Three Essays in Institutional Trading and Corporate Finance:

Author: Yuyuan Zhu

Boston College Electronic Thesis or Dissertation, 2017

Boston College

Wallace E. Carroll Graduate School of Management

Department of Finance

THREE ESSAYS IN INSTITUTIONAL TRADING AND CORPORATE FINANCE

a dissertation by

Yuyuan Zhu

submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Finance

May 2017

Acknowledgments

I am deeply grateful to my advisor, Dr. Thomas J. Chemmanur, who has been the most supportive and has always given me the freedom to pursue my research interests without any constraints. His advice on research, career and even life have been priceless. Consciously or unconsciously, he has taught me how good research shall be done. His guidance inspires me to continue my research in finance and grow as a researcher. He is a tremendous mentor.

I also want to thank Dr. Jonathan Reuter, Dr. Leonard Kostovetsky, Dr. Alan Marcus, and Dr. Robert Taggart for all the valuable and insightful comments which have provoked new thoughts on my research. I am also thankful to all the faculty members at Department of Finance of Boston College, my fellow PhD students, for providing an intellectually stimulating research environment. I have benefited immensely over the past years.

Lastly, and most importantly, I would like to thank my parents and wife Tingting Nian, whose love, encouragement and faithful support have been so instrumental for me, particularly during the tough time of my Ph.D. pursuit. Without your love, this dissertation would have never been possible. I am indebted to them forever.

Contents

THREE ESSAYS IN INSTITUTIONAL TRADING AND CORPORATE FINANCE

by

Yuyuan Zhu

Finance Department, Boston College

Dissertation Committee:

Professor Thomas Chemmanur (Chair)

Professor Leonard Kostovetsky

Professor Alan Marcus

Professor Jonathan Reuter

Professor Robert Taggart

Abstract

My dissertation is comprised of three chapters. In this first chapter, I study the effect of social connections on mutual fund investors' information production and accuracy of their signals. 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). 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.

In the second chapter, co-authored with Thomas Chemmanur, Yingzhen Li, and Jie Xie, we propose a "noisy signaling" hypotheses of open market share repurchase (OMSR) programs, where the equity market equilibrium that prevails after OMSR program announcements is a partial pooling rather than a fully separating equilibrium. We argue that two complementary mechanisms, namely, actual share repurchases by firms and information production by institutions, serve to reduce the residual equity market information asymmetry facing firms subsequent to OMSR program announcements. We test the implications of this noisy signaling hypothesis using transaction-level data on trading by institutions and by a subsample of identified hedge funds, and find strong support for the above hypothesis.

In the third chapter, co-authored with Thomas Chemmanur, and Jiekun Huang, we analyze how the geographical locations of institutions affect their investments in IPOs and various characteristics of the IPOs that they invest in. We argue that institutions geographically close to each other may free-ride on each other's information when evaluating IPOs, resulting in IPOs dominated by geographically clustered institutions reflecting less accurate information signals compared to those dominated by geographically dispersed institutions. We find that the equity holdings of institutions in IPOs are influenced more by the investments made by neighboring institutions. We show that an increase in the geographical dispersion of the institutions investing in an IPO is associated with higher IPO price revisions, higher firm valuations at offering and secondary market, larger IPO initial returns, greater long-run post-IPO stock returns lower information asymmetry facing an IPO firm in the equity market. Finally, the predictive power of institutional trading post-IPO for subsequent long-run stock returns and earnings surprises for the first fiscal-year end after the IPO is greater for geographically isolated institutions compared to those that are geographically clustered.

Chapter 1

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

1.1 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)).

[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).

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 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

[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.

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 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.

[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) 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)).

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 independent information 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

[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.

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 several variables, which summarize a fund manager's alpha generation during his entire career in the mutual fund industry, to proxy for the fund manager's “skill". While I find these “skill" variables have significant predictive power for fund returns, the predictive power of centrality 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 1.2 discusses the related literature. In Section 1.3, I describe the data and the construction of the mutual fund sample used in my empirical analysis. In Section 1.4, I study whether social connections have an effect on the portfolio holdings and trades for mutual funds. In Section 1.5, 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 1.6, I construct my PMC measure based on mutual fund holdings, and I study whether this PMC measure has predictive power for future abnormal

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

stock returns and earnings surprises. Section 1.7 concludes.

1.2 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 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,

[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 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 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 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.

1.3 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 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

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

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.

1.4 Social Connections and Mutual Fund Portfolios

1.4.1 Measuring overlap

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

$$PortOverlap_{i,j,t} = \sum_{k \in H_t} \min\{w_{i,k,t}, w_{j,k,t}\} \tag{1.1}$$

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

$$BuyOverlap_{i,j,t} = \frac{\sum_{k \in T_t} \min\{I^+_{i,k,t}, I^+_{j,k,t}\}}{\min\{\sum_{k \in T_t} I^+_{i,k,t}, \sum_{k \in T_t} I^+_{j,k,t}\}} \tag{1.2}$$

$$SellOverlap_{i,j,t} = \frac{\sum_{k \in T_t} \min\{I^-_{i,k,t}, I^-_{j,k,t}\}}{\min\{\sum_{k \in T_t} I^-_{i,k,t}, \sum_{k \in T_t} I^-_{j,k,t}\}} \tag{1.3}$$

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.

1.4.2 Summary Statistics of Fund Pairs

Table 1.1 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 1.1 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.

1.4.3 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

$$PortOverlap_{i,j,t} = \alpha + \beta Connected_{i,j,t} + \delta SameCity_{i,j,t} + \Gamma' Controls_{i,j,t} + \epsilon_{i,j,t} \qquad (1.4)$$

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 1.2 shows the coefficient estimates and t-statistics for various specifications of equation 1.4. 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 1.4. In Table 1.3, 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.

1.4.4 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 1.4 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 1.4 provide evidence against the hypothesis that the abnormal overlap in portfolio and trades between connected mutual funds is driven by unobserved managers' preferences.

1.5 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

[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.

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

1.5.1 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

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

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 1.5 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

$$NetFlow_t = \frac{TNA_t - TNA_{t-1}(1 + R_t)}{TNA_{t-1}} \tag{1.5}$$

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

1.5.2 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 ($EigenvectorCentrality_t$ and $DegreeCentrality_t$) on a list of fund characteristics ($Log(FundSize)_t$, $Log(FamilySize)_t$, $NumMgrs_t$, and $Log(Age + 1)_t$) and manager characteristics ($ManagerSAT_t$, $ManagerMBA_t$). Month fixed effects are included. Standard errors are clustered at the fund family level.

Table 1.6 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. In model (3) and (4), I also control for $MgrDollarAlphaHist_t$, which is cumulative dollar weighted alpha generated by fund managers (equally weighed across all managers at fund level). Interestingly, the coefficients for $MgrDollarAlphaHist_t$ are negative and highly statistically significant, suggesting that low-skilled fund managers are more likely to have higher centrality measures. I will explore the implication of this relationship further in Section 1.5.5.

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 $Turnover_t$, $ActiveShare_t$,[11] and $ExpenseRatio_t$ on the eigenvector centrality measure.

Table 1.7 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. In addition, mutual funds with higher eigenvector centrality have lower Active Share, and therefore they are more likely to be closet indexers. On the other hand, I do not find a significant relationship between eigenvector centrality measure and $Turnover_t$ or $ExpenseRatio_t$.

[11] $ActiveShare_t$ is Active Share measure defined in Petajisto (2013), which represents the share of portfolio holdings that differ from the benchmark index holdings. It is downloaded from Antti Petajisto's website.

1.5.3 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

$$r_{i,t+1} = \alpha + \beta Centrality_{i,t} + \gamma X_{i,t} + \epsilon_{i,t+1} \tag{1.6}$$

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.[12] 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 1.8 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.[13] 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),

[12] My results are robust to using net returns instead of gross returns.

[13] 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.

implying a one-standard-deviation increase in $EigenvectorCentrality_t$ predicts a decrease of 2.4 basis 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 1.8, 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.[14]

[14]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

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 quintile 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 months after the portfolio formation date. The equal-weighted (Panel A) and value-weighted (Panel B) returns of these portfolios are presented in Table 1.10. The 1-5 quintile spread is the zero-investment long-short portfolio that is long on quintile one and short on quintile five. In columns (1)-(4), I use Fama-French-Carhart 4-factor alpha based on the gross fund returns (before fee). In columns (5)-(8), I use Fama-French-Carhart 4-factor alpha based on the net fund returns (after fee). In columns (1)-(4) of Panel A, the 1-5 decile spread is 5.0 basis points for the 1-month horizon, 16.8 basis points for the 3-month horizon, 32.8 basis points for the 6-month horizon, and 63.5 basis points for 12-month horizon. These 1-5 decile spreads are highly statistically significant.[15] In columns (5) -(8), I find similar results when the 4-factor alphas of the portfolios are calculated based on after-fee fund net returns. In Panel B, where the portfolio returns are value-weighted (by fund size), I find even larger 1-5 quintile spread. In columns (1)-(4) of Panel B, the 1-5 decile spread is 7.3 basis points for the 1-month horizon, 26.9 basis points for the 3-month horizon, 54.9 basis points for the 6-month horizon, and 100.3 basis points for 12-month horizon. It implies that the negative relationship between negative centrality and future fund alpha is especially strong among funds with large total net assets (TNA).

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

size and find equally strong results.

[15]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.

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.

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

$$\begin{aligned} r_{i,t+1} - r^A_{j,t+1} = \alpha + \beta(&EigenvectorCentrality_{i,t} - EigenvectorCentrality^A_{j,t}) \\ &+ \gamma(X_{i,t} - X^A_{j,t}) + \epsilon_{i,t+1} \end{aligned} \tag{1.7}$$

$$r^A_{j,t+1} = \alpha + \beta EigenvectorCentrality^A_{j,t} + \gamma X^A_{j,t} + \epsilon_{j,t+1} \tag{1.8}$$

where $r^A_{j,t+1}$ 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, $EigenvectorCentrality^A_{j,t}$ and $X^A_{j,t}$ also represent the family average of their corresponding variable during period t. Table 1.9 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 1.9, 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.

1.5.5 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),[16] and the finding of an inverse relationship between mutual fund centrality and future fund performance in section 1.5.3 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 1.11. 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. The first variable is $MgrDollarAlphaHist_t$, which equals to the cumulative dollar weighted Cahart 4-factor alpha generated by a fund manager. The second variable is $MgrAlphaRankHist_t$. To construct $MgrAlphaRankHist_t$, I rank Fama-French-Carhart 4-factor alpha of every fund in every month 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.$MgrAlphaRankHist_t$ is defined as the cumulative average of a manager’s past 4-factor alpha percentile rankings. If a fund has multiple managers, I average $MgrDollarAlphaHist_t$ and $MgrAlphaRankHist_t$ equally across all managers in the fund. 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 $MgrDollarAlphaHist_t$ or $MgrAlphaRankHist_t$. The empirical result is presented in columns (2)-(3) and (5)-(6) of Table 1.11. The coefficients for $MgrDollarAlphaHist_t$ and $MgrAlphaRankHist_t$ are 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 $MgrDollarAlphaHist_t$ and $MgrAlphaRankHist_t$

[16]The negative relationship between past performance of managers and centrality is shown in Table 1.6.

as control variables. 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 1.12. 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.

1.5.6 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:

$$\begin{aligned} NetFlow_{i,t} &= \alpha + \beta_0 EigenvectorCentrality_{i,t-1} + \beta_1 Return_{t-1} \\ &\quad + \beta_2 Return_{t-1} \times EigenvectorCentrality_{i,t-1} + \gamma X_{i,t-1} + \epsilon_{i,t} \end{aligned} \tag{1.9}$$

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 1.13. 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%[17] 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 $R_{t-1} \times EigenvectorCentrality_{t-1}$ 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 1.5.3 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.[18]

1.6 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.

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

[18] An alternative explanation is that managers in higher centrality funds have less incentive to produce alpha because investors' flow into these funds are less responsive to past performance. However, we do not have evidence showing difference in clienteles between well connected funds and less 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.[19] 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.

1.6.1 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* funds and *peripheral* funds are represented as $CTR_{k,t}$ and $PER_{k,t}$, respectively.[20] I construct a PMC measure, which is defined as the difference in average portfolio weights between *peripheral* funds and *central* funds,

$$PMC_{k,t} = \frac{PER_{k,t} - CTR_{k,t}}{2} \tag{1.10}$$

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

Table 1.14 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^{21}. 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

[19]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).

[20]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.

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

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

Panel B of Table 1.14 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.

1.6.2 Forecast Stock Returns

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

$$r_{k,t+j-1,t+j} = \alpha + \beta_1^1 PER_{k,t} + \beta_2^1 CTR_{k,t} + \gamma X_{k,t} + \epsilon_{k,t,t+1} \tag{1.11}$$

$$r_{k,t+j-1,t+j} = \alpha + \beta_1^2 PMC_{k,t} + \beta_2^2 ALL_{k,t} + \gamma X_{k,t} + \epsilon_{k,t,t+1} \tag{1.12}$$

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 1.15, 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 1.15. 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.[22] 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 1.15, 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

[22] 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

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.[23]

The coefficient for PMC is also positive and statistically significant in Panel B and Panel C of Table 1.15, 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 1.15).

In conclusion, the results in Table 1.15 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 *central* funds, and that information is gradually impounded into stock prices through the portfolio rebalancing by these funds.

[23] 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).

1.6.3 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,

$$SUE_{k,t+j} = \alpha + \beta_1^1 PER_{k,t} + \beta_2^1 CTR_{k,t} + \gamma X_{k,t} + \epsilon_{k,t,t+1} \tag{1.13}$$

$$SUE_{k,t+j} = \alpha + \beta_1^2 PMC_{k,t} + \beta_2^2 ALL_{k,t} + \gamma X_{k,t} + \epsilon_{k,t,t+1} \tag{1.14}$$

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,

$$SUE_{k,t+j} = \frac{EPS^A_{k,t+j} - EPS^E_{k,t}}{P_t} \tag{1.15}$$

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 1.16 presents the results of Fama-MacBeth (1973) regressions of model 1.13. 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.

1.7 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.[24] 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.

[24] 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.

