$ManagerSAT_t$	0.006***	0.005	-0.004	-0.003	0.012	-0.000
	(2.65)	(0.88)	(0.69)	(0.76)	(0.59)	(0.69)
$ManagerMBA_t$	0.011	0.003	0.033*	-0.008	-0.083	-0.000
	(1.63)	(0.16)	(1.89)	(0.98)	(1.35)	(0.85)
$Log(ManagerTenure+1)_t$	-0.022***	-0.020**	0.031***	-0.018***	-0.258***	0.000
	(5.18)	(2.19)	(3.81)	(4.10)	(7.55)	(1.13)
$Log(TNA)_t$						-0.001***
						(5.80)
$Log(FamilySize)_t$						-0.000***
						(3.43)
Month Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R^2	0.060	0.012	0.079	0.053	0.030	0.061
No. of observations	276,951	276,951	276,951	276,951	271,399	277,746

Table 8: **Predicting Mutual Fund Returns with Centrality: Fama-MacBeth Regression**
The sample includes actively managed U.S. equity mutual funds between 1996 and 2010 (I restrict the samples to those with Morningstar category in the 3 by 3 size/value grid and having monthly return information in CRSP survivor-bias-free mutual fund database). The dependent variables are monthly fund gross return and Fama-French-Carhart four-factor alpha (factor loadings are calculated using monthly fund returns of prior 36 months). $Log(TNA)_t$ is the natural logarithm of total net assets of the fund. $Log(FamilySize)_t$ is the natural logarithm of total net assets of all active equity funds in the fund family.$NumMgrs_t$ is the number of managers managing the fund during month t. $NetFlow_t$ is calculated as $NetFlow_t = \frac{TNA_t - TNA_{t-1}(1+R_t)}{TNA_{t-1}}$, where R_t is the net return of the fund during month t. $TurnoverRatio_t$ is the minimum of aggregated sales or aggregated purchases of securities, divided by the average 12-month Total Net Assets of the fund. Age_t is the age of the fund since inception. $ManagerSAT_t$ is the median SAT of matriculants at the manager's undergraduate institution. $ManagerMBA_t$ is a dummy variable which equals to one if the manager has an MBA degree and zero otherwise. β_{MKT}, β_{SMB}, β_{HML}, and β_{UMD} are estimates from monthly rolling regressions of gross fund returns on Fama-French-Carhart four factors (MKT, SMB, HML, and UMD) using a 24-month window. Coefficients of Fama-MacBeth (1973) regressions are reported. t-statistics, which are in parentheses, are adjusted (12 monthly lags) for serial correlation and heteroscedasticity following Newey-West (1987). ***, **, and * indicate significance at the 1, 5, and 10 percent levels, respectively.

	Gross Return			Fama-French-Carhart Alpha		
	(1)	(2)	(3)	(4)	(5)	(6))
$EigenvectorCentrality_t$	-1.335	-1.424***	-1.361***	-1.411***	-1.772***	-1.798***
	(1.39)	(3.70)	(2.99)	(2.77)	(4.65)	(3.33)
$Log(TNA)_t$		-0.013**	-0.015***		-0.015**	-0.016***
		(2.50)	(2.95)		(2.56)	(2.93)
$Log(FamilySize)_t$		0.015***	0.010***		0.011***	0.007*
		(5.01)	(3.16)		(3.21)	(1.93)
$NumMgrs_t$		0.002	0.000		0.003	0.003
		(0.61)	(0.19)		(0.99)	(0.82)
$Log(Age+1)_t$		-0.021*	-0.019		-0.019*	-0.017
		(1.83)	(1.55)		(1.74)	(1.44)
$NetFlow_t$		0.716***	0.836***		0.714***	0.821***
		(3.12)	(4.05)		(2.80)	(3.32)

$NetFlow_t^2$		-0.946**	-0.903**		-0.930**	-0.909**
		(2.40)	(2.58)		(2.07)	(2.16)
$TurnoverRatio_t$		0.019	0.025		0.013	0.009
		(0.89)	(1.13)		(0.43)	(0.27)
$\beta_{MKT,t}$		0.153	0.143			
		(0.41)	(0.39)			
$\beta_{SMB,t}$		0.300	0.285			
		(1.30)	(1.23)			
$\beta_{HML,t}$		0.189	0.170			
		(0.64)	(0.57)			
$\beta_{UMD,t}$		0.024	-0.004			
		(0.06)	(0.01)			
$ManagerSAT_t$			0.021***			0.020***
			(4.09)			(3.33)
$ManagerMBA_t$			0.048***			0.036**
			(4.04)			(2.59)
Average R^2	0.003	0.415	0.424	0.002	0.029	0.031
No. of months	180	180	180	180	180	180
No. of observations	327,222	290,573	266,657	304,009	291,649	267,617