Table 1.1: **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 1.2: **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 1.3: **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 Manage (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 1.4: **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 1.5: **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
$FamilySize_t$	18474	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
$ManagerTenure_t$	5.69	5.27	4.33	1.00	11.67
$ManagerAge_t$	48	9	47	38	60
$R_t(GrossReturn, \%)$	0.70	5.77	1.13	-6.41	7.10
$\alpha_t^{4f}(GrossReturn, \%)$	0.03	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	-0.09	-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 1.6: **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 $EigenvectorCentrality_t$ and $DegreeCentrality_t$. $Log(FundSize)_t$ 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. $MgrDollarAlphaHist_t$ is cumulative dollar weighted alpha generated by every fund manager. If the fund is managed by multiple managers, $ManagerSAT_t$, $ManagerMBA_t$, $ManagerTenure_t$, $ManagerAge_t$ and $MgrDollarAlphaHist_t$ are averaged equally 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) $EigenvectorCentrality_t$	(4) $DegreeCentrality_t$
$Log(TNA)_t$	0.000*	0.003**	0.000	0.001
	(1.66)	(2.42)	(0.61)	(0.46)
$Log(FamilySize)_t$	0.002***	0.007***	0.002***	0.007***
	(7.00)	(5.25)	(5.71)	(3.74)
$NumMgrs_t$	0.002***	0.016***	0.002***	0.013***
	(12.25)	(13.77)	(10.53)	(9.69)
$Log(Age+1)_t$	-0.001***	-0.009***	-0.002***	-0.014***
	(3.08)	(3.40)	(4.92)	(4.27)
$ManagerSAT_t$	0.001***	0.005***		
	(2.71)	(2.98)		
$ManagerMBA_t$	0.003***	0.022***		
	(5.01)	(5.07)		

$Log(ManagerTenure + 1)_t$	-0.001***	-0.007**		
	(2.81)	(2.29)		
$Log(ManagerAge)_t$	-0.000	0.004		
	(0.06)	(0.36)		
$MgrDollarAlphaHist_{t-1}$			-0.079***	-0.702***
			(3.46)	(4.50)
Month Fixed Effects	Yes	Yes	Yes	Yes
Adjusted R^2	0.215	0.264	0.127	0.096
No. of observations	259,903	259,903	309,264	309,264

Table 1.7: **Mutual Fund Behavior and Centrality**

This table presents the results for the pooled panel regression for the relationship between mutual fund behavior and the eigenvector centrality measure. The dependent variables are β_{MKT}, β_{SMB}, β_{HML}, β_{UMD}, $Turnover_t$, $ActiveShare_t$, and $ExpenseRatio_t$. $ActiveShare_t$ is downloaded from Antti Petajisto's website, and is defined in Petajisto (2013). $ManagerSAT_t$ 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. $Log(FundSize)_t$ 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. 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) $\beta_{MKT,t}$	(2) $\beta_{SMB,t}$	(3) $\beta_{HML,t}$	(4) $\beta_{UMD,t}$	(5) $TurnoverRatio_t$	(6) $ActiveShare_t$	(7) $ExpenseRatio_t$
$EigenvectorCentrality_t$	1.142***	-0.747*	-0.766*	0.585***	0.358	-1.591***	-0.003
	(6.09)	(1.80)	(1.71)	(2.68)	(0.21)	(5.73)	(0.43)
$ManagerSAT_t$	0.006***	0.005	-0.004	-0.003	0.012	0.005	-0.000
	(2.64)	(0.89)	(0.70)	(0.76)	(0.59)	(1.56)	(0.69)
$ManagerMBA_t$	0.011	0.003	0.033*	-0.008	-0.083	-0.002	-0.000
	(1.64)	(0.16)	(1.90)	(0.98)	(1.35)	(0.21)	(0.85)
$Log(ManagerTenure+1)_t$	-0.022***	-0.020**	0.031***	-0.018***	-0.258***	0.016***	0.000
	(5.16)	(2.20)	(3.82)	(4.10)	(7.55)	(3.33)	(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	Yes
Adjusted R^2	0.061	0.012	0.078	0.053	0.030	0.052	0.061
No. of observations	276,951	276,951	276,951	276,951	271,399	49,288	277,746

Table 1.8: **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 1.9: **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 1.10: **Predicting Mutual Fund Returns with Centrality: Portfolio Sorts**
This table reports future Fama-French-Carhart 4-factor alpha for 5 quintile mutual fund portfolios formed based on the eigenvector centrality measure. At the start of each calendar month, I sort funds into quintile portfolios (low to high) based on the eigenvector centrality measure at the end of last month. In Panel A, I calculate equal-weighted Fama-French-Carhart 4-factor alpha of each portfolio over the next one month, three months, six months, and twelve months after portfolio formation. In Panel B, I calculate Fama-French-Carhart 4-factor alpha weighted by fund size for each portfolio over the next one month, three months, six months, and twelve months after portfolio formation. The 1-5 decile spread is the zero-investment long-short portfolio that is long on quintile one and short on quintile five. In columns (1) - (4), I use Fama-French-Carhart 4-factor alpha based on gross fund returns (before fee). In columns (5) - (8), I use Fama-French-Carhart 4-factor alpha based on net fund returns (after fee). 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.

	$\alpha^{4f}_{t+1}(GrossReturn, \%)$				$\alpha^{4f}_{t+1}(NetReturn, \%)$			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Quintile	1 Month	3 Months	6 Months	12 Months	1 Month	3 Months	6 Months	12 Months
Panel A: equal-weighted portfolio								
1(Low)	0.057	0.198*	0.446*	0.867**	-0.048	-0.122	-0.213	-0.513
	(1.43)	(1.71)	(1.91)	(2.01)	(1.19)	(1.05)	(0.91)	(1.19)
2	0.042	0.139	0.309	0.676	-0.057	-0.163	-0.309	-0.621
	(0.98)	(1.08)	(1.15)	(1.26)	(1.31)	(1.25)	(1.14)	(1.15)
3	0.019	0.086	0.221	0.386	-0.076**	-0.204*	-0.374*	-0.861**
	(0.52)	(0.78)	(0.99)	(0.95)	(2.00)	(1.83)	(1.64)	(2.09)
4	0.035	0.095	0.245	0.558	-0.064	-0.205	-0.371	-0.732
	(0.67)	(0.62)	(0.78)	(0.88)	(1.23)	(1.33)	(1.18)	(1.18)
5(High)	0.007	0.030	0.118	0.233	-0.092*	-0.271*	-0.500*	-1.058*
	(0.14)	(0.21)	(0.39)	(0.41)	(1.86)	(1.85)	(1.68)	(1.89)
Low - High	0.050*	0.168***	0.328***	0.635***	0.044*	0.148***	0.287***	0.545***
	(1.91)	(3.87)	(5.22)	(6.56)	(1.67)	(3.44)	(4.60)	(5.76)
Panel B: value-weighted portfolio								
1(Low)	0.014	0.077	0.224	0.425	-0.064	-0.162	-0.265	-0.596
	(0.26)	(0.48)	(0.71)	(0.71)	(1.16)	(1.02)	(0.85)	(1.03)

2	-0.005	-0.006	0.003	-0.048	-0.077	-0.227	-0.447*	-0.982**
	(0.08)	(0.04)	(0.01)	(0.10)	(1.41)	(1.54)	(1.65)	(1.98)
3	-0.035	-0.072	-0.056	-0.210	-0.109***	-0.299***	-0.519***	-1.172***
	(1.08)	(0.81)	(0.33)	(0.79)	(3.40)	(3.35)	(3.00)	(4.26)
4	-0.021	-0.060	-0.152	-0.371	-0.095**	-0.286**	-0.612**	-1.325***
	(0.49)	(0.49)	(0.63)	(0.82)	(2.21)	(2.27)	(2.46)	(2.80)
5(High)	-0.059	-0.193	-0.325	-0.578	-0.139***	-0.436***	-0.822***	-1.604***
	(1.20)	(1.36)	(1.18)	(1.21)	(2.82)	(3.08)	(3.01)	(3.42)
Low - High	0.073	0.269***	0.549***	1.003***	0.075*	0.274***	0.557***	1.008***
	(1.62)	(3.21)	(4.58)	(6.14)	(1.67)	(3.28)	(4.68)	(6.27)

Table 1.11: **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). $MgrAlphaRankHist_t$ is the cumulative average of the percentile ranking for every managers in our sample. $MgrDollarAlphaHist_t$ is cumulative dollar weighted alpha generated by every fund manager. If a fund has multiple managers, I average $MgrAlphaRankHist_t$ and $MgrDollarAlphaHist_t$ 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)	(5)	(6)
$EigenvectorCentrality_t$	-1.721***	-1.243***	-1.227***	-2.130***	-1.588***	-1.601***
	(3.45)	(2.78)	(2.73)	(3.83)	(2.98)	(2.96)
$Log(ManagerTenure + 1)_t$	0.005			0.006		
	(0.34)			(0.42)		
$MgrAlphaRankHist_t$		0.011***			0.012***	
		(5.64)			(4.98)	
$MgrDollarAlphaHist_t$			0.120***			0.121***
			(5.56)			(4.13)
$Log(TNA)_t$	-0.021***	-0.020***	-0.017***	-0.022***	-0.021***	-0.018***

	(3.78)	(3.49)	(3.27)	(3.70)	(3.75)	(3.42)
$Log(FamilySize)_t$	0.016***	0.009***	0.010***	0.013***	0.007**	0.008**
	(4.62)	(3.21)	(3.25)	(3.23)	(2.18)	(2.23)
$NumMgrs_t$	0.001	0.001	0.001	0.003	0.003	0.003
	(0.29)	(0.51)	(0.48)	(0.91)	(1.06)	(1.03)
$Log(Age + 1)_t$	-0.017	-0.009	-0.014	-0.016	-0.006	-0.012
	(1.29)	(0.82)	(1.24)	(1.29)	(0.56)	(1.18)
$NetFlow_t$	0.805***	0.625***	0.635***	0.805***	0.573***	0.578***
	(3.79)	(3.61)	(3.76)	(3.16)	(2.82)	(2.96)
$NetFlow_t^2$	-0.749**	-0.669**	-0.682**	-0.768*	-0.572	-0.559
	(2.14)	(2.05)	(2.01)	(1.86)	(1.48)	(1.43)
$TurnoverRatio_t$	0.028	0.025	0.024	0.011	0.010	0.010
	(1.17)	(1.20)	(1.19)	(0.32)	(0.33)	(0.35)
$ManagerSAT_t$	0.021***	0.019***	0.020***	0.020***	0.018***	0.020***
	(4.20)	(3.66)	(3.87)	(3.44)	(2.88)	(3.27)
$ManagerMBA_t$	0.046***	0.042***	0.047***	0.033***	0.030**	0.035***
	(3.90)	(3.80)	(4.10)	(2.62)	(2.32)	(2.63)
$\beta_{MKT,t}$	0.124	0.172	0.173			
	(0.34)	(0.48)	(0.47)			
$\beta_{SMB,t}$	0.291	0.267	0.277			
	(1.26)	(1.17)	(1.21)			
$\beta_{HML,t}$	0.165	0.189	0.186			
	(0.56)	(0.63)	(0.62)			
$\beta_{UMD,t}$	-0.023	0.015	0.012			
	(0.06)	(0.04)	(0.03)			
Average R^2	0.429	0.428	0.429	0.033	0.039	0.042
No. of months	180	180	180	180	180	180
No. of observations	261,811	266,657	266,657	262,749	267,242	267,242

Table 1.12: **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 1.5.5 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 1.13: **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 1.14: **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 1.15: **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 1.16: **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

Chapter 2

Noisy Signaling through Open Market Share Repurchase Programs and Information Production by Institutions

2.1 Introduction

In recent years, the number of firms undertaking stock repurchases has increased dramatically, while the proportion of firms distributing value through cash dividends has declined (see, e.g., Fama and French (2001)). Open-market share repurchases (OMSRs) constitute around 90% of the stock repurchases consummated in recent years: see, e.g., Comment and Jarrell (1991). An interesting question in this context is regarding the precise economic mechanism through which OMSR programs maximize shareholder value. The predominant rationale for share repurchase programs provided by the existing theoretical literature is that they serve to signal firm insiders' private information about the intrinsic value of the firm to outsiders in the equity market: see, e.g., Ofer and Thakor (1987) or Constantinides and Grundy (1989). However, there are certain important modifications of the traditional signaling paradigm that need to be made if we are to apply it to the case of OMSR programs.

Signaling models of share repurchase assume that the firm commits to repurchase a certain number of shares, as is the case in practice in Dutch auction or fixed-price tender offer repurchase programs. However, there is no such commitment to buy a specified number of shares in an OMSR program: the announcement of such a program involves announcing the authorization by the firm's board to repurchase a certain number of shares, not a commitment to buy these shares. This means that an important assumption underlying theoretical models of signaling private information using stock repurchases does not hold in the case of OMSR programs, implying that stock repurchases using OMSR programs may not be able to signal insider private information fully to the equity market. In particular, since top managers' compensation is often linked to stock price, firm managers have an incentive to boost stock prices in the short run, even if they believe that their equity is correctly valued or even overvalued. Given this incentive, and if there is no cost incurred by firm managers for not repurchasing a significant fraction of the shares announced in an OMSR program, even the managers of overvalued firms have an incentive to announce an OMSR program (in order to mimic the behavior of undervalued firms) but actually repurchase only a small fraction of the announced shares (or none at all).

The fact that OMSR programs may not be able to fully signal firm insiders' private information, however, does not necessarily imply that such programs are not able to convey any information at all to outside investors in the equity market. In this paper, we argue that as long as firm managers suffer a moderate reputational or other cost arising from the firm's actual repurchase falling short of the number of shares authorized in the OMSR program announcement, OMSR programs will be able to convey a noisy signal that the firm's equity is undervalued to outside shareholders.[1] In

[1]It is worth noting that the completion rate of actual OMSR programs announced in practice is broadly consistent with firms facing a moderate per share cost of not actually repurchasing the number of shares announced in the OMSR program (as required by our noisy signaling hypothesis). In particular, the evidence in our sample is that, on average, firms actually repurchase 80.11 percent of the shares announced in the OMSR program in the one-year period after the announcement. The evidence documented in various other papers in the existing literature point to a substantial completion rate in OMSR programs, on average: see, e.g., Stephens and Weisbach (1998), who document that firms acquire on average 74 to 82 percent of the number of shares announced within three years of the repurchase

other words, while an OMSR announcement itself may not fully eliminate the undervaluation of the firm's equity, it is likely to reduce it, with additional information being conveyed gradually over time as the firm subsequently repurchases a larger and larger number of its shares. We further argue that, in this setting, institutional investors play an important role complementary to OMSR program announcements and actual share repurchases by firms in reducing the information asymmetry faced by firms in the equity market (and therefore their equity undervaluation): institutions are able to accomplish this by producing information about announcing firms and trading in their equity after the announcement of OMSR programs. We will refer to the above hypothesis as the "noisy signaling hypothesis" of OMSR programs. The objective of this paper is to propose the above novel hypothesis about OMSR programs and to test its implications in the unique setting of information production and trading by institutional investors, using a detailed transaction-level institutional trading database.

The economic setting we consider to develop our empirical analysis can be described as follows. Consider a situation where the insiders of a firm, having private information about its intrinsic value, are considering whether or not to undertake an OMSR program. For concreteness, consider three types of firms: those with the highest intrinsic value (type G), medium intrinsic value (type M), and the lowest intrinsic value (type B). Prior to an OMSR program announcement, all three types of firms are pooled together (priced at the average value across types), so that the type G and type M firms are undervalued while the type B firm is overvalued. This means that the higher type firms have an incentive to announce an OMSR program to reduce their undervaluation: we argue that, even if there is no commitment to buy back all the shares announced in the program, announcing an OMSR program will convey a noisy signal of higher intrinsic value as long as there is at least a moderate reputational or other cost per share to firm management of having a shortfall in the number of shares actually repurchased relative to the target number of shares announced in the OMSR program ("shortfall cost" from now on). We argue that, in this setting, the type G and type

announcement.

M (higher intrinsic value) firms will announce an OMSR program while the type B firms (lowest intrinsic value) firm will not. Further, there will be an announcement effect (abnormal stock return) following such an announcement. However, the undervaluation of the highest firm type will not be completely eliminated by an OMSR program, since, after the announcement, the type G and type M firm will remain pooled together. Consequently, we argue that, after the OMSR program announcement, while both types will repurchase shares, the highest intrinsic value (type G) firms will repurchase a larger number of their own shares than the medium intrinsic value (type M) firms, so that their equity undervaluation is further reduced.