Table 9: **Predicting Mutual Fund Returns with Centrality: Fixed Effects**
In this table, I study the predictive power of centrality for fund returns with family fixed effects and city fixed effects. The dependent variable is Fama-French-Carhart 4-factor alpha for all regressions. Specifically, I estimate between estimator and within estimator for family fixed effects and city fixed effects separately. In columns (1) and (2), I keep only fund samples if its affiliated family has at least two funds in each calendar month. In column (1) "within family" estimation, all variables are demeaned at the fund family level for each calendar month. In column (2) "between family" estimation, all variables are family level averages for each calendar month. In columns (3) and (4), I keep only fund samples if its affiliated city is in the top 100 according to total net assets in the calendar month. In column (3) "within city" estimation, all variables are demeaned at the city level for each calendar month. In column (4) "between city" estimation, all variables are city level averages for each calendar month. Coefficients of Fama-MacBeth (1973) regressions are reported. t-statistics, which are in parentheses, are adjusted (12 monthly lags) for serial correlation and heteroscedasticity following Newey-West (1987). ***, **, and * indicate significance at the 1, 5, and 10 percent levels, respectively.

	Dependent Variable: Fama-French-Carhart 4-factor alpha			
	Within Family	Between Family	Within City	Between City
	(1)	(2)	(3)	(4)
$EigenvectorCentrality_t$	-0.877**	-2.214***	-1.970***	-1.546**
	(2.14)	(3.57)	(4.21)	(2.11)
$Log(TNA)_t$	-0.030***	0.007	-0.017***	0.001
	(3.79)	(0.69)	(2.97)	(0.06)
$Log(FamilySize)_t$		0.005	0.013***	0.004
		(0.84)	(3.71)	(0.43)
$NumMgrs_t$	0.002	0.001	0.006*	-0.016**
	(0.49)	(0.18)	(1.89)	(1.98)
$Log(Age + 1)_t$	0.004	-0.020	-0.017	-0.014
	(0.31)	(0.71)	(1.44)	(0.47)
$NetFlow_t$	0.618**	0.861**	0.775***	1.185**
	(2.27)	(2.24)	(2.95)	(2.08)
$NetFlow_t^2$	-0.625	-1.807**	-1.200***	-0.558
	(1.40)	(2.52)	(2.61)	(0.39)
$TurnoverRatio_t$	0.038	0.002	0.027	-0.025
	(1.14)	(0.03)	(0.88)	(0.63)
Average R^2	0.025	0.083	0.030	0.135
No. of months	180	180	180	180
No. of observations	257,431	41,727	272,610	17,808

Table 10: **Predicting Mutual Fund Returns with Centrality: Portfolio Sorts**
This table reports future Fama-French-Carhart four factor alpha for 10 deciles of the past centrality measure. At the start of each calendar month, we sort funds into decile portfolios based on the eigenvector centrality measure at the end of last month. Next, we calculate equal-weighted Fama-French-Carhart four factor alpha over the next one month, three months, six months, and twelve months after portfolio formation. The 1-10 decile spread is the zero-investment long-short portfolio that is long on decile one and short on decile ten. t-statistics, which are in parentheses, are adjusted (12 monthly lags) for serial correlation and heteroscedasticity following Newey-West (1987). ***, **, and * indicate significance at the 1, 5, and 10 percent levels, respectively.

Decile	1 Month	3 Months	6 Months	12 Months
1(Low)	0.057	0.178*	0.428**	0.869***
	(1.56)	(1.71)	(2.09)	(2.59)
2	0.035	0.142	0.351	0.854
	(0.82)	(1.13)	(1.31)	(1.58)
3	0.028	0.139	0.346	0.718
	(0.61)	(1.02)	(1.24)	(1.31)
4	0.049	0.127	0.254	0.434
	(1.27)	(1.10)	(1.05)	(0.96)
5	0.018	0.078	0.175	0.342
	(0.48)	(0.71)	(0.81)	(0.77)
6	0.017	0.077	0.227	0.491
	(0.36)	(0.59)	(0.87)	(1.02)
7	0.027	0.094	0.314	0.667
	(0.54)	(0.61)	(0.94)	(1.07)
8	0.024	0.062	0.178	0.524
	(0.46)	(0.39)	(0.57)	(0.82)
9	-0.005	0.018	0.094	0.249
	(0.11)	(0.13)	(0.33)	(0.45)
10(High)	0.005	0.008	0.073	0.139
	(0.09)	(0.05)	(0.23)	(0.23)
Low - High	0.052*	0.169***	0.354***	0.731***
	(1.67)	(3.18)	(4.49)	(5.64)

Table 11: **Predicting Mutual Fund Returns with Centrality: Control for Manager Tenure and Historical Performance**

The sample includes active managed U.S. equity mutual funds between 1996 and 2010 (I restrict the sample to those with Morningstar category in the 3 by 3 size/value grid and having monthly return information in CRSP survivor-bias-free mutual fund database). The dependent variables are monthly fund gross return and Fama-French-Carhart four-factor alpha (factor loadings are calculated using monthly fund returns of prior 36 months). $Log(TNA)_t$ is the natural logarithm of total net assets of the fund. $Log(FamilySize)_t$ is the natural logarithm of total net assets of all active equity funds in the fund family.$NumMgrs_t$ is the number of managers managing the fund during month t. $NetFlow_t$ is calculated as $NetFlow_t = \frac{TNA_t - TNA_{t-1}(1+R_t)}{TNA_{t-1}}$, where R_t is the net return of the fund during month t. $TurnoverRatio_t$ is the minimum of aggregated sales or aggregated purchases of securities, divided by the average 12-month Total Net Assets of the fund. Age_t is the age of the fund since inception. $ManagerSAT_t$ is the median SAT of matriculants at the manager's undergraduate institution. $ManagerMBA_t$ is a dummy variable which equals to one if the manager has an MBA degree and zero otherwise. $Log(ManagerTenure + 1)$ is defined as the natural logarithm of the number of years the manager have been working in the fund plus 1. I rank 4-factor alpha of every fund and assign a percentile ranking (higher percentile, better performance). I average the percentile ranking for all funds managed by all managers in our sample. $MgrPerformanceHist_t$ is the historical average of manager's percentile average at the fund level. If a fund has multiple managers, I average $MgrPerformanceHist_t$ equally across managers in the fund level. β_{MKT}, β_{SMB}, β_{HML}, and β_{UMD} are estimates from monthly rolling regressions of gross fund returns on Fama-French-Carhart four factors (MKT, SMB, HML, and UMD) using a 24-month window. Coefficients of Fama-MacBeth (1973) regressions are reported. t-statistics, which are in parentheses, are adjusted (12 monthly lags) for serial correlation and heteroscedasticity following Newey-West (1987). ***, **, and * indicate significance at the 1, 5, and 10 percent levels, respectively.

	Gross Return		Fama-French-Carhart Alpha	
	(1)	(2)	(3)	(4)
$EigenvectorCentrality_t$	-1.349***	-1.245***	-1.785***	-1.583***
	(3.13)	(2.79)	(3.53)	(2.98)
$Log(ManagerTenure + 1)_t$	-0.001		-0.000	
	(0.09)		(0.01)	
$MgrPerformanceHist_t$		1.126***		1.162***
		(5.64)		(4.98)
$Log(TNA)_t$	-0.014***	-0.020***	-0.015***	-0.021***
	(2.94)	(3.50)	(2.96)	(3.77)
$Log(FamilySize)_t$	0.009***	0.009***	0.007**	0.007**
	(3.18)	(3.22)	(2.06)	(2.19)
$NumMgrs_t$	0.000	0.001	0.003	0.003
	(0.18)	(0.52)	(0.85)	(1.05)
$Log(Age + 1)_t$	-0.019	-0.009	-0.017	-0.005

	(1.38)	(0.80)	(1.35)	(0.51)
$NetFlow_t$	0.833***	0.628***	0.814***	0.571***
	(4.04)	(3.63)	(3.30)	(2.88)
$NetFlow_t^2$	-0.910**	-0.670**	-0.903**	-0.579
	(2.60)	(2.04)	(2.15)	(1.51)
$TurnoverRatio_t$	0.024	0.025	0.008	0.010
	(1.09)	(1.20)	(0.24)	(0.33)
$ManagerSAT_t$	0.021***	0.019***	0.020***	0.018***
	(4.24)	(3.65)	(3.45)	(2.86)
$ManagerMBA_t$	0.047***	0.042***	0.035**	0.031**
	(4.08)	(3.79)	(2.59)	(2.32)
$\beta_{MKT,t}$	0.145	0.174		
	(0.40)	(0.48)		
$\beta_{SMB,t}$	0.285	0.268		
	(1.23)	(1.17)		
$\beta_{HML,t}$	0.170	0.187		
	(0.57)	(0.63)		
$\beta_{UMD,t}$	-0.006	0.015		
	(0.02)	(0.04)		
Average R^2	0.425	0.428	0.032	0.039
No. of months	180	180	180	180
No. of observations	266,657	266,657	267,617	267,242

Table 12: **Predicting Mutual Fund Returns with Centrality: Directed Social Connections**

The sample includes actively managed U.S. equity mutual funds between 1996 and 2010 (I restrict the sample to those with Morningstar category in the 3 by 3 size/value grid and having monthly return information in CRSP survivor-bias-free mutual fund database). The dependent variable is Fama-French-Carhart four-factor alpha (calculated based on fund gross returns). See section 5.5 for the definition of $EigenvectorCentrality(In)_t$, $DegreeCentrality(In)_t$, $EigenvectorCentrality(Out)_t$ and $DegreeCentrality(Out)_t$. Coefficients of Fama-MacBeth (1973) regressions are reported. t-statistics, which are in parentheses, are adjusted (12 monthly lags) for serial correlation and heteroscedasticity following Newey-West (1987). ***, **, and * indicate significance at the 1, 5, and 10 percent levels, respectively.