We argue in Section 2.3.1 that the above partial pooling equilibrium is the one that is most likely to prevail in the equity market after an OMSR program announcement for a very wide range of the per share shortfall cost incurred by firm insiders.[2] Given the nature of this equilibrium, there is room for information production about intrinsic firm values by institutional investors, who will trade on this information in the equity market. The precision of this information produced by institutions is likely to be lower than that of the private information (about their own firm's intrinsic value) held by firm insiders, so that, while information production helps institutions reduce their information disadvantage with respect to firm insiders, it does not eliminate it. Since the information produced by institutions gets reflected in firms' stock prices through their trading, institutional trading will further reduce the information asymmetry faced by the highest intrinsic value (type G) firms (and therefore the undervaluation of their equity). In summary, we hypothesize that there are three complementary mechanisms that serve to reduce the information asymmetry facing firms (and thereby the undervaluation of their equity) in OMSR programs: first, the OMSR program announcement itself; second, actual share repurchases by firms in the open market following the announcement; and third, information production and trading by institutions subsequent to the

[2]In particular, we argue in Section 2.3.1 that a fully pooling equilibrium is likely to prevail only if the per share shortfall cost is close to zero; similarly, a fully separating equilibrium is likely to prevail only if the per share shortfall cost is extremely high. Further, we point out that the implications of these two types of equilibria are contradicted by the empirical evidence on the announcement effects of OMSR programs and the completion rate of actual share repurchases following OMSR program announcements.

OMSR program announcement. We discuss this economic setting in more detail in Section 2.3.1 where we develop a theoretical framework incorporating the above ingredients, based on which we develop testable hypotheses for our empirical analysis (in Section 2.3.2).

We address the following four sets of research questions in the above economic setting. The first set of research questions is regarding the ability of institutions to produce valuable information about a firm prior to its announcing an OMSR program. We address this question empirically by analyzing whether institutional trading prior to an OMSR program announcement has predictive power for the announcement effect of such a program. The second set of research questions is regarding institutions' ability to produce valuable information about a firm immediately after its announcing an OMSR program. The answer to this question gives us insight into the nature of the equilibrium that prevails in the equity market after an OMSR program announcement: clearly, there is no room of information production by institutions in the event this equilibrium is fully separating, thus resolving all information asymmetry upon the announcement of the program. We address this question empirically, in two steps. First, by analyzing the predictive power of institutional trading immediately after an OMSR program announcement for the subsequent long-run performance of the firm's equity. Second, by analyzing whether institutions are able to make abnormal profits by trading in the announcing firm's equity after the announcement of an OMSR program. If institutions indeed have a residual information advantage over retail investors after the announcement of an OMSR program, they should be able to translate this information advantage into abnormal profits by trading in the firm's equity.

The third research question is regarding the interaction between the information production by institutions and the actual share repurchases by firms after an OMSR program announcement. If institutions are able to produce valuable information about the undervaluation of firms' equity after an OMSR program announcement (and buy more equity in more undervalued firms), while more undervalued firms repurchase a larger number of their own shares after the announcement, then institutional net buy should be positively related to the number of shares actually repurchased by announcing firms. We address this question by empirically analyzing the predictive power of institutional trading immediately after an OMSR program announcement for the actual share repurchases made by the announcing firm in the subsequent period. The fourth and final research question is how the information produced by institutions interacts with the private information held by insiders (conveyed to the equity market noisily through the OMSR program announcement and the actual share repurchases of firms) to affect the information asymmetry facing the firm. We address this question empirically by analyzing how institutional trading immediately after an OMSR program announcement affects the change in information asymmetry faced by the firm from before the announcement of an OMSR program to after.

We are able to address the above four research questions directly, given our transaction-level institutional trading data. While we conduct our empirical analysis using trading by our entire sample of institutions around OMSR programs, some of the institutions in our sample may not have the ability (or inclination) to produce information. Therefore, we also conduct our analysis using trading by a subsample of institutions, namely, hedge funds, whose avowed objective is to produce information and to trade on this information in order to generate positive abnormal returns, and who are likely to be less constrained in their trading relative to other institutions in our overall sample.

We make use of a detailed transaction-level institutional trading database provided by Abel Noser Solutions (formerly Ancerno Ltd., or Abel/Noser Corporation) to address the above research questions. Our data includes transactional-level institutional trading data spanning twelve years from January 2003 to September 2011 originated from 868 different institutions, with an aggregate annualized trading principal of around \$9 trillion on all U.S. domestic equity. For an average open-market repurchase event, our sample institutions collectively account for about 12% of the CRSP-reported trading volume within the two-year period surrounding the open-market repurchase announcement. With this dataset, we are able to track institutional trading both before and after an open-market share repurchase announcement. We are also able to compute realized institutional trading profitability net of explicit trading costs (i.e., brokerage commissions) and implicit trading costs (i.e., market impact). Throughout this paper, we use a variable we call "Net Buy" to measure institutional trading. We define Net Buy as the number of shares purchased by institutions minus the number of shares sold by institutions, normalized by the number of shares outstanding.

Our paper provides a number of new results on the effect of information production and trading by institutional investors around OMSR programs on the valuation of the equity of firms announcing such programs, thereby yielding considerable insight into the mechanism through which OMSR programs help to reduce the information asymmetry faced by announcing firms. We organize our empirical tests and results into five parts, corresponding to five different empirical analyses we undertake to address the four sets of research questions outlined above.

First, we study, for the first time in the literature, the informativeness of institutional trading before the announcement of an open-market repurchase program. We find that institutional trading before an open-market repurchase announcement has considerable predictive power for the announcement effect of these programs.[3] This result holds for trading by the entire sample of institutions as well as for trading by hedge funds. A larger extent of net buying by institutional investors prior to OMSR program announcements is significantly associated with a smaller announcement effect. This suggests that institutional investors are indeed able to produce valuable information about the intrinsic values of firms announcing OMSR programs: since the information produced by institutions gets reflected in the equity prices of firms as a result of institutional trading, the undervaluation of firms with greater institutional net buying is reduced to a greater extent prior to the announcement, so that the stock market reaction to OMSR program announcements by such firms will be smaller.

Second, we study the predictive power of institutional trading immediately after OMSR program announcements (over the next month) for the firm's subsequent long-run (one year) performance, again for the first time in the literature. We find that hedge fund trading immediately after an open-

[3]This result is robust to controlling for various variables that have been found in the prior literature to be able to predict announcement effects of open-market share repurchase programs, including prior firm performance and insider trading.

market share repurchase announcement has considerable predictive power for the firm's subsequent long-run stock performance: a 1% increase in hedge fund net buying is associated with about 4.5% increase in the firm's abnormal stock return over the subsequent one-year period. This result is robust to controlling for various variables capturing publicly available information, as well as the extent of trading in the firm's equity by insiders.

Third, we study the realized profitability of institutional trading after OMSR program announcements, using actual transaction prices and net of brokerage commissions, for the first time in the literature. We find that institutions make positive abnormal profits by trading in the firms' equity after the announcement of OMSR programs, even after taking commissions and other trading costs into account. This is the case when the information conveyed by the announcement of an OMSR program is noisier (i.e., when the size of the OMSR program is smaller or when the firm actually repurchases a smaller number of shares subsequent to the announcement). This result holds not only for trading by our hedge fund subsample, but also for trading by our entire sample of institutions. In terms of economic magnitude, over the one-year horizon after an OMSR program announcement, our sample institutions on average realize a risk-adjusted return of 1% when the size of the OMSR program is smaller (i.e., below the sample median), and they realize a risk-adjusted return of 0.8% when the firm actually repurchases less subsequent to the announcement (i.e., below the sample median). The profitability of trading by hedge funds in the same time horizon (one year) is even larger. These results suggest that the information produced by institutional investors (especially hedge funds) after an open-market repurchase announcement that we documented earlier translates into real trading profits when the information conveyed by the OMSR program announcement made by the firm is noisier (i.e., when the size of the OMSR program is smaller or when the firm actually repurchases less subsequent to the announcement).

The above two results, on the predictive power of institutional trading for subsequent stock returns and the realized profitability of institutional trading, respectively, together show that institutions are able to generate a residual information advantage over retail investors even after the announcement of an OMSR program. The fact that institutions are able to produce valuable information about the intrinsic value of firms announcing OMSR programs suggests that the equilibrium prevailing in the equity market after the announcement of such a program is a partial pooling (rather than a fully separating) equilibrium, since there would be no room for information production by institutions in a fully separating equilibrium (where all information asymmetry about firm value is resolved) upon the OMSR program announcement itself.

Fourth, we study the predictive power of institutional trading immediately after OMSR program announcements for the actual share repurchases made by firms in the subsequent period, again for the first time in the literature. We find that institutional trading (by institutions in our entire sample as well as by our subsample of hedge funds) immediately after an OMSR program announcement (either over a month or one quarter horizon) has considerable predictive power for the subsequent actual share repurchases made by the firm: a 1% increase in institutional net buying over the next quarter is associated with about a 3% increase in the firm's actual repurchase over the subsequent two-quarter period.[4] The above result is consistent with the noisy signaling hypothesis that we advance in this paper. In particular, the positive relation we document between institutional trading and actual stock repurchases is consistent with both the above variables serving as complements to OMSR program announcements in reducing the information asymmetry (and therefore the equity undervaluation) of firms making OMSR program announcements.

Fifth and finally, we examine how institutional trading after OMSR program announcements affects the information asymmetry faced by announcing firms in the equity market. We find that institutional trading over the two-quarter period immediately after an open-market repurchase program announcement is associated with a significant reduction in information asymmetry faced by the firm around the OMSR program announcement (i.e., from before the announcement to after) where information asymmetry is measured using four proxies widely used in the literature, namely,

[4]This result is robust to controlling for various variables capturing publicly available information, as well as the extent of trading in the firm's equity by insiders. Unless otherwise mentioned, the economic magnitude refers to that of trading by our entire sample of institutions.

analyst forecast errors, analyst forecast dispersions, coefficient of variation of analyst forecasts, and bid-ask spreads. Thus, greater net buying by institutional investors of the equity of firms announcing OMSR programs is associated with a greater reduction in analyst forecast errors; greater reduction in analyst forecast dispersions; greater reduction in the coefficient of variation of analyst forecasts; and a greater reduction in the bid-ask spreads of the announcing firms' equity. This result provides direct evidence showing that institutional trading subsequent to an OMSR program announcement serves a role complementary to the announcement itself in reducing the information asymmetry facing firms announcing OMSR programs, thus providing further support to our noisy signaling hypothesis.

The rest of the paper is organized as follows. Section 2.2 relates this paper to the existing literature and discusses its contribution relative to this literature. Section 2.3.1 develops a theoretical framework analyzing the complementary role of OMSR program announcements, subsequent actual share repurchases, and institutional trading after OMSR program announcements in reducing the information asymmetry faced by firms; Section 2.3.2 develops testable hypotheses based on the above theoretical framework. Section 2.4 describes the data and various variables used in our empirical analysis. Section 2.5 presents our empirical tests and results. Section 2.6 concludes.

2.2 Relation to the Existing Literature and Contribution

Our paper is related to two strands in the theoretical literature. The theoretical literature closest to our paper is the one modeling institutions as information producers and the implications of such information production for stock repurchases, dividends, and equity issues. Brennan and Thakor (1990) develop a theoretical model assuming that large investors such as institutions have the ability to produce information about firms in the context of their choice of distribution method between open-market repurchases, dividends, and fixed-price tender offers. They, however, do not assume that firm insiders have any private information about intrinsic firm value, and the objective of their paper is to analyze how the presence of both informed investors (such as institutions) and uninformed investors (such as retail investors) among a firm's shareholders affect its choice of payout methods between dividends, OMSR programs, and tender offer repurchases. Nevertheless, our paper may be viewed as empirically analyzing an important assumption of their model, namely, that institutional investors are able to produce information about firms undergoing OMSR programs. Allen, Bernardo, and Welch (2000) develop a theoretical model incorporating the role of institutions as information producers. In their model, institutions have the ability to produce information about the intrinsic values of firms, and are at a greater advantage relative to retail investors in buying shares in firms paying dividends, since dividends are taxed for individuals but untaxed for institutions. In this setting, firms prefer to pay taxable dividends rather than repurchase shares in equilibrium, since paying taxable dividends allows them to reveal their true values to outside investors making use of institutions' ability to produce information about true firm value. Chemmanur and Jiao (2011) develop a model of institutional trading and information production around SEOs. We adapt their economic setting to OMSRs when developing a theoretical framework that incorporates information production by institutions after OMSR program announcements.

The theoretical signaling literature on stock repurchases is also related to our paper. This literature argues that, in an asymmetric information setting, undervalued firms are able to credibly and fully separate themselves from overvalued firms using stock repurchases: see, e.g., Ofer and Thakor (1987), Constantinides and Grundy (1989), Vermaelen (1981), or Persons (1994).[5] Oded (2005) points out that, unlike Dutch auction or fixed-price tender offers, OMSR programs do not pre-commit firms to acquire shares, and that many firms buy back only a fraction of the dollar value announced, thus calling into the question the ability of OMSR programs to signal true firm value.

[5]McNally (1999) develops a signaling model of OMSR programs in a setting similar to that of Leland and Pyle (1977). He assumes that entrepreneurs are risk averse and do not tender their own shares, so that a share repurchase increases entrepreneurs' proportionate equity holdings in the firm, triggering an increase in equity value due to the signaling effect of such an increase (similar to the signaling effect of entrepreneurs' equity holdings in Leland and Pyle (1977)). However, he assumes that the announcement of the target number of shares in an OMSR program is a "commitment to action," thus assuming away the difference between OMSR programs and the other two forms of repurchase (Dutch auction and fixed-price tender offers) that exists in practice.

He, however, goes on to develop a signaling model of OMSR programs where firms face a trade-off between the long-run gains from the informed trading that the option to repurchase shares creates and the short-run costs from the market's accounting for this adverse selection.[6] Under this trade-off, only good firms announce open-market repurchase programs, so that the announcement of an OMSR program acts as a credible signal (i.e., yielding a fully separating equilibrium).[7] In contrast to the above literature, we argue that OMSR program announcements are not able to fully signal true firm value: i.e., the equilibrium prevailing in the equity market after an OMSR program announcement is not a fully separating equilibrium but a partial pooling equilibrium. Further, the empirical evidence we document here, that institutions are able to produce valuable information about firms announcing OMSR programs and generate abnormal profits from trading on this information, also contradicts the notion that OMSR program announcements fully convey firm insiders' private information to the equity market: there is no room for information production by institutions in a fully separating equilibrium.[8]

Our paper is also related to the large empirical literature on stock repurchases in general and OMSR programs in particular. A number of early papers show that the prices of firms that announce a stock repurchase program increase significantly in the short run (e.g., Dann (1981); Vermaelen (1981)) and in the long run (e.g., Ikenberry, Lakonishok, and Vermaelen (1995)). Comment and Jarrell (1991) study the relative signaling power of Dutch-auctions, self-tender offers and open-market repurchases, and show that, while the announcement effects of OMSR programs are positive, they provide weaker signals of stock undervaluation (weaker announcement effects) compared to the other two forms of repurchase. Comment and Jarrell (1991) also point out that larger OMSR program

[6]See also Ikenberry and Vermaelen (1996) for a discussion of this option under symmetric information.

[7]This result, however, holds only under the rather strong assumption that the stock value distribution of a good (higher intrinsic value) firm has a higher variance than that of a bad (lower intrinsic value) firm, so that the option to repurchase the shares of a good firm is more valuable than the corresponding option of a bad firm.

[8]Bayar, Chemmanur, and Liu (2014) develop a theoretical analysis of a firm's choice between dividend payments and share repurchases to pay out cash (in a setting of heterogeneous beliefs between firm insiders and outsiders as well as among outsiders). They also develop a theoretical rationale for the positive long-run stock returns following stock repurchases that has been documented in the empirical literature. See Allen and Michaely (2003) for an extensive review on the payout policy literature.

announcements are viewed as stronger signals. A number of papers have also studied actual share repurchases in OMSR programs and compared the number of shares actually repurchased relative to the target number announced in OMSR programs. Stephens and Weisbach (1998) document that for OMSR programs announced between 1981 to 1990, firms acquire on average 74 to 82 percent of the shares announced as repurchase targets within three years of the repurchase announcement. Ben-Rephael, Oded, and Wohl (2014) show that disclosure of firms' actual repurchase activity following OMSR program announcements lead to a positive and significant abnormal stock return, consistent with actual share repurchases contributing to a reduction in the residual information asymmetry facing firms even after an OMSR program announcement. Busch and Obernberger (2017) find that actual share repurchases in OMSR programs increase stock price efficiency and the information content of stock prices. Despite the above large body of empirical evidence, there has been no empirical analysis of the role played by information production and trading by institutions in mitigating the residual information asymmetry faced by firms after the announcement of OMSR programs in the existing literature: this is our focus here.[9,10]

Our paper makes several important contributions to the literature at a conceptual as well as at an empirical level. First, ours is the first paper in the literature to propose a noisy signaling hypothesis of OMSR programs. Thus, we are the first to argue that, in contrast to the existing literature which has theoretically demonstrated a separating equilibrium after share repurchase programs in general, the equilibrium in the equity market after an OMSR program announcement is likely to be a partial pooling equilibrium, where (in a setting with a continuum of types or a discrete type setting with three or more types) the highest firm types pool by announcing an OMSR program while the lowest types do not announce such a program. In this context, we argue that there is room for some

[9]Two other contemporaneous papers also study trading by institutional investors around stock repurchases. DeLisle, Morscheck, and Nofsinger (2014) document that institutional investors are net sellers during share repurchases. In a similar spirit, Huang and Zhang (2013) show that institutions sell after share repurchase announcements. Neither of these papers analyze and test hypotheses regarding the information production role of institutional investors, which is our focus here.

[10]Our paper is also distantly related to the empirical literature analyzing institutional trading around corporate events other than stock repurchases: see, e.g., Gibson, Safieddine, and Sonti (2004) and Chemmanur, He, and Hu (2009), who empirically analyze institutional trading around SEOs.

information transmission from firm insiders to equity market investors through an OMSR program announcement, even in the absence of a commitment by the firm to buy back the entire amount of shares announced, as long as the firm or its insiders suffer a moderate shortfall cost per share: i.e., the firm incurs such a cost if the actual number of shares repurchased falls short of the target number of shares announced. Finally, we conjecture that, in such a partial pooling equilibrium, there are two mechanisms that play a role complementary to OMSR program announcements in further mitigating the residual information asymmetry faced by the firm even after the OMSR program announcement. The first such mechanism is actual share repurchases by firms after OMSR program announcements: we argue that more undervalued firms repurchase a larger number of shares after OMSR program announcements, thereby conveying further information about intrinsic firm value to the equity market. The second such mechanism is information production by institutions and trading by them making use of this information after OMSR program announcements. We argue that, based on their information production, institutions buy more equity in more undervalued firms, and institutions' information getting reflected in stock prices as a result of their trading further reduces the residual information asymmetry facing firms after OMSR program announcements.

The second contribution made by our paper lies in testing the implications of the above noisy signaling hypothesis of OMSR program announcements for information production and trading by institutions. The results of our empirical analysis provide considerable support for the noisy signaling hypothesis. First, the fact that institutions are able to produce valuable information subsequent to OMSR program announcements (as evidenced by the predictive power of institutional trading and the realized profitability of such trading) provides support for the notion that the equilibrium prevailing in the equity market is a partial pooling rather than a fully separating equilibrium: there would be no room for information production by institutions in a separating equilibrium, since, in such an equilibrium, all information asymmetry is resolved upon the announcement of the OMSR program itself. Second, the positive relationship that we document between institutional net buying and actual share repurchases by firms provides further support for the noisy signaling hypothesis.

This positive relationship is likely to be induced by the fact that firms that are more undervalued after OMSR program announcements repurchase more of their own shares, while institutions net buy a larger number of shares after OMSR program announcements in firms that they believe to be more undervalued. Third, we provide direct evidence supporting the complementary role of information production and trading by institutions in reducing the residual information asymmetry facing firms announcing OMSR programs, by showing that the reduction in information asymmetry facing firms from before an OMSR program announcement to after such an announcement is greater when net buying by institutions in the firm's equity immediately after the announcement is greater.

Our paper also makes a third contribution to the literature by documenting the predictive power of institutional trading prior to OMSR program announcements for the announcement effects of such programs. This, along with the other empirical results discussed earlier, allows us to empirically confirm a crucial assumption made by the Brennan and Thakor (1990) model: we are able to show that institutional investors are indeed able to produce credible information about the intrinsic values of firms around OMSR programs.

2.3 Theoretical Framework and Testable Hypotheses

2.3.1 Theoretical Framework

In this section we briefly develop a theoretical framework that allows us to analyze the noisy signaling role of OMSR programs, and the complementary role of information production and trading by institutions in conveying information from firm insiders to uninformed outsiders (e.g., retail investors) in the equity market. We will use this theoretical framework to develop testable hypotheses in the next section.

Consider a situation where the insiders of a firm, having private information about its intrinsic value, are deciding whether or not to undertake an OMSR program. If insiders choose to announce an OMSR program, it may convey a signal to outsiders that the firm's equity is undervalued relative to its intrinsic value as assessed by firm insiders (based on their private information): whether the signal is fully revealing or partially revealing, or whether there will be any information content to this signal at all, will depend upon the nature of the equilibrium that prevails in the equity market after the OMSR program announcement, as we discuss below. In other words, the announcement of an OMSR program may convey firm insiders' private information only partially to outsider shareholders, thereby reducing (but not necessarily eliminating) the undervaluation of the firm's equity relative to its intrinsic value. The fact that, even after an OMSR program announcement, there may be residual undervaluation of the firm's equity leaves room for information production by institutions about the firm's intrinsic value: i.e., information production by institutions getting reflected in the stock price may play a role (complementary to the noisy signal conveyed by the OMSR program) in reducing the undervaluation of the announcing firm's equity.

We now analyze more precisely the nature of the equilibrium that prevails in the equity market upon the announcement of an OMSR program. The setting we study is the following. There are three types of firms: Good (type G) with intrinsic value V_G; Medium (type M) with intrinsic value V_M; and Bad (type B) with intrinsic value V_B; $V_G > V_M > V_B$. While firm insiders know the true type of their own firm, outsiders know only the probability distribution across firm types: they assess that any firm in the equity market is a type G with probability γ_G, type M with probability γ_M, and type B with probability γ_B, $\gamma_G + \gamma_M + \gamma_B = 1$. Since outsiders cannot fully distinguish between the three types of firms, the share price of any firm prior to an OMSR announcement will be the pooling value across the three types of firms: i.e., it will be $\gamma_G V_G + \gamma_M V_M + \gamma_B V_B$.

The timeline of events (depicted in Figure 2.1) is the following. At time 0, a firm chooses whether or not to announce an OMSR program and announces it if it finds it optimal to do so. If the firm chooses to actually repurchase any of the shares authorized in the OMSR program, it does so between time 0 and time 1, with the actual repurchases completed at time 1. Investors in the equity market (both institutional and retail investors) come to know the number of shares (if any) actually repurchased by the firm at time 1.[11]

We assume that the firm and/or its top managers suffer a per share reputation cost arising from any shortfall in the number of shares actually repurchased relative to the number of shares authorized in the OMSR program announcement: the aggregate reputation cost suffered by a firm (or its top managers) is given by the product of the number of shares by which the firm's actual repurchase falls short of the number announced (authorized) in the OMSR program and the shortfall cost per share. The firm incurs this aggregate shortfall cost (if any) at time 1, soon after the actual number of shares repurchased becomes known. We do not take a position on the magnitude of this per share shortfall cost: we argue that the nature of the equilibrium in the equity market will depend upon whether this magnitude is large, moderate, or small.

In the long-run (time 2), the true value of the firm becomes revealed exogenously (i.e., the information asymmetry between firm insiders and outsiders is eliminated at time 2, as the firm's operating performance becomes known to outsiders over time). As we discuss in more detail below, institutional investors may produce information (at a cost) about the firm's intrinsic value between the announcement of an OMSR program (time 0) and the completion of actual share repurchases (if any) by the firm (and the public disclosure of the number of shares repurchased) at time 1.

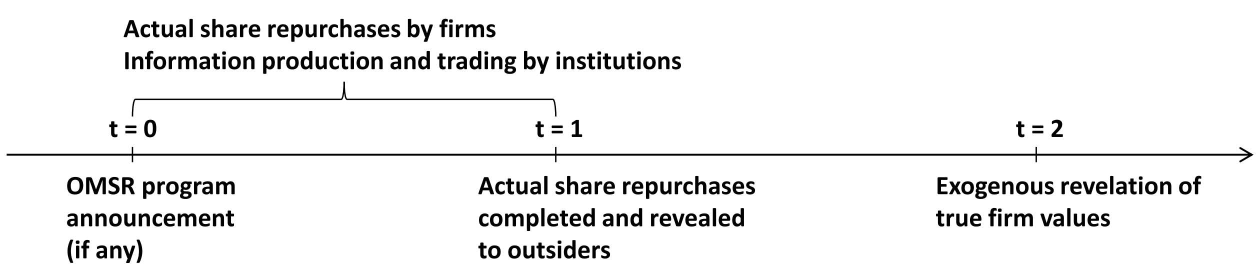

Figure 2.1: Timeline of Events

The objective of each firm in deciding whether or not to announce an OMSR program at time

[11]Beginning from 2004, U.S. firms are required to make quarterly disclosures of actual share repurchases and average prices paid.

0, and the actual number of shares to repurchase, is to maximize a weighted average of its equity values in the short run (time 0), medium run (time 1), and the long-run (time 2), net of any reputation cost incurred by the firm at time 1 due to the shortfall in shares actually repurchased. The weights placed on the valuation at each date is determined exogenously (not alterable by firm managers).

Outsiders in the equity market consist of two types of investors.[12] The first type of investors are institutional investors, who have the ability to produce noisy information about the firm (at a cost). The precision of information produced by institutions is lower than that of the private information held by firm insiders, so that, while information production helps institutions reduce their information disadvantage with respect to firm insiders, it does not eliminate it. We assume that institutions trade on the information they produce: they buy shares that they believe to be undervalued based on the information they have produced. The second type of investors are retail investors, who do not have any ability to produce information about the intrinsic value of the firm, and are therefore at a disadvantage with respect to both institutions and insiders. Retail investors are essentially liquidity traders in the equity market in the economic setting we study here, similar to their role in market microstructure models such as Kyle (1985). The price of the firm's stock in the equity market is set by a market-maker who is uninformed to begin with, but who sets the stock price to break even (after observing the aggregate order flow of trades in the firm's equity), again similar to the price-setting rule in market microstructure models. The aggregate order flow observed by the market-maker in the equity market in our setting comes from three sources: trading by institutional investors; actual share repurchases by firms; and trading by retail investors (uninformed liquidity traders). While the market-maker cannot fully separate informed and uninformed trades, the price of the firm's equity will reflect, to some degree, the information held by institutional investors as well as that contained in the actual repurchases made by firms. The information flow between firm

[12] We adapt the setting of Chemmanur and Jiao (2011) to analyze an equity market with information production by institutional investors around OMSR program announcements, and how this information gets reflected in stock prices. Chemmanur and Jiao (2011), however, focus on theoretically analyzing the implications of institutional trading around seasoned equity offerings. In the interest of conserving space, we choose not to develop a formal theoretical model here to analyze institutional trading around OMSR program announcements, but instead adapt the theoretical analysis of Chemmanur and Jiao (2011) to the stock repurchase setting.

insiders, institutional investors, retail investors and the resulting determination of the firm's stock price (by the market-maker) in the economic setting we postulate here is depicted in Figure 2.2.

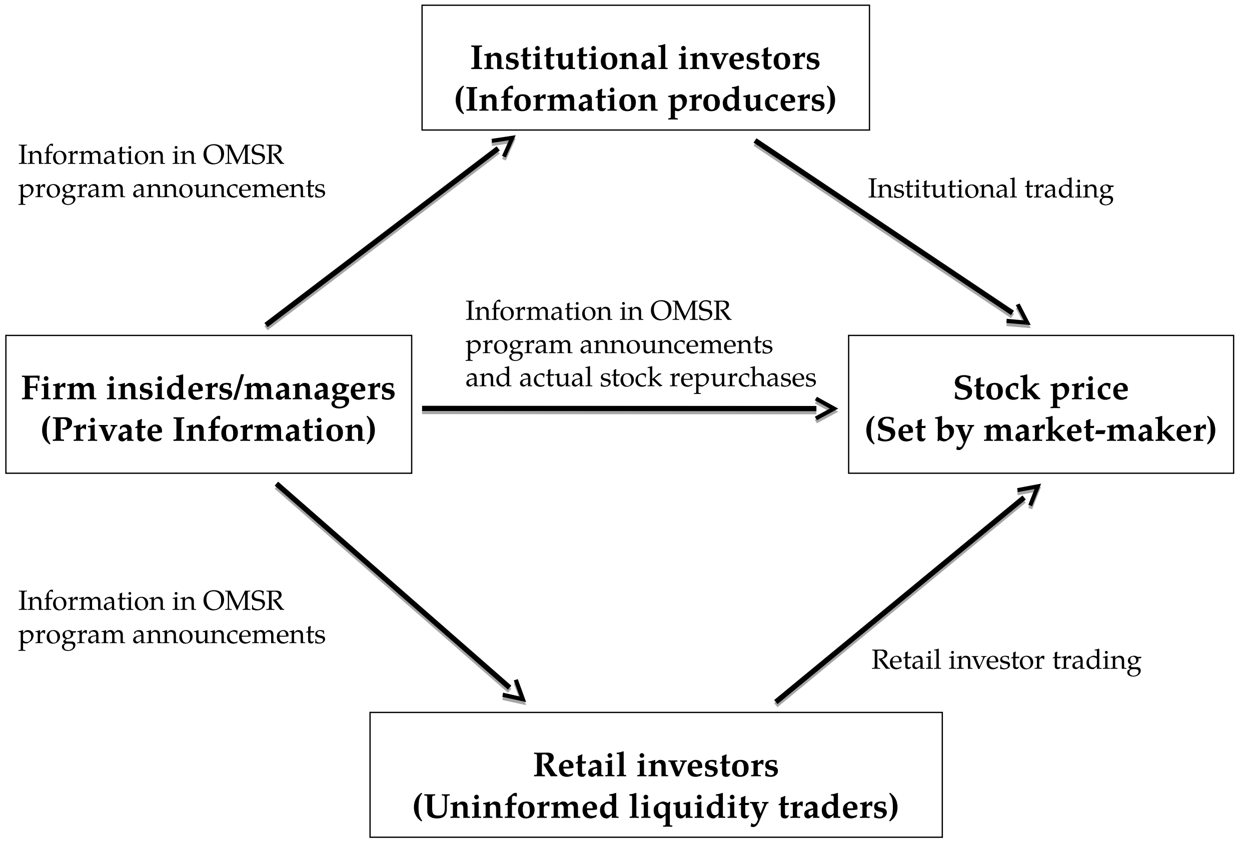

Figure 2.2: Information flow between firms, institutions, uninformed investors, and the stock price

We now characterize the equilibrium in the above setting as a function of the magnitude of the repurchase shortfall cost incurred by the firm. We describe three possible equilibria. We start with the equilibrium which prevails when the per share shortfall cost is moderate.