	(1)	(2)	(3)	(4)
$DegreeCentrality(In)_t$	-0.290***			
	(3.81)			
$DegreeCentrality(Out)_t$		-0.326***		
		(3.98)		
$EigenvectorCentrality(In)_t$			-0.988***	
			(3.00)	
$EigenvectorCentrality(Out)_t$				-1.277***
				(3.29)
$Log(TNA)_t$	-0.014**	-0.015**	-0.014**	-0.015**
	(2.48)	(2.55)	(2.44)	(2.55)
$Log(FamilySize)_t$	0.009**	0.009***	0.009**	0.010***
	(2.48)	(2.62)	(2.58)	(2.87)
$NumMgrs_t$	0.003	0.004	0.003	0.005
	(0.79)	(1.02)	(0.88)	(1.40)
$Log(Age+1)_t$	-0.018*	-0.019*	-0.018*	-0.019*
	(1.68)	(1.74)	(1.65)	(1.77)
$NetFlow_t$	0.717***	0.716***	0.725***	0.725***
	(2.76)	(2.77)	(2.78)	(2.80)
$NetFlow_t^2$	-0.928**	-0.908**	-0.943**	-0.911**
	(2.05)	(2.02)	(2.07)	(2.04)
$TurnoverRatio_t$	0.012	0.012	0.012	0.013
	(0.41)	(0.42)	(0.41)	(0.42)
Average R^2	0.029	0.029	0.028	0.029
No. of months	180	180	180	180
No. of observations	291,649	291,649	291,649	291,649

Table 13: **Relationship between Centrality and Fund Flows**

This table reports the results of pooled regression on the relationship between fund flows and centrality. The dependent variable is $NetFlow$, calculated as $NetFlow_t = \frac{TNA_t - TNA_{t-1}(1+R_t^g)}{TNA_{t-1}}$, where R_t^g is the gross return of the fund during month t. α_{t-1}^{4f} is monthly Fama-French-Carhart 4-factor alpha lagged by one month. R_{t-1} is monthly lagged net return of the fund. $EigenvectorCentrality$ is the eigenvector centrality of the fund. $ReturnVol$ is the standard deviation of the monthly gross return of the fund (using a 24-month window). Month fixed effects are included. Robust standard errors, reported in parentheses, are two-way clustered at the family and month levels. ***, **, and * indicate significance at the 1, 5, and 10 percent levels, respectively.

	$NetFlow_t$					
	(1)	(2)	(3)	(4)	(5)	(6))
α_{t-1}^{4f}	0.226***	0.243***	0.464***			
	(11.13)	(12.02)	(9.39)			
R_{t-1}				0.237***	0.246***	0.316***
				(8.66)	(9.10)	(8.72)
$EigenvectorCentrality_{t-1}$		-0.076***	-0.076***		-0.070***	-0.070***
		(3.42)	(3.43)		(3.29)	(3.28)
$\alpha_{t-1}^{4f} \times EigenvectorCentrality_{t-1}$		-1.251**	-1.461**			
		(2.06)	(2.48)			
$R_{t-1} \times EigenvectorCentrality_{t-1}$					-0.612**	-0.655***
					(2.51)	(2.69)
$\alpha_{t-1}^{4f} \times Log(Age+1)_{t-1}$			-0.107***			
			(6.33)			
$R_{t-1} \times Log(Age+1)_{t-1}$						-0.035***
						(5.09)
$ReturnVol_{t-1}$	-0.093***	-0.092***	-0.092***	-0.081**	-0.079**	-0.077**
	(2.62)	(2.59)	(2.59)	(2.19)	(2.15)	(2.10)

$Log(Age+1)_{t-1}$	-0.011***	-0.011***	-0.012***	-0.011***	-0.011***	-0.011***
	(18.04)	(18.13)	(18.16)	(18.22)	(18.29)	(18.34)
$Log(FundSize)_{t-1}$	-0.001**	-0.001**	-0.001**	-0.001**	-0.001**	-0.001**
	(2.27)	(2.20)	(2.20)	(2.32)	(2.25)	(2.26)
$Log(FamilySize)_{t-1}$	0.001**	0.001***	0.001***	0.001**	0.001***	0.001***
	(2.29)	(2.72)	(2.74)	(2.27)	(2.71)	(2.71)
$TurnoverRatio_{t-1}$	0.001	0.001	0.001	0.001	0.001	0.001
	(0.64)	(0.67)	(0.66)	(0.62)	(0.64)	(0.63)
$ExpenseRatio_{t-1}$	0.100***	0.099***	0.095***	0.103***	0.103***	0.101***
	(4.97)	(4.94)	(4.60)	(4.98)	(4.95)	(4.79)
Month Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R^2	0.037	0.037	0.038	0.041	0.041	0.042
No. of observations	290,256	290,256	290,256	290,310	290,310	290,310