Equilibrium One (*Moderate repurchase shortfall cost per share*):

In this equilibrium, only a type G or a type M firm announces an OMSR program at time 0.[13]

[13] The number of shares authorized (and announced) in the OMSR program by the type M will be the same as that by the type G, since, otherwise it will reveal its true type. In other words, we an think of the type G as determining the number of shares to be announced in the OMSR program and the type M mimicking it by announcing the same number of shares.

A type B firm will not announce such a program since its incentive compatibility (truth-telling) condition is satisfied. In other words, it is optimal for a type B firm to reveal its true type by refraining from making such an OMSR announcement. The trade-off faced by a type B firm is the following. On the one hand, if the type B firm makes an OMSR program announcement, it can prevent a drop in its share price from the pooled value across types to its true value. On the other hand, in this case, the type B firm will not buy back any shares between time 0 and time 1 even if it announces an OMSR program, since its shares will be overvalued relative to intrinsic value between time 0 and time 1, so that actually repurchasing shares will be prohibitively costly for the firm, resulting in the type B firm having to incur a high aggregate shortfall cost at time 1. Given that the firm's true value is revealed exogenously at time 2 (so that its equity value will equal its true value at time 2 even if it announces an OMSR program at time 0), it can be shown that the value of the type B firm's objective will be strictly lower (for moderate values of the shortfall cost) if it announces an OMSR program at time 0 compared to the case where it refrains from making any such announcement. In summary, the type B firm does not announce an OMSR program at time 0, thereby revealing its true type.[14]

Given the type B firm's behavior in the equilibrium, outside investors in the stock market recompute a firm's stock value upon an OMSR program announcement as a weighted average of the intrinsic values of the type G and the type M firms, the weights being the Bayesian updated probabilities of the firm being of type G and type M respectively (conditional on such an announcement). The announcement effect (abnormal stock return) of an OMSR program is therefore positive, since the equity value of the firm upon the announcement is strictly higher than the fully pooling value prevailing before the announcement. Subsequent to an OMSR program announcement, the type G firm continues to be undervalued, though less than before the announcement; the type M firm, however, is overvalued. Given that its shares continue to be undervalued, the type G firm

[14]It is straightforward to write down the formal incentive compatibility conditions of the type B and type M firm at time 0, which are consistent with the type B choosing not to announce an OMSR program (thus separating from the other two types) while the type M chooses to announce such a program (thus pooling with the type G). We choose not to present these here due to space limits.

repurchases the entire number of shares announced in the OMSR program between time 0 and time 1.

In contrast to the type G, the type M firm repurchases only a certain fraction of the number of shares announced in the OMSR program since its shares are overvalued, resulting in actual share repurchase being costly for that firm. The number of shares repurchased by the type M firm between time 0 and time 1 will therefore reflect the trade-off faced by that firm between buying back its overvalued shares, and incurring the repurchase shortfall cost at time 1.[15] Given that the type G and type M firms actually repurchase different numbers of shares in the open market, once the actual number of shares repurchased is revealed to outside investors at time 1, the stock prices of the two types of firms will signal their true intrinsic values (with the type G firm's stock price going up (since its true value is now fully revealed) and the type M firm's price going down). At time 2 (long-run), the information asymmetry facing all firms is resolved exogenously, so that their equity market value will be equal to their intrinsic values at this date.

The role of institutions in the above equilibrium is that of information production and trading on the information produced between time 0 and time 1. In the above equilibrium, institutions are able to produce noisy information that allow them to distinguish partially between a type M firm (overvalued) and a type G firm (undervalued). Institutions then buy equity in firms that they believe to be undervalued and sell equity in those they believe to be overvalued. Since the information contained in institutional trading gets reflected in the stock price through the inference and price-setting process of the market-maker (see Chemmanur and Jiao (2011) or Kyle (1985) for details), institutional trading between time 0 and time 1 plays a role complementary to OMSR program announcements and actual share repurchases by firms in reducing the undervaluation of the equity of a type G firm (and in reducing the overvaluation of the equity in a type M firm).[16]

[15] Thus, the actual number of shares repurchased by the type M firm between time 0 and time 1 will be a function of the repurchase shortfall cost.

[16] Trading by institutions using the information they have produced and its effect on the stock price of the two firm types (type G and type M) may, in turn, affect the number of shares actually repurchased by them. However, the behavior of two firm types we specified under equilibrium remains qualitatively unchanged. Thus, as long as the

Further, institutional buying of shares will be correlated with actual share repurchases by firms, since institutions buy equity in undervalued (type G) firms and, as discussed above, a type G firm actually repurchases more of its own equity than a type M firm between time 0 and time 1.

Equilibrium Two (*Very low repurchase shortfall cost per share*):

In this equilibrium, all three types pool by announcing an OMSR program.[17] There is therefore no announcement effect at time 0: the stock price of any firm announcing an OMSR program remains the same as before the announcement. Between time 0 and time 1, the type B firm does not repurchase any shares, and suffers a low aggregate shortfall cost at time 1. The type B finds it optimal to announce an OMSR program at time 0 since this enables it to keep its equity value at the overvalued level at time 0 by pooling with the type G and type M, while incurring only a low aggregate shortfall cost at time 1, once it is revealed that it did not repurchase any of the shares announced in the OMSR program. Consequently, the value of the type B firm's objective is higher in this equilibrium if it announces an OMSR program but does not repurchase any shares between time 0 and time 1 (in other words, its truth-telling condition at time 0 is not satisfied when the per share shortfall cost is low). The type G firm actually repurchases all the shares it announced in the OMSR program at time 0, since its shares continue to be undervalued after the announcement (between time 0 and time 1). The type M firm may repurchase all shares announced in the OMSR program (if its shares are undervalued after the OMSR program announcement between time 0 and time 1) or only a fraction of the shares it announced (if the pooling value prevailing after the OMSR program announcement is above its intrinsic value, so that its equity is overvalued). In the latter case, its trade-off in determining the actual repurchase fraction is similar to that discussed above under equilibrium one. At time 2, true firm values are exogenously revealed, so that the equity values of all three firm types equal their intrinsic values.

---

undervaluation of the type G firm is not fully eliminated due to institutional trading, the firm will repurchase all the shares it has announced in the OMSR program; similarly, while the precise number of shares that the type M firm repurchases between time 0 and time 1 may change due to the reduction in overvaluation of the type M firm's equity due to institutional trading, its equilibrium behavior remains qualitatively unchanged.

[17] The type M and type B firms pool with the type G firm by announcing the same number of shares as in the type G firm's OMSR program announcement, since they will otherwise reveal their true types.

Note that this equilibrium is inconsistent with the empirical evidence documented by the existing literature (as well as that in this paper), since this evidence shows that OMSR programs have a positive announcement effect (see, e.g., Table 1 of this paper or previous papers such as Comment and Jarrell (1991) or Vermaelen (1981)).

Equilibrium Three (*Very high repurchase shortfall cost per share*):

In this case, the equilibrium is fully separating, so that each type of firm fully reveals its type at time 0. In this equilibrium, the type G firm announces an OMSR program at time 0 and repurchases the announced number of shares between time 0 and time 1. The type M firm also announces an OMSR program at time 0, but for a smaller number of shares; it repurchases this smaller announced number of shares between time 0 and time 1, and therefore avoids incurring any repurchase shortfall cost. The type B firm does not announce any OMSR program at time 0, since, given the large per share shortfall cost assumed here, any valuation benefit arising from pooling with the type M or the type G at time 0 is overcame by its cost of buying back its own overvalued shares (if it chooses to actually repurchase some shares between time 0 and time 1) or its aggregate shortfall cost it incurs at time 1 (if it choose not to actually repurchase the number of shares announced at time 0). In other words, the value of the type B's objective if it reveals its true type at time 0 itself is greater than if it attempts to pool with the type M or type G. The announcement effect of an OMSR program is positive for both the type G and type M firm (and therefore positive on average for all firms announcing OMSR programs).[18]

Further, since this equilibrium is fully separating at time 0 (i.e., all information asymmetry is resolved upon announcement), the magnitude of the announcement effect in an OMSR program will be as high as in other types of repurchases, such as fixed-price tender offers. This prediction is clearly inconsistent with the existing empirical literature: see, e.g., Comment and Jarrell (1991) and Vermaelen (1981), who compare the signalling power of fixed-price tender offers, Dutch auctions, and OMSRs, and conclude that OMSRs have the smallest announcement effect. Further, this

[18]Here we are assuming that V_B is sufficiently smaller than V_M, and γ_G, γ_M, and γ_B are such that the type M firm is undervalued at the pooling price prevailing before the OMSR program announcement.

equilibrium does not allow any room for costly information production by institutions after an OMSR program announcement, since the equilibrium is fully separating, so that all information asymmetry is resolved upon the OMSR program. This, in turn, implies that, if this equilibrium prevails, institutions are unlikely to have an information advantage over retail investors in the equity market after an OMSR program announcement, so that there will be no meaningful reward to institutions engaging in costly information production.

2.3.2 Testable Hypotheses

In this section, we use the theoretical framework developed in Section 2.3.1 to develop testable hypotheses to analyze the predictive power of institutional trading around open-market stock repurchases for the announcement effect of stock repurchases; actual shares repurchased (as against the authorized repurchase in the OMSR program announced); the long-run stock return performance of the firm's equity subsequent to the announcement of an open-market repurchase; and finally, the abnormal profits realized by institutional investors (net of all transaction costs) by trading in the equity of firms subsequent to their announcement of the OMSR program. We also develop testable hypotheses for the relationship between institutional trading immediately after an OMSR program announcement and the change in the information asymmetry faced by the firm from before the announcement of the OMSR program to after.

We rely on equilibrium one discussed in Section 2.3.1 to develop our testable hypotheses. This is because we view this equilibrium as the most plausible one in practice, since, as we discussed in the previous section, the other two equilibria are inconsistent with the empirical evidence on the announcement effect of OMSR programs. Since the real world is continuous in terms of intrinsic firm values, we use a continuous analog of the three type model characterized in Section 3 to develop testable hypotheses: in other words, we can think of type G, type M, and type B firms that we discussed in Section 2.3.1 as intervals of continuous firm types behaving differently at different points in time, as discussed in equilibrium one in Section 2.3.1. Finally, equilibrium one is even more likely to prevail in this continuous type version of our theoretical framework than in the discrete type framework, since, even if the per share repurchase shortfall cost that we assume in Section 2.3.1 is rather small in practice, there will always be a set of firm types that can be identified as behaving similar to the type B firm in equilibrium one (partial pooling equilibrium) that we discussed in section 2.3.1.[19]

Our first hypothesis deals with the relationship between institutional trading prior to an OMSR program announcement and the announcement effect of such a program. Before the announcement of an OMSR program, the equity of all firms will be priced at the pooling value across firm types (as we discussed in Section 2.3.1). Consider now the scenario where institutional investors produce information about intrinsic firm values and trade in the firm's equity prior to the OMSR program announcement. The stock price will reflect the additional information contained in trading by institutions, with the price of the firm's equity falling lower if the net buy by institutions (number of shares bought minus number of shares sold) is negative, while it will rise higher if their net buy is positive. In other words, effect of the information produced pre-OMSR program announcement by institutions getting reflected in stock prices is to reduce the extent of pooling across firm types, reducing the undervaluation of higher type firms while reducing the overvaluation of lower type firms.[20] At this point, if the firm announces an OMSR program, stock market investors will further

[19]It is also worth noting that the completion rate of actual OMSR programs announced in practice is broadly consistent with firms facing a moderate per share cost of not actually repurchasing the number of shares announced in the OMSR program (as in our equilibrium one). In particular, the evidence in our sample is that, on average, firms actually repurchase 80.11 percent of the shares announced in the OMSR program in the one-year period after the announcement (the evidence documented in various other papers in the existing literature is broadly similar). This is inconsistent with firms behaving as if there is no repurchase shortfall cost (as in equilibrium two) or a very high shortfall cost (as in equilibrium three). In the former scenario, we would expect the number of shares actually repurchased as a fraction of shares announced in the OMSR program to be much smaller; in latter scenario, we would expect almost all firms announcing OMSR programs to repurchase one hundred percent of the shares announced in the program. We do not observe either of the above scenarios in practice.

[20]We do not incorporate information production and trading by institutions prior to the announcement of an OMSR program in the theoretical framework developed in Section 2.3.1. However, it is easy to incorporate the effect of this information production and trading into our theoretical framework by introducing an additional date prior to the announcement of an OMSR program, namely, date -1, with pre-OMSR program institutional information production and trading occurring between time -1 and time 0. As we discuss in the main text, the effect of pre-announcement information production and trading by institutions is to reduce the extent of pooling across types that prevails at time 0, so that the pooling across types prevailing at time 0 (the date of the OMSR program announcement) would be lower

positively update the value of the firm's equity, knowing that higher intrinsic value firms are more likely to announce an OMSR program than lower intrinsic value firms, and that the decision to repurchase (or not) is made by firm insiders who have private information about intrinsic firm value. This means that the magnitude of the announcement effect, which will reflect the difference in the firm's stock price from immediately before the repurchase announcement to immediately after, will be negatively related to net buying by institutional investors. The intuition here is that if the institutional net buying prior to the OMSR program announcement is larger, the reduction in the undervaluation of higher types that has already occurred prior to the OMSR program announcement is more, so that the stock market reaction to the OMSR program announcement itself will be smaller. This will be the first hypothesis that we test here (**H1**).

We now turn to trading by institutions subsequent to the announcement of an open-market repurchase program. If the announcement of an OMSR program conveys firm insiders' private information to the equity market, the price of the firm's stock immediately after the repurchase announcement will reflect this information. However, as we discussed under equilibrium one in Section 2.3.1, the OMSR program announcement may only be a noisy signal of firm insiders' private information. If this is indeed the case, there is room for further information production by institutions about intrinsic firm value even after an OMSR program announcement, giving institutional investors a residual information advantage over retail investors even after the announcement of an OMSR program. In this case, trading by institutions after an OMSR announcement will have predictive power for the firm's future stock returns. This is the second hypothesis that we test here (**H2**).

If (as we postulated under **H2**) institutions indeed have an information advantage over retail investors when they trade in a firm's equity subsequent to its OMSR program announcement, we would expect this information advantage to translate into abnormal profits realized by institutions. This is therefore the next hypothesis that we test here (**H3**). The residual information advantage

than it would otherwise be in the absence of such information production and trading by institutions. In other words, while at time 0 three types: type G, M, and B pool in our theoretical framework (discussed in Section 2.3.1), at time -1 even lower intrinsic value firm types than the type B may be pooling with higher type firms, so that the extent of undervaluation of higher type firms may be even more severe at time -1 than at time 0.

of institutional investors over retail investors will be greater as the information conveyed by the repurchase announcement itself is weaker or more noisy: i.e., if the repurchase program announced is smaller (as a fraction of total shares outstanding) or if the number of shares actually repurchased is smaller.[21] This, in turn, implies that the abnormal profits made by institutions from trading in the firms' equity after an OMSR program announcement will also be greater for smaller repurchase programs. This is the next hypothesis that we test here (**H4**).

We now turn to developing a testable hypothesis regarding the relation between institutional trading after an OMSR program announcement and the number of shares actually repurchased by the firm. To develop this hypothesis, recall first from our discussion of equilibrium one in Section 2.3.1 that higher intrinsic value firms will actually repurchase a larger number of shares (out of the total number announced in the OMSR program). If, in the above setting, institutions are able to produce information about the extent of undervaluation of firms' equity (i.e., about intrinsic firm value), and buy more of the equity in firms where the extent of the undervaluation is greater (i.e., in higher type firms), then the extent of institutional net buying immediately after an OMSR announcement will be positively related to the amount of shares actually repurchased by the firm. Thus, a greater net buy of the firms' equity by institutions after an OMSR program announcement will be positively related to the actual repurchases made by the firm in the subsequent period (**H5**).

Finally, we examine how institutional trading after an OMSR program announcement affects the information asymmetry faced by a firm. Clearly, if OMSR program announcements serve as noisy signals of firm insiders' private information to outsiders in the equity market, then the information asymmetry faced by the firm will be reduced: i.e, the extent of information asymmetry faced by the firm in the equity market subsequent to an OMSR announcement will be lower than that before the repurchase announcement. The question we examine here, however, is the effect of the interaction between the signal conveyed by the OMSR program announcement and the information conveyed by institutional trading immediately after the announcement of the repurchase program on the change

[21] Comment and Jarrell (1991) point out that larger OMSR program announcements act as stronger signals, based on their announcement effects.

in information asymmetry facing the firm. In particular, the reduction in the information asymmetry facing the firm will be greater when the noisy signal conveyed by the repurchase announcement and the information conveyed to the equity market by institutional trading reinforce each other (which will be the case when the institutional net buy immediately after the repurchase announcement is positive). On the other hand, the reduction in information asymmetry facing the firm will be smaller when the noisy signal conveyed by the repurchase announcement and the information conveyed to the equity market by institutional trading oppose each other (which will be the case when the institutional net buy immediately after the repurchase announcement is negative). In summary, the reduction in information asymmetry facing the firm following the announcement of an OMSR program will be positively related to the institutional net buy immediately after the repurchase program announcement (**H6**).

2.4 Data and Summary Statistics

2.4.1 OMSR Program Data

The data on OMSR programs in this study comes from several sources. Our initial sample of OMSR program announcements from January 2004 to December 2010 comes from the SDC Platinum Database of Mergers and Acquisitions. We then exclude announcements such that the repurchase may be executed through tender offer, private negotiation, or Dutch auction.[22] If a firm makes multiple OMSR program announcements in the same calendar year, we only keep the first announcement. We also require that accounting information from Compustat and stock return information from CRSP are available for the firms in our OMSR data.

Our data on U.S. firms' actual share repurchases comes from Quarterly Compustat, which is made available by the regulatory changes to Rule 10b-18 of the Securities Exchange Act of 1934 in

[22]The purpose is to eliminate repurchase programs that may be executed through a combination of methods (e.g., open-market and private negotiation).

2003.[23] Beginning from 2004, U.S. firms are required to make quarterly disclosures of actual share repurchases and average prices paid. We then match this actual repurchase data from Quarterly Compustat with the data on OMSR announcements we obtained from SDC.

Table 2.1 reports summary statistics of our OMSR data. We have about 3,000 open-market repurchase programs announced from January 2004 to December 2010. The average OMSR program size, defined as the dollar amount value of the OMSR program normalized by the market capitalization of the firm, is 7.94%, consistent with prior studies in the literature (e.g., Peyer and Vermaelen (2009)). We find a significant 1.74% average abnormal return in the 3-day window around an OMSR announcement, which is also consistent with the empirical findings in the literature (e.g., Babenko, Tserlukevich, and Vedrashko (2012)). Over one-year period following an OMSR announcement, our sample firms' actual repurchases on average account for about 80.11% of the OMSR program size announced. This is largely consistent with prior findings that firms complete a significant portion of the repurchase programs within the one-year period after the announcement (e.g., Stephens and Weisbach (1998)).

2.4.2 Institutional Trading Data

We obtain transaction-level institutional trading data from Abel Noser Solutions, a leading execution quality measurement service provider for institutional investors. The data are similar to those used by several microstructure studies on institutional trading costs, for example, Keim and Madhavan (1995), Conrad, Johnson, and Wahal (2001), and Jones and Lipson (2001). To the best of our knowledge, this is the first paper to use institutional trading data to study institutional investors' trading behavior around OMSR programs.

The data cover equity trading transactions by a large sample of institutions from January 2003 to

[23]Earlier studies (e.g., Stephens and Weisbach (1998), Fama and French (2001), and Grullon and Michaely (2002)) have used a variety of other CRSP- and Compustat-based measures to estimate actual share repurchases by U.S. firms. These estimations invariably suffer from different measurement biases. For a detailed discussion of these measures, see Banyi, Dyl, and Kahle (2008).

September 2011. For each transaction, the data include the date of the transaction, the stock traded, the number of shares traded, the dollar principal traded, commissions paid by the institution, and whether it is a buy or sell by the institution. The data are provided to us under the condition that the names of all institutions are removed from the data. However, identification codes are provided enabling us to separately identify all institutions. Sample institutions are either investment managers or plan sponsors. Within investment managers, hedge funds are identified by merging management companies in Abel Noser with a list of hedge funds provided by Thomson Reuters. Please see the Appendix for details of this matching algorithm.

Table 2.2 reports summary statistics of our institutional trading data. We have 868 institutions in our sample, with 372 of them being investment managers and 496 of them being plan sponsors. Within the group of investment managers, 162 of them are identified as hedge fund companies (including institutions that have both hedge funds and non-hedge fund businesses). In aggregate, these institutions have an annualized trading volume of around 304 billion shares and an annualized trading principal of around $9 trillion. In association with these trading activities, our sample institutions in aggregate incur an annualized commission expense of about $7.5 billion. If we consider a two-year trading horizon surrounding the OMSR announcement dates in our OMSR data, on average (for each OMSR event), our sample institutions in aggregate execute about 24,510 transactions, with a trading principal of about $3.5 billion, and account for about 12% of the trading volume reported by CRSP.

2.5 Empirical Tests and Results

2.5.1 The Relation between Pre-Announcement Institutional Trading and OMSR Program Announcement Effects

Hypothesis **H1** predicts that net buying from institutions prior to OMSR program announcements will be negatively related to OMSR program announcement effects. In this subsection, we make use of institutional trading data and examine the relationship between institutional trading before OMSR program announcements and the announcement effects of OMSR programs.

Table 2.3 presents the results of our OLS analysis. The dependent variable is the cumulative abnormal return over a 3-day period around OMSR announcements (i.e., the announcement return described in Table 2.1). Following the literature on open-market share repurchases, we calculate the announcement effects based on a market model, where the market beta is estimated with daily returns over the 6-month period ending one trading day before OMSR program announcements. The variable of interest is *Net Buy* from institutions. For ease of interpretation, from this subsection onwards, *Net Buy* is expressed in percentage rather than in basis points. In Panel A, *Net Buy* is aggregated over all sample institutions over the 12-month period before OMSR program announcements, whereas in Panel B, *Net Buy* is aggregated over all hedge funds over the 12-month period before OMSR program announcements.

From Model (1) of Panel A, we can see that the coefficient on *Net Buy* is negative and statistically significant. This is consistent with **H1**, suggesting that institutional trading prior to announcements of OMSR programs leads to the information produced by institutions about the intrinsic values of firms getting reflected in the announcing firms' stock prices before the announcement itself, so that the actual announcement effect of the OMSR program is smaller. We control for the variables that have been found in the literature to be able to explain OMSR program announcement effects, such as the size of the OMSR programs and the past stock return of the firm. In a recent paper, Babenko, Tserlukevich, and Vedrashko (2012) find that insider trading and insider holdings provide additional explanatory power regarding OMSR program announcement returns. In Models (2) and (3) of Panel A, we incrementally control for these variables and the coefficients on *Net Buy* remain negative and statistically significant.

In Panel B, we aggregate net buying by hedge funds, a subsample of all institutions, over the 12-month period before OMSR program announcements and perform a similar multivariate analysis as in Panel A. The coefficients on *Net Buy* remain negative and statistically significant, and the economic magnitude here is larger than that in Panel A, where *Net Buy* is calculated using trading by all sample institutions.

To summarize, we find evidence in this subsection that is consistent with hypothesis **H1**. Institutional trading before OMSR program announcements has predictive power for the announcement effect of such programs, in the sense that an algebraically lower institutional net buying before an OMSR program announcement is associated with a larger announcement effect. This result holds even after we control for variables capturing publicly available information such as the size of the OMSR program, prior firm performance, and insider trading. This evidence suggests that institutional trading prior to an OMSR program announcement indeed reflects the information produced by institutional investors about the intrinsic value of the firm. Additionally, we find evidence suggesting that hedge funds, as a subgroup of institutional investors, possess somewhat more accurate information regarding the intrinsic values of firms compared to the average for institutional investors in our overall sample. Overall, our empirical results in this section suggest that institutions are able to produce valuable information about the intrinsic values of firms announcing OMSR programs, consistent with a crucial assumption of the model of Brennan and Thakor (1990).

2.5.2 The Relation between Institutional Trading after OMSR Program Announcements and Subsequent Stock Returns

In the previous subsection, we examined the relationship between institutional trading before OMSR program announcements and the announcement effect of such programs. From this subsection onwards, we focus on institutional trading immediately after OMSR program announcements and examine the informativeness of such trading. We first investigate whether institutional trading immediately after OMSR program announcements predicts the subsequent stock return performance of firms (**H2**).

Table 2.4 reports the results of our multivariate analysis. The variable of interest is institutional trading, measured by *Net Buy*, after OMSR program announcements. We focus on institutional trading over the one month period after an OMSR program announcement. The dependent variable is the buy-and-hold abnormal return over the 12-month period in percentage points subsequent to the measurement period of *Net Buy*. Buy-and-hold abnormal return is the buy-and-hold raw return minus the buy-and-hold return of the Fama-French 25 portfolio matched on size and book-to-market. In Panel A, *Net Buy* is aggregated over all sample institutions over the one month period after OMSR program announcements, whereas in Panel B, *Net Buy* is aggregated over all hedge funds over the one month period after such announcements.[24]

In Model (1) of Panel A and Panel B in Table 2.4, we control for the size of the OMSR program, the amount of shares actually repurchased by the firm during the two fiscal quarters after the announcement, and various variables capturing different aspects of the firm's characteristics. The coefficient on *Net Buy* in Model (1) of Panel A is not statistically significant, suggesting that we do not find evidence that trading by institutions in our overall sample has predictive power about the future stock returns of repurchasing firms. Meanwhile, the coefficient on *Net Buy* in Model (1) of

[24]We focus only on hedge funds that participate in trading on the stocks of OMSR announcing firms on the announcement day. This group of hedge funds, that timely respond to OMSR announcement news, are presumably more informed about the announcing firm than other hedge funds that do not trade on the announcement day.

Panel B is positive and statistically significant, suggesting that trading by hedge funds as a group has strong predictive power for future stock returns of the repurchasing firms. In Models (2) - (4) of both Panel A and Panel B, we separately control for contemporaneous insider trading, industry fixed effects, and year fixed effects, and in Model (5) of both Panel A and Panel B, we control for these variables together. We thus find that hedge funds, as a subgroup of institutional investors, have a unique information advantage in their post-OMSR program announcement trading in terms of predicting the long-run stock returns of repurchasing firms.

In summary, we find evidence that trading by a subgroup of institutional investors, namely, hedge funds, after OMSR program announcements has predictive power for the subsequent stock return performance of OMSR program announcing firms. This is consistent with hypothesis **H2**, suggesting that some institutions, notably hedge funds, possess a residual information advantage over retail investors about the firm's intrinsic value, even after the announcement of an OMSR program.[25] Overall, this evidence suggests that the equilibrium in the equity market after an OMSR program announcement is such that there is room for the production of valuable information about the intrinsic value of firms by at least one subgroup of institutions, namely, hedge funds.

2.5.3 Profitability of Institutional Trading after OMSR Program Announcements

In the previous subsection, we presented evidence that institutional (hedge fund) trading after OMSR program announcements has predictive power for the subsequent long-run stock return performance of firms. This result suggests that institutions are able to produce private information about the intrinsic values of firms announcing OMSR programs, even after the announcement of such programs which may partially convey firm insiders' private information as well. In this subsection, we investigate hypotheses **H3** and **H4** by analyzing whether institutions are able to

[25]The result that aggregate institutional trading does not have predictive power for future stock returns may be partially explained by the fact that some institutions may also sell shares passively to provide liquidity for firms buying back shares in the open market: see, e.g., Huang and Zhang (2013).

use the information they have produced to realize abnormal trading profits after OMSR program announcements. Further, we analyze whether institutional investors make larger abnormal profits when the information conveyed by an OMSR program is more noisy (i.e., when the size of the OMSR program is smaller or when the firm repurchases a smaller number of shares).

We make use of our transaction-level institutional trading data to calculate trading profits, capital committed, and investment returns earned by sample institutions. Following Chemmanur, He, and Hu (2009), we consider a “raw" measure as well as a risk-adjusted measure, where we use the cumulative returns from the corresponding Fama-French 25 portfolio matched on size and book-to-market to discount profits and capital committed back to the first day of the trading horizon. We focus on institutional trading over the four quarters immediately after OMSR program announcements.

Table 2.5 presents the results of our analysis. In Panel A, we include trading for all institutions, and in Panel B, we include trading only for our identified hedge fund subsample. Further, we split our data by the noisiness of the signal conveyed by the OMSR program announcement: i.e., smaller versus larger OMSR programs; and by the subsequent actual share repurchase: i.e., smaller versus larger number of shares actually repurchased. In Panel A1 and B1, we split by the OMSR program size. We expect that a larger OMSR program size sends a stronger signal to the market, which leaves less room for information production by market participants (both institutional and retail investors) after such an OMSR program announcement. Therefore, we expect that institutional investors have less informational advantage compared to retail investors (and hence make smaller abnormal trading profits) when the OMSR program size is larger. In Panel A2 and B2, we split by the cumulative actual share repurchases made by the firm during the first two fiscal quarters after the OMSR program announcement. If a firm actually repurchases more, we expect that more information about firm value is already impounded into stock prices. Therefore, we expect institutional investors to have less of an informational advantage compared to retail investors (and hence make smaller abnormal trading profits) when the firm actually repurchases more shares.