Table 14: **Summary Statistics of Sample Stocks**

The sample includes common stocks from NYSE/AMEX and NASDAQ between 1996 and 2010. For each quarter t, I classify fund i with above median eigenvector centrality as *central* fund and below median eigenvector centrality as *peripheral* fund. The average portfolio weights in stock k for *central* investors and *peripheral* investors are represented as $CTR_{k,t}$ and $PER_{k,t}$, respectively. PMC factor is constructed as the difference in average portfolio weights between *peripheral* funds and *central* funds, $PMC_{k,t} = \frac{PER_{k,t} - CTR_{k,t}}{2}$. $ALL_{k,t}$ is the average portfolio weights of stock k for all funds in the sample. $\Delta BREADTH_t$ is the change in breadth of ownership from the end of quarter $t-1$ to quarter t. ΔIO_t is the change in fraction of shares outstanding of a stock held by 13F institutions from the end of quarter $t-1$ to quarter t. $LOG(SIZE)_t$ is the log market capitalization at the end of quarter t. BK/MKT_t is the most recently available observation of the book-to-market ratio at the end of quarter t. $MOM12_t$ is the raw stock return for the last 12 months excluding the recent one month. XTR_t is the quarterly share turnover (volume normalized by shares outstanding) adjusted for the average share turnover of the firm's exchange. CTR_t, PER_t, PMC_t, ALL_t, and $\Delta BREADTH_t$ are expressed in basis points ($\times 10,000$). ΔIO_t is expressed in percentage terms ($\times 100$). Size quintiles are determined using NYSE breakpoints.

Panel A: Means and standard deviations

	PER_t	CTR_t	PMC_t	ALL_t	$\Delta BREADTH_t$	ΔIO_t	$LOG(SIZE)_t$	BK/MKT_t	MOM_t	XTR_t
Size Quintile 1										
Mean	29.54	13.59	7.67	22.33	-0.43	0.15	4.54	0.91	0.05	-0.24
Std. dev.	32.19	19.19	15.04	22.04	31.53	8.59	1.00	1.10	0.88	1.85
Median	19.64	6.28	3.80	16.45	0.00	0.06	4.66	0.68	-0.06	-0.59
No. of obs.	105,057	107,384	105,035	105,137	108,986	109,744	109,744	101,065	107,316	109,744
Size Quintile 2										
Mean	48.47	32.12	8.26	40.65	8.39	1.17	6.16	0.62	0.23	0.42
Std. dev.	31.31	29.17	17.63	24.96	53.34	11.14	0.50	0.64	0.89	2.05
Median	43.52	27.33	6.81	37.18	6.61	0.78	6.13	0.51	0.09	-0.11
No. of obs.	47,108	47,222	46,653	47,238	46,716	47,400	47,400	42,701	45,587	47,400
Size Quintile 3										
Mean	61.51	45.00	8.45	53.66	14.60	1.08	7.01	0.56	0.30	0.63
Std. dev.	30.84	36.82	20.34	28.52	70.15	9.93	0.44	0.57	1.02	2.22
Median	57.90	40.66	7.74	50.62	11.21	0.73	6.98	0.46	0.13	0.06
No. of obs.	32,574	32,613	32,319	32,622	32,306	32,760	32,760	29,913	31,625	32,759
Size Quintile 4										
Mean	71.09	54.55	8.42	63.31	18.13	0.56	7.90	0.53	0.30	0.62
Std. dev.	30.46	38.76	20.13	29.92	89.64	9.78	0.44	0.45	1.05	2.05
Median	68.37	50.24	8.17	60.54	14.65	0.49	7.91	0.42	0.15	0.07
No. of obs.	26,302	26,309	26,122	26,315	26,226	26,497	26,497	24,693	25,841	26,496
Size Quintile 5										
Mean	91.41	74.77	7.29	83.03	27.28	0.26	9.55	0.47	0.27	0.28
Std. dev.	33.03	43.34	19.45	35.44	143.98	9.65	0.93	0.38	0.83	1.67
Median	88.34	70.02	7.53	79.12	21.14	0.30	9.37	0.37	0.15	-0.14
No. of obs.	21,850	21,078	22,382	21,199	22,694	22,778	22,778	21,953	22,514	22,778
Total										
Mean	48.34	31.78	7.94	40.62	8.07	0.53	6.05	0.72	0.17	0.15
Std. dev.	37.71	35.57	17.44	32.76	68.19	9.57	1.83	0.86	0.92	1.99
Median	42.69	24.30	5.82	35.43	0.00	0.30	5.93	0.54	0.05	-0.29
No. of obs.	232,891	234,606	232,511	232,511	236,928	239,179	239,179	220,325	232,883	239,177

Panel B: Contemporaneous correlations

	PER_t	CTR_t	PMC_t	ALL_t	$\Delta BREADTH_t$	ΔIO_t	$LOG(SIZE)_t$	BK/MKT_t	MOM_t	XTR_t
PER_t	1.000									
CTR_t	0.530	1.000								
PMC_t	0.510	-0.459	1.000							
ALL_t	0.879	0.870	0.038	1.000						
$\Delta BREADTH_t$	0.232	0.204	0.036	0.250	1.000					
ΔIO_t	0.076	0.068	0.011	0.083	0.262	1.000				
$LOG(SIZE)_t$	0.577	0.581	0.015	0.662	0.123	0.043	1.000			
BK/MKT_t	-0.186	-0.168	-0.024	-0.202	-0.013	-0.013	-0.251	1.000		
MOM_t	0.195	0.173	0.029	0.211	0.250	0.112	0.131	0.048	1.000	
XTR_t	0.096	0.129	-0.031	0.128	0.005	-0.021	0.186	-0.097	0.163	1.000

Panel C: Autocorrelations and cross-autocorrelations

	PER_{t-1}	CTR_{t-1}	PMC_{t-1}	ALL_{t-1}	$\Delta BREADTH_{t-1}$	ΔIO_{t-1}	$LOG(SIZE)_{t-1}$	BK/MKT_{t-1}	MOM_{t-1}	XTR_{t-1}
PER_t	0.837	0.477	0.356	0.747	0.170	0.064	0.538	-0.177	0.142	0.067
CTR_t	0.477	0.579	-0.053	0.588	0.162	0.057	0.550	-0.161	0.150	0.108
PMC_t	0.365	-0.051	0.437	0.189	0.009	0.013	0.011	-0.024	-0.001	-0.034
ALL_t	0.746	0.590	0.171	0.769	0.185	0.066	0.594	-0.186	0.163	0.095
$\Delta BREADTH_t$	0.129	0.120	0.014	0.142	0.086	0.025	0.078	-0.010	0.110	-0.016
ΔIO_t	0.023	0.019	0.008	0.026	0.013	-0.215	0.028	-0.008	0.041	-0.023
$LOG(SIZE)_t$	0.556	0.562	0.016	0.614	0.140	0.042	0.984	-0.250	0.134	0.153
BK/MKT_t	-0.200	-0.181	-0.028	-0.212	-0.031	-0.024	-0.278	0.894	-0.034	-0.103
MOM_t	0.195	0.168	0.038	0.211	0.268	0.120	0.109	0.087	0.639	0.134
XTR_t	0.109	0.142	-0.023	0.140	0.053	0.032	0.169	-0.094	0.190	0.737

Table 15: **Forecasting Returns using Fund Holdings**

The sample includes common stocks from NYSE/AMEX and NASDAQ between 1996 and 2010. In columns (1) and (2), the dependent variable is quarterly raw cumulative return for sample stocks. In columns (3) and (4), the dependent variable is quarterly Fama-French-Carhart 4-factor alpha for sample stocks. In columns (5) and (6), the dependent variable is quarterly DGTW-adjusted cumulative return for sample stocks. Panel A studies the stock returns over the next quarter $t+1$ following the focal quarter t. Panel B studies the stock returns over the quarter $t+2$ following the focal quarter t. Panel C studies the stock returns over the quarter $t+3$ following the focal quarter t. Panel D studies the stock returns over the quarter $t+4$ following the focal quarter t. For each quarter t, I classify fund i with above median eigenvector centrality as *central* fund and below median eigenvector centrality as *peripheral* fund. The average portfolio weights in stock k for *central* investors and *peripheral* investors are represented as $CTR_{k,t}$ and $PER_{k,t}$, respectively. PMC factor is constructed as the difference in average portfolio weights between *peripheral* funds and *central* funds, $PMC_{k,t} = \frac{PER_{k,t} - CTR_{k,t}}{2}$. $ALL_{k,t}$ is the average portfolio weights of stock k for all funds in the sample. $\Delta BREADTH_t$ is the change in breadth of ownership from the end of quarter $t-1$ to quarter t. ΔIO_t is the change in fraction of shares outstanding of a stock held by 13F institutions from the end of quarter $t-1$ to quarter t. XTR_t is the quarterly share turnover (volume normalized by shares outstanding) adjusted for the average share turnover of the firm's exchange. Coefficients of Fama-MacBeth (1973) regressions are reported. t-statistics, which are in parentheses, are adjusted (using 4 lags) for serial correlation and heteroscedasticity following Newey-West (1987). ***, **, and * indicate significance at the 1, 5, and 10 percent levels, respectively.