We examine the realized investment return by institutions after OMSR program announcements in Table 2.5. We calculate two return measures, i.e., *return on buy principal* and *return on maximum investment*. For each return measure, we calculate a raw return measure without risk adjustment and a risk-adjusted return measure by discounting using benchmark returns of the corresponding Fama-French 25 portfolio matched on size and book-to-market. For example, *raw return on buy principal* is defined as raw profit divided by buy principal and *risk-adjusted return on maximum investment* is defined as risk-adjusted profit divided by risk-adjusted maximum investment. As we can see in Table 2.5 Panel A1, when the OMSR program size is smaller, institutions earn positive and statistically significant investment returns. Specifically, institutions on average realize a *risk-adjusted return on maximum investment* of 0.77% when the OMSR program size is below the sample median. On the other hand, when the OMSR program size is larger (above the sample median), institutions make zero and sometimes negative investment returns. The difference in investment returns realized by institutions between small and large OMSR programs is statistically significant. In Panel A2, we split our sample by the actual shares repurchased by the firm during the first two fiscal quarters after the OMSR program announcement, and show that, when firms actually repurchase less (below the median), institutions earn positive and statistically significant investment returns. Specifically, institutions on average realize a *risk-adjusted return on maximum investment* of 1.03% when the actual repurchase is below the sample median. On the other hand, when firms actually repurchase more (above the sample median), institutions make zero and sometimes negative investment returns. The difference in investment returns realized by institutions between small and large actual share repurchases is also statistically significant.

We also examine the realized investment returns by hedge funds after OMSR program announcements using similar return measures, and the results are presented in Panel B of Table 2.5. Consistent with our findings for all institutions, we find that hedge funds earn positive and statistically significant investment returns when the OMSR program size is smaller (Panel B1), or when the firm's actual repurchase is smaller (Panel B2). In comparison to the results reported in Panel A, hedges funds realize even higher investment returns than that realized by the average institution in our overall sample. When the OMSR program size is below the sample median, we find hedged funds on average realize 1.17% *risk-adjusted return on maximum investment* (versus 0.77% realized by average institutions); when the firm's actual share repurchase is below the sample median, we find that hedge funds on average realize 2.11% *risk-adjusted return on maximum investment* (versus 1.03% realized by the average institution in our overall sample).

In summary, we find that institutional investors realize abnormal investment returns in trading the stock of OMSR program announcing firms when the size of the OMSR program is smaller and when the firm makes smaller actual share repurchases subsequent to the OMSR program announcement (**H3** and **H4**). This suggests that institutional investors are able to produce valuable information about the intrinsic values of firms announcing OMSR programs, and are able to translate this information advantage into realized trading profits, when the information conveyed by an OMSR program announcement itself is more noisy, or when the number of shares actually repurchased after the OMSR program announcement is smaller. We also find that hedge funds, who are more specialized information producers compared to institutions as a whole, hold even more of an information advantage and realize greater investment profits. Overall, our empirical results in this subsection, together with our results presented in Section 2.5.2, suggest that the equilibrium in the equity market after an OMSR program announcement is a partial pooling equilibrium, which leaves room for the production of valuable information by institutions. Further, our results suggest that the precision of the information produced by institutions depends upon the noisiness of the signal conveyed by the OMSR program announcement and the information conveyed by the number of shares actually repurchased. Thus, our results presented in the subsection provide considerable support for the noisy signaling hypothesis of OMSR programs.

2.5.4 Institutional Trading after OMSR Program Announcements and Actual Share Repurchases

In this subsection, we investigate the relationship between institutional trading after OMSR program announcements and the firm's actual share repurchase activities. Specifically, we test hypothesis **H5** by examining whether institutional net buying after OMSR program announcements is positively related to the actual share repurchases made by firms subsequent to these announcements.

Table 2.6 presents the results of our multivariate analysis. In Panel A, we aggregate institutional trading (*Net Buy*) over the first month and the first quarter after an OMSR program announcement. The dependent variable is the firm's actual share repurchases during the two fiscal quarters after the measurement period of institutional net buying, and is defined as the number of shares repurchased by the firm normalized by the number of shares outstanding. In Panel A, the coefficient on *Net Buy* is positive and statistically significant for the one quarter horizon (it is positive but insignificant for the one month horizon). This suggests that aggregate trading by all institutions during the quarter after OMSR program announcements is positively related to the future actual share repurchases by firms.

We control for variables capturing different firm characteristics. In particular, we control for several variables that affect the firm's financial ability to repurchase. Specifically, we control for the firms' cash holdings, capital expenditures, R&D expenses, and whether these firms are paying dividends. All of the coefficients on these variables have the expected signs and are statistically significant in most cases. In Models (2)-(5) and (7)-(10), we additionally control for different combinations of insider trading, insider holdings, industry fixed effects, and year fixed effects. The coefficients on *Net Buy* remain positive and statistically significant. In fact, the magnitude and statistical significance of the coefficients on *Net Buy* remain similar across all the models we consider here, suggesting that institutional trading possesses additional predictive power that is not captured by these control variables regarding the firm's subsequent actual share repurchases.

We also examine the relationship between net buying by hedge funds after OMSR program announcements and firms' subsequent actual shares repurchases: the results are presented in Panel B of Table 2.6. We find that the coefficient on *Net Buy* (aggregated trading by hedge funds) is positive and statistically significant for the one month horizon (it is positive but insignificant for the one quarter horizon). This suggests that aggregate trading by all hedge funds during the first month after OMSR announcements is positively related to the future actual share repurchases by firms.

In summary, in this subsection we present evidence that institutional trading after an OMSR program announcement has predictive power for the subsequent actual share repurchases made by the firm, in the sense that greater net buying by institutional investors after an OMSR announcement is associated with greater actual share repurchases by the firm in the subsequent period. This is consistent with hypothesis **H5**, suggesting that the information produced by institutional investors after an OMSR program announcement is correlated with the residual undervaluation of firms announcing such programs. Since firms that are more undervalued after an OMSR program announcement are likely to undertake larger actual share repurchases according to the noisy signaling hypothesis, this induces a positive relationship between institutional net buying after OMSR program announcements and actual share repurchases. Overall, our results presented in this subsection provide further support for the noisy signaling hypothesis.

2.5.5 Institutional Trading and the Change in Equity Market Information Asymmetry around OMSR Program Announcements

In this subsection, we empirically examine the relationship between institutional trading after OMSR program announcements and the change in the information asymmetry facing firms from before an OMSR program announcement to after. Specifically, we test hypothesis **H6** by examining whether institutional net buying after an OMSR program announcement is associated with a larger decrease in the extent of information asymmetry facing firms in the equity market.

Table 2.7 presents the results of our multivariate analysis. Our main variable of interest is *Net Buy*, measured as the aggregate institutional net buying during the first two quarters after an OMSR program announcement. The dependent variable is the difference in the measures of firms' information asymmetry from before OMSR program announcements to after. In Panel A, B, and C, we use three different measures of information asymmetry based on I/B/E/S analyst earnings forecasts; in Panel D, we use a fourth measure of information asymmetry, namely, the bid-ask spread. For each firm announcing an OMSR program, we retrieve analyst earnings forecasts for the fiscal year end before the OMSR program announcement date (at least one year before the announcement date), and after the OMSR program announcement date (at least one year after the announcement date). In Panel A, the information asymmetry measure is the mean-squared error of analysts' forecasts (*MSE*). We measure forecast error as the absolute difference between the average forecasted earnings and the actual earnings per share divided by the price per share at the time of the forecast. In Panel B, the information asymmetry measure is the standard deviation of analyst forecasts (*Dispersion*). In Panel C, the information asymmetry measure is the coefficient of variation of analyst forecasts (*COV*), which is defined as the ratio of standard deviation in analyst forecasts to the absolute value of the average of analyst forecasts. In Panel D, we use the bid-ask spread (*BidAskSpread*) as a measure of information asymmetry. We first calculate the daily bid-ask spread as the average of all quoted spreads (the difference between the log ask price and log bid price) during normal trading hours on the day based on NYSE TAQ data, and average the daily bid-ask spread over the one-year horizon before the OMSR program announcement, as well as the one-year horizon following the second quarter after the OMSR program announcement.

As we can see from model (1) from Table 2.7 Panel A, the coefficient on *Net Buy* is negative and statistically significant, suggesting that a larger *Net Buy* is associated with a decrease in analyst forecast error from before the OMSR program announcement to after. This result is robust to controlling for firm size, book-to-market ratio, and past stock return. In model (2), we add industry and year fixed effects; in model (3) and (4), we additionally control for *Actual Repurchase* (actual shares repurchased by the firm during the first two fiscal quarters after the OMSR program announcement), *Insider Net Buy* (aggregate net purchase by top level insiders of the firm during the two fiscal quarters after the OMSR program announcement), as well as *Insider Holding* (aggregate stock holding of top level insiders of the firm at the end of the most recent fiscal year before the OMSR program announcement). The coefficients on *Net Buy* remain positive and statistically significant. Similarly, in Panel B, where we use standard deviation of analysts' forecasts as the measure of information asymmetry, we find that a larger *Net Buy* is associated with a decrease in *Dispersion* from before the OMSR program announcement to after; and in Panel C, we find that a larger *Net Buy* is associated with a decrease in *COV* (coefficient of variation of analysts' forecasts) from before the OMSR program announcement to after. Finally, in Panel D, we find that a larger *Net Buy* is associated with a decrease in the bid-ask spread of the firm's equity from before the OMSR program announcement to after.

In summary, in this subsection we present evidence that is consistent with hypothesis **H6**. This evidence provides direct support for the role of information production and trading by institutions in reducing the residual information asymmetry facing firms after an OMSR program announcement. In particular, we show that the reduction in information asymmetry facing firms in the equity market from before an OMSR program announcement to after such an announcement is greater when net buying by institutions in the firm's equity immediately after the announcement is greater.

2.6 Conclusion

In this paper, we have accomplished three objectives.

First, we have proposed a noisy signaling hypothesis of OMSR programs. Thus, in contrast to the existing literature which has theoretically demonstrated a separating equilibrium after share repurchase programs in general, we argued that the equilibrium in the equity market after an OMSR program announcement is likely to be a partial pooling equilibrium, where (in a setting with a continuum of types or a discrete type setting with three or more types) the highest firm types pool by announcing an OMSR program while the lowest types do not announce such a program. In this context, we argued that there is room for some information transmission from firm insiders to equity market investors through an OMSR program announcement, even in the absence of a commitment by the firm to buy back the entire amount of shares announced, as long as the firm or its insiders suffer a moderate shortfall cost per share: i.e., they suffer such a cost if the actual number of shares repurchased falls short of the target number of shares announced. Finally, we conjectured that, in such a partial pooling equilibrium, there are two mechanisms that play a role complementary to OMSR program announcements in further reducing the information asymmetry faced by the firm even after such announcements. The first such mechanism is actual share repurchases made by firms after OMSR program announcements: we argued that more undervalued firms repurchase a larger number of shares after OMSR program announcements, thereby conveying further information about intrinsic firm value to the equity market. The second such mechanism is information production by institutions and trading by them making use of this information after OMSR program announcements. We argued that, based on their information production, institutions buy more equity in more undervalued firms, and their information getting reflected in stock prices as a result of their trading further reduces the residual information asymmetry facing firms after OMSR program announcements.

Second, we tested the implications of the above noisy signaling hypothesis of OMSR program announcements for information production and trading by institutions. The results of our empirical analysis provide considerable support for the noisy signaling hypothesis. First, the fact that institutions are able to produce valuable information subsequent to OMSR program announcements (as evidenced by the predictive power of institutional trading and the realized profitability of such trading) provides support for the notion that the equilibrium prevailing in the equity market is a partial pooling rather than a fully separating equilibrium: there would be no room for information production in a separating equilibrium, since, in such an equilibrium, all information asymmetry is resolved upon the announcement of the OMSR program itself. Second, the positive relationship that we documented between institutional net buying and actual share repurchases by firms provides further support for the noisy signaling hypothesis. This positive relationship is likely to be induced by the fact that firms that are more undervalued after OMSR program announcements repurchase more of their own shares, while institutions net buy a larger number of shares after OMSR program announcements in firms that they believe to be more undervalued. Third, we provided direct evidence supporting the complementary role of information production and trading by institutions in reducing the residual information asymmetry facing firms after OMSR program announcements, by showing that the reduction in information asymmetry facing firms from before an OMSR program announcement to after such an announcement is greater when net buying by institutions in the firm's equity immediately after the announcement is greater.

Finally, our paper documented the predictive power of institutional trading prior to an OMSR program announcement for the announcement effect of such a program. This, along with the other empirical results discussed earlier, allows us to empirically confirm a crucial assumption made by the Brennan and Thakor (1990) model by showing that institutional investors are indeed able to produce valuable information about the intrinsic values of firms around OMSR programs.

Appendix: Identifying hedge funds within the Abel Noser sample

We identify hedge funds in the Abel Noser sample by merging management companies in Abel Noser with a list of hedge funds provided by Thomson Reuters. The client manager code along with the institutional manager code allows for the identification of a particular institutional investor.

The second dataset we use is the list of hedge funds provided by Thomson Reuters. This list is comprehensive as it classifies all 13F filers. We verify the quality of Thomson Reuters hedge fund classification by checking Form ADV filed by institutions. In particular, following Brunnermeier and Nagel (2004) and Griffin and Xu (2009), we classify an institution as a hedge fund if more than half of its investors are categorized as high net worth individuals or pooled investment vehicles in item 5.D. In addition, we require that the manager charge a performance-based fee (item 5.E).

To merge management companies in the Abel Noser sample with hedge funds in Thomson Reuters, we compute each institution's quarterly change in stock ownership (in number of shares) for each stock, denoted by ΔIO. For each pair of a Thomson Reuters hedge fund and an Abel Noser management company, we calculate the difference of ΔIO by two institutions and map the Thomson Reuters hedge fund to the Abel Noser management company with the closest ΔIO. Finally, we manually verify the matches identified above, using fund names from the Thomson Reuters and a manager name list disclosed by Abel Noser in 2011.

Our classification identifies 162 hedge funds in the whole sample. Since our identification is based on management companies, it is likely that our hedge fund sample includes some institutions that have both hedge funds and non-hedge funds business. Therefore, it is more appropriate to refer to hedge funds in our sample as hedge fund management companies: however, for brevity of presentation, we call them hedge funds in our analyses. b

Table 2.1: **Summary Statistics of Open-Market Share Repurchases.**
This table presents summary statistics of the open-market share repurchase programs (OMSRs) data from January 2004 to December 2010. *OMSR Program Size* is the value of the OMR program (in dollar amount), normalized by the market capitalization of the firm as of the most recent month-end before the announcement date. *Announcement Effect* is measured as the three-day ([0,2]) abnormal return, where date 0 is the announcement date. Abnormal returns are calculated based on a market model, where the market beta is estimated using returns over 126 trading days ending one trading day before the announcement (that is, [-126, -1]). *Actual Completion* is the actual repurchase by the firm during one-year period after the OMSR program announcement date as a percentage of OMSR Program Size. *Log Total Assets* is the natural logarithm of total assets at the most recent fiscal year end before the announcement; similarly, *Log B/M* is the natural logarithm of Book-to-Market ratio; *Cash Holdings* is the firm's cash holdings normalized by total assets; *R&D Expenses* is the firm's R&D expenses normalized by total assets; *Dividend Paying Dummy* is a dummy variable which equals one if the firm pays out dividends and zero otherwise. *Prior Quarter Market-adj Return* is the market adjusted stock return during the one-quarter period before the announcement. *Prior Year Market-adj Return* is the market adjusted stock return during the one-year period before the announcement.