Panel A: Quarter 1

	Raw Return		Fama-French-Carhart Alpha		DGTW-adjusted Return	
	(1)	(2)	(3)	(4)	(5)	(6)
PER_t^H	0.793**		0.454		0.706**	
	(2.38)		(1.41)		(2.45)	
CTR_t^H	-1.075**		-1.495*		-1.097**	
	(2.08)		(1.88)		(2.32)	
PMC_t^H		2.126***		2.216**		2.023***
		(3.98)		(2.57)		(3.99)
ALL_t^H		-0.332		-1.099		-0.435
		(0.46)		(1.32)		(0.69)
$\Delta BREADTH_t$	0.531**	0.532**	-0.028	-0.028	0.337	0.337
	(2.46)	(2.45)	(0.08)	(0.08)	(1.59)	(1.60)
XTR_t	-0.004	-0.004	-0.003*	-0.003*	-0.003*	-0.003*
	(1.61)	(1.61)	(1.89)	(1.88)	(1.74)	(1.73)
ΔIO_t	-0.028*	-0.028*	-0.051**	-0.052**	-0.032**	-0.033**
	(1.96)	(1.97)	(2.62)	(2.63)	(2.20)	(2.23)
BK/MKT_t	0.003	0.003	-0.003	-0.003	-0.003	-0.003
	(0.65)	(0.65)	(0.81)	(0.81)	(1.21)	(1.21)
MOM_t	-0.001	-0.001	0.005	0.005	0.001	0.001
	(0.12)	(0.12)	(0.65)	(0.64)	(0.19)	(0.19)
$Log(Size)_t$	-0.001	-0.001	-0.000	-0.000	0.002	0.002
	(0.44)	(0.42)	(0.22)	(0.19)	(1.60)	(1.63)
Average R^2	0.052	0.052	0.030	0.030	0.023	0.023
No. of quarters	60	60	60	60	60	60
No. of observations	214,128	214,128	188,633	188,633	210,893	210,893

Panel B: Quarter 2

	Raw Return		Fama-French-Carhart Alpha		DGTW-adjusted Return	
	(1)	(2)	(3)	(4)	(5)	(6)
PER_t^H	0.856***		0.787***		0.678***	
	(3.05)		(3.28)		(2.76)	
CTR_t^H	-0.973**		-0.943		-0.965**	
	(2.10)		(1.50)		(2.39)	
PMC_t^H		2.126***		1.980***		1.904***
		(4.28)		(3.28)		(4.17)
ALL_t^H		-0.251		-0.261		-0.386
		(0.43)		(0.39)		(0.81)
$\Delta BREADTH_t$	0.167	0.177	-0.065	-0.058	0.046	0.054
	(0.66)	(0.70)	(0.22)	(0.19)	(0.20)	(0.23)
XTR_t	-0.003	-0.003	-0.002*	-0.002*	-0.003*	-0.003*
	(1.54)	(1.54)	(1.93)	(1.93)	(1.76)	(1.76)
ΔIO_t	-0.006	-0.006	0.003	0.003	-0.012	-0.012
	(0.55)	(0.56)	(0.20)	(0.20)	(1.15)	(1.17)
BK/MKT_t	0.003	0.003	-0.000	-0.000	-0.003	-0.003
	(0.74)	(0.73)	(0.10)	(0.11)	(1.12)	(1.12)
MOM_t	-0.007	-0.007	0.006*	0.006*	-0.001	-0.001
	(0.58)	(0.58)	(1.99)	(1.99)	(0.27)	(0.26)
$Log(Size)_t$	0.000	0.000	-0.000	0.000	0.003**	0.003**
	(0.03)	(0.11)	(0.02)	(0.05)	(2.21)	(2.30)
Average R^2	0.048	0.048	0.023	0.023	0.018	0.018
No. of quarters	60	60	60	60	60	60
No. of observations	209,599	209,599	184,909	184,909	205,855	205,855

Panel C: Quarter 3

	Raw Return		Fama-French-Carhart Alpha		DGTW-adjusted Return	
	(1)	(2)	(3)	(4)	(5)	(6)
PER_t^H	0.520*		0.710**		0.461**	
	(1.86)		(2.58)		(2.00)	
CTR_t^H	-0.896**		-0.664		-0.767**	
	(2.02)		(1.14)		(2.05)	
PMC_t^H		1.654***		1.595***		1.448***
		(3.48)		(2.93)		(3.55)
ALL_t^H		-0.485		-0.050		-0.401
		(0.84)		(0.07)		(0.83)
$\Delta BREADTH_t$	0.106	0.111	0.084	0.087	0.124	0.129
	(0.39)	(0.41)	(0.39)	(0.41)	(0.76)	(0.79)
XTR_t	-0.003	-0.003	-0.002	-0.002	-0.002	-0.002
	(1.20)	(1.20)	(1.58)	(1.58)	(1.30)	(1.29)
ΔIO_t	-0.044***	-0.044***	-0.032**	-0.032**	-0.038***	-0.038***
	(3.03)	(3.03)	(2.56)	(2.57)	(3.14)	(3.14)
BK/MKT_t	0.002	0.002	-0.001	-0.001	-0.004	-0.004
	(0.60)	(0.59)	(0.20)	(0.20)	(1.17)	(1.17)
MOM_t	-0.008	-0.008	0.004	0.004	-0.003	-0.003
	(1.24)	(1.23)	(1.14)	(1.14)	(1.44)	(1.42)
$Log(Size)_t$	-0.000	-0.000	-0.001	-0.001	0.002	0.002
	(0.14)	(0.08)	(0.42)	(0.36)	(1.56)	(1.65)
Average R^2	0.045	0.045	0.023	0.023	0.016	0.016
No. of quarters	60	60	60	60	60	60
No. of observations	205,094	205,094	181,222	181,222	200,853	200,853

Panel D: Quarter 4

	Raw Return		Fama-French-Carhart Alpha		DGTW-adjusted Return	
	(1)	(2)	(3)	(4)	(5)	(6)
PER_t^H	0.262		0.718**		0.248	
	(0.95)		(2.48)		(1.13)	
CTR_t^H	-0.501		-0.098		-0.207	
	(1.00)		(0.13)		(0.45)	
PMC_t^H		0.867		0.879		0.578
		(1.55)		(1.38)		(1.11)
ALL_t^H		-0.348		0.571		-0.072
		(0.60)		(0.67)		(0.14)
$\Delta BREADTH_t$	-0.367	-0.358	-0.229*	-0.223*	-0.404***	-0.395**
	(1.52)	(1.47)	(1.72)	(1.69)	(2.67)	(2.61)
XTR_t	-0.002	-0.002	-0.002*	-0.002*	-0.002	-0.002
	(1.10)	(1.10)	(1.78)	(1.78)	(1.33)	(1.33)
ΔIO_t	0.005	0.005	0.004	0.004	0.006	0.006
	(0.38)	(0.38)	(0.29)	(0.28)	(0.53)	(0.52)
BK/MKT_t	0.001	0.001	0.000	0.000	-0.003	-0.003
	(0.21)	(0.21)	(0.01)	(0.01)	(0.96)	(0.96)
MOM_t	-0.004	-0.004	0.000	0.000	-0.002	-0.002
	(0.92)	(0.90)	(0.08)	(0.06)	(0.89)	(0.88)
$Log(Size)_t$	-0.000	-0.000	-0.000	-0.000	0.002**	0.002**
	(0.16)	(0.10)	(0.22)	(0.20)	(2.24)	(2.39)
Average R^2	0.042	0.042	0.022	0.022	0.016	0.016
No. of quarters	60	60	60	60	60	60
No. of observations	200,635	200,635	177,600	177,600	195,934	195,934

Table 16: **Forecasting Earnings Surprises**

The sample includes common stocks from NYSE/AMEX and NASDAQ between 1996 and 2010. The dependent variable is earnings announcement surprises (SUE) over the first quarter (columns (1) and (2)) and the second quarter (columns (3) and (4)). We define the SUE as the difference between the actual and consensus EPS, scaled by the share price at the beginning of the quarter. Consensus EPS is the median of latest analyst forecasts issued within 90 days prior to the earning announcement date. Each quarter t, I classify fund i with above median eigenvector centrality as *central* fund and below median eigenvector centrality as *peripheral* fund. The average portfolio weights in stock k for *central* investors and *peripheral* investors are represented as $CTR_{k,t}$ and $PER_{k,t}$, respectively. PMC factor is constructed as the difference in average portfolio weights between *peripheral* funds and *central* funds, $PMC_{k,t} = \frac{PER_{k,t} - CTR_{k,t}}{2}$. $ALL_{k,t}$ is the average portfolio weights of stock k for all funds in the sample. Coefficients of Fama-MacBeth (1973) regressions are reported. t-statistics, which are in parentheses, are adjusted (using 4 lags) for serial correlation and heteroscedasticity following Newey-West (1987). ***, **, and * indicate significance at the 1, 5, and 10 percent levels, respectively.

	Qtr 1		Qtr 2	
	(1)	(2)	(3)	(4))
PER_t^H	0.387***		0.542***	
	(3.02)		(3.37)	
CTR_t^H	-0.742**		-0.462*	
	(2.44)		(1.73)	
PMC_t^H		1.179***		1.118***
		(2.92)		(2.78)
ALL_t^H		-0.341		0.063
		(1.58)		(0.35)
BK/MKT_t	-0.006**	-0.006**	-0.007**	-0.007**
	(2.41)	(2.41)	(2.28)	(2.28)
MOM_t	0.010***	0.010***	0.012***	0.012***
	(3.78)	(3.79)	(3.89)	(3.90)
$Log(Size)_t$	0.004***	0.004***	0.004***	0.004***
	(7.63)	(7.80)	(10.97)	(11.32)
Average R^2	0.040	0.040	0.042	0.042
No. of quarters	60	60	60	60
No. of observations	170,013	170,013	159,129	159,129