Variable	N	Mean	Median	Std. Err.
OMSR Program Size	2988	7.94%	6.19%	0.12%
Announcement Effect [0, 2]	2988	1.74%	1.39%	0.18%
Actual Completion	2988	80.11%	67.71%	67.25%
Log Total Assets	2988	7.15	7.05	0.04
Log B/M	2988	-0.87	-0.81	0.01
Cash Holdings	2988	0.18	0.10	0.00
R&D Expenses	2988	0.03	0.00	0.00
Dividend Paying Dummy	2988	0.52	1.00	0.01
Prior Quarter Market-adj. Return	2988	-6.18%	-5.49%	0.31%
Prior Year Market-adj. Return	2988	-2.58%	-7.58%	0.73%

Table 2.2: **Summary Statistics of Institutional and Hedge Fund Trading Data around OMSR Programs.**
This table presents summary statistics of the institutional trading sample from January 2003 to September 2011. Annualized number of transactions, annualized trading volume, annualized principal traded, and annualized commission expense are computed based on all U.S. domestic equity traded by sample institutions from January 2003 to September 2011. Sample mean, median, and total are presented. Trading around OMSRs is the aggregated trading by sample institutions during the two-year period surrounding OMSR announcement dates in our OMSR data. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	All Institutions	Investment Managers	Plan Sponsors	Hedge Funds
Number of Institutions	868	372	496	162
Annualized Number of Transactions (thousands)				
Mean	75.42	134.64	31.01	26.87
Median	7.59	26.73	4.77	1.10
Total	65,464.75	50,085.37	15,379.38	4,353.14
Annualized Trading Volume (millions)				
Mean	350.69	600.57	163.29	170.18
Median	37.75	133.74	12.96	10.80
Total	304,402.64	223,412.34	80,990.29	27,569.45
Annualized Principal Traded ($ millions)				
Mean	10,318.32	17,910.06	4,624.51	5,085.31
Median	1,060.36	3,614.01	374.07	272.43
Total	8,956,297.68	6,662,542.45	2,293,755.23	823,819.42
Annualized Commission Expense ($ millions)				
Mean	8.65	15.55	3.47	4.86
Median	0.87	3.67	0.40	0.24
Total	7,506.46	5,786.06	1,720.40	788.02
Trading around OMRs ([-4Q, 4Q])				
Number of Institutions trading around OMRs	849	360	489	157
Number of Transactions per OMR (thousands)	24.51	21.49	3.02	4.95
Trading Volume per OMR (millions)	100.02	86.01	14.61	25.06
Principal Traded per OMR ($ millions)	3,501.76	3,000.47	522.43	909.38
Commission Expense per OMR ($ millions)	2.58	2.26	0.33	0.72

Table 2.3: **The Predictive Power of Pre-Announcement Institutional Trading for OMSR Program Announcement Returns.**

This table presents OLS regression analysis of the predictive power of trading by all institutions (and a subsample of hedge funds) for the OMSR program announcement returns between 2004 and 2010. Panel A (B) presents the results using trading by all institutions (hedge funds). The dependent variable is OMSR program announcement returns, measured as the three-day ([0,2]) abnormal return, where date 0 is the announcement date. Abnormal returns are calculated based on a market model, where the market beta is estimated using daily returns over 6-month period ending one trading day before the announcement. In Panel A, *Net Buy* is the aggregate net buying, scaled by number of shares outstanding of the firm, by our institutions sample during the one-year periods before the OMSR program announcement; in Panel B, *Net Buy* is the aggregate net buying, scaled by number of shares outstanding of the firm, by our hedge funds sample during the one-year periods before the OMSR program announcement. *OMSR Program Size* is the value of the OMR program (in dollar amount), normalized by the market capitalization of the firm as of the most recent month-end before the announcement date; *Log Total Assets* is the natural logarithm of total assets of firm at the most recent fiscal year end before the announcement; *Log B/M* is the natural logarithm of Book-to-Market ratio of firm at the most recent fiscal year end before the announcement; *Industry Adj. ROA* is firm's EBIT/Total Asset minus two-digit SIC industry's median EBIT/Total Asset, at the most recent fiscal year end before the announcement; *Cash Holdings* is firm's cash holdings normalized by total assets, at the most recent fiscal year end before the announcement; *Past Stock Return* is the market adjusted stock return during the one-year period before the announcement; *Leverage* is firm's total liabilities normalized by total assets, at the most recent fiscal year end before the announcement; *Dividend Paying Dummy* is a dummy variable which equals one if firm pays out dividends during the most recent fiscal year before the announcement, and equals zero otherwise; *Insider Net Buy* is the aggregate net buy from top level insiders of firm during the one-year period before the announcement; *Insider Holding* is the aggregate stock holding of top level insiders of firm at the beginning of the one-year period before the announcement. Industry (two-digit SIC code) and year fixed effects are included. t-statistics are in parentheses. ***, **, and * represent statistical significance at the 1, 5, and 10 percent levels.

Panel A: Trading by all institutional investors

	Dep. Var.: OMSR Announcement Returns		
	(1)	(2)	(3)
Net Buy	-0.0005**	-0.0005**	-0.0005**
	(-2.06)	(-2.03)	(-2.03)
OMSR Program Size	0.0005***	0.0005***	0.0005***
	(3.41)	(3.42)	(3.42)
Log Total Asset	-0.0021***	-0.0021***	-0.0020***
	(-3.68)	(-3.70)	(-3.53)
Log B/M	0.0025	0.0024	0.0024

	(1.33)	(1.29)	(1.28)
Industry Adj. ROA	-0.0114	-0.0106	-0.0107
	(-0.87)	(-0.81)	(-0.81)
Cash	0.0004	0.0006	0.0006
	(0.06)	(0.09)	(0.08)
Past Stock Return	0.0068**	0.0068**	0.0068**
	(2.05)	(2.04)	(2.04)
Leverage	0.0069	0.0068	0.0068
	(1.07)	(1.05)	(1.05)
Dividend Dummy	0.0003	0.0003	0.0002
	(0.17)	(0.13)	(0.11)
Insider Net Buy		0.0006	0.0006
		(0.74)	(0.74)
Insider Holding			0.0000
			(0.17)
Industry Fixed Effects	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes
Observations	2,853	2,853	2,853
R-squared	0.05	0.05	0.05

Panel B: Trading by hedge funds

	Dep. Var.: OMSR Announcement Returns		
	(1)	(2)	(3)
Net Buy	-0.0013**	-0.0013**	-0.0013**
	(-2.53)	(-2.51)	(-2.51)
OMSR Program Size	0.0005***	0.0005***	0.0005***
	(3.44)	(3.45)	(3.46)
Log Total Asset	-0.0020***	-0.0021***	-0.0020***
	(-3.65)	(-3.67)	(-3.50)
Log B/M	0.0025	0.0025	0.0024
	(1.33)	(1.29)	(1.28)
Industry Adj. ROA	-0.0119	-0.0112	-0.0112

	(-0.91)	(-0.85)	(-0.85)
Cash	0.0005	0.0008	0.0007
	(0.07)	(0.11)	(0.09)
Past Stock Return	0.0070**	0.0070**	0.0070**
	(2.10)	(2.09)	(2.09)
Leverage	0.0068	0.0067	0.0066
	(1.05)	(1.03)	(1.02)
Dividend Dummy	0.0005	0.0004	0.0004
	(0.25)	(0.21)	(0.19)
Insider Net Buy		0.0006	0.0006
		(0.76)	(0.77)
Insider Holding			0.0000
			(0.19)
Industry Fixed Effects	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes
Observations	2,853	2,853	2,853
R-squared	0.05	0.05	0.05

Table 2.4: **The Predictive Power of Institutional Trading after OMSR Program Announcements for Subsequent Long-run Stock Returns.**

This table presents OLS regression analysis of the predictive power of institutional trading after an OMSR program announcement for the subsequent one-year stock return performance of a firm. *Net Buy* is the aggregate net buying from all institutions (hedge funds) during the first month after the OMSR program announcement in panel A (B). The dependent variable is cumulative stock return measured in percentage points over the one-year period subsequent to the measurement period of *Net Buy*, adjusted by the return of matched Fama-French 5×5 size and book-to-market portfolio return. *OMSR Program Size* is the value of the OMSR program (in dollar amount), normalized by the market capitalization of the firm as of the most recent month-end before the announcement date; *Past Stock Return* is the market adjusted stock return during the one-year period before the announcement; *Log Total Assets* is the natural logarithm of total assets of firm at the most recent fiscal year end before the announcement; *Industry Adj. ROA* is firm's EBIT/Total Asset minus 2-digit SIC industry's median EBIT/Total Asset, at the most recent fiscal year end before the announcement; *Cash* is firm's cash holdings normalized by total assets, at the most recent fiscal year end before the announcement; *Cash Flow* is firm's cash flow normalized by total assets; *S&P 500 Dummy* is an indicator variable that equals one for firms in the S&P 500 Index and zero otherwise; *Log B/M* is the natural logarithm of Book-to-Market ratio of firm at the most recent fiscal year end before the announcement; *Insider Net Buy* is the aggregate net buy from top level insiders of firm during the one-year period before the announcement; *Insider Holding* is the aggregate stock holding of top level insiders of firm at the beginning of the one-year period before the announcement. Industry (two-digit SIC code) and year fixed effects are included. t-statistics are in parentheses. ***, **, and * represent statistical significance at the 1, 5, and 10 percent levels.

Panel A: Trading by all institutional investors

	Dependent Variable: Size and book-to-Market adjusted buy-and-Hold return over 12 months after the measurement period of *Net Buy*				
	(1)	(2)	(3)	(4)	(5)
Net Buy	-0.0000	-0.0000	0.0000	0.0000	0.0000
	(-0.31)	(-0.08)	(0.03)	(0.03)	(0.03)
OMSR Program Size	-0.0026***	-0.0023**	-0.0025***	-0.0025***	-0.0025***
	(-2.97)	(-2.54)	(-2.67)	(-2.67)	(-2.67)
Past Stock Return	0.5143***	0.5711***	0.5680***	0.5680***	0.5680***
	(12.22)	(10.39)	(10.35)	(10.33)	(10.33)
Log Total Asset	-0.0069	-0.0047	-0.0048	-0.0048	-0.0047
	(-1.43)	(-0.98)	(-0.92)	(-0.92)	(-0.90)
Industry Adj. ROA	-0.2341*	-0.2353*	-0.2484*	-0.2485*	-0.2483*
	(-1.83)	(-1.91)	(-1.88)	(-1.88)	(-1.88)
Cash	0.0760**	0.0614*	0.0578	0.0578	0.0576

	(2.06)	(1.69)	(1.37)	(1.37)	(1.36)
Cash Flow	0.3502***	0.3363***	0.3122***	0.3123***	0.3123***
	(2.86)	(2.89)	(2.67)	(2.67)	(2.67)
S&P 500 Dummy	0.0385**	0.0366**	0.0369*	0.0369*	0.0368*
	(2.10)	(2.01)	(1.92)	(1.91)	(1.91)
Log B/M	-0.0082	-0.0204*	-0.0212*	-0.0212*	-0.0212*
	(-0.71)	(-1.80)	(-1.82)	(-1.82)	(-1.82)
Insider Net Buy				-0.0012	-0.0010
				(-0.04)	(-0.04)
Insider Holding					0.0001
					(0.12)
Industry Fixed Effects	No	No	Yes	Yes	Yes
Year Fixed Effects	No	Yes	Yes	Yes	Yes
Observations	2,789	2,789	2,789	2,789	2,789
R-squared	0.27	0.31	0.32	0.32	0.32

Panel B: Trading by hedge funds

	Dependent Variable: Size and book-to-Market adjusted buy-and-Hold return over 12 months after the measurement period of *Net Buy*				
	(1)	(2)	(3)	(4)	(5)
Net Buy	0.0005**	0.0005**	0.0005**	0.0005**	0.0005**
	(2.15)	(2.35)	(2.25)	(2.25)	(2.25)
OMSR Program Size	-0.0026***	-0.0022**	-0.0024***	-0.0024***	-0.0024***
	(-2.91)	(-2.48)	(-2.63)	(-2.63)	(-2.63)
Past Stock Return	0.5146***	0.5715***	0.5681***	0.5681***	0.5681***
	(12.26)	(10.43)	(10.39)	(10.37)	(10.37)
Log Total Asset	-0.0068	-0.0047	-0.0046	-0.0046	-0.0045
	(-1.42)	(-0.97)	(-0.89)	(-0.89)	(-0.86)
Industry Adj. ROA	-0.2354*	-0.2367*	-0.2520*	-0.2521*	-0.2519*
	(-1.84)	(-1.93)	(-1.91)	(-1.91)	(-1.90)
Cash	0.0745**	0.0595	0.0569	0.0568	0.0567
	(2.02)	(1.63)	(1.35)	(1.35)	(1.34)

Cash Flow	0.3500***	0.3353***	0.3118***	0.3118***	0.3119***
	(2.86)	(2.88)	(2.66)	(2.66)	(2.66)
S&P 500 Dummy	0.0379**	0.0359**	0.0357*	0.0358*	0.0356*
	(2.07)	(1.98)	(1.86)	(1.85)	(1.85)
Log B/M	-0.0087	-0.0211*	-0.0218*	-0.0218*	-0.0218*
	(-0.75)	(-1.86)	(-1.87)	(-1.87)	(-1.87)
Insider Net Buy				-0.0012	-0.0010
				(-0.04)	(-0.04)
Insider Holding					0.0001
					(0.12)
Industry Fixed Effects	No	No	Yes	Yes	Yes
Year Fixed Effects	No	Yes	Yes	Yes	Yes
Observations	2,789	2,789	2,789	2,789	2,789
R-squared	0.27	0.31	0.32	0.32	0.32

Table 2.5: **Profitability of Institutional Trading after OMSR Program Announcements.**

This table reports univariate results of the profitability of institutional trading around open-market share repurchase programs (OMSRs). We consider the trading horizon starting from the first fiscal quarter after the OMSR announcement and ending at the fourth fiscal quarter after the OMSR announcement. Panel A (B) reports the profitability of trading by all institutions (hedge funds). Panel A1 (B1) reports the result where we split the sample by the cumulative actual repurchase by the firm during the first two fiscal quarters after the OMSR announcement. Panel A2 (B2) reports the result where split the sample by the OMSR program size, defined as the dollar amount value of the OMSR program, normalized by the market capitalization of the firm as of the most recent month-end before the announcement date. *Raw Profit* is the total raw profit earned by institutions using actual transaction prices net of commissions, with the net position marked to market at the end of the trading horizon. *Buy Principal* is the sum of the actual dollar amount of all the buy transactions including commissions spent by sample institutions during the trading horizon. *Maximum Investment* is the maximum dollar amount committed to trading the sample firms' shares during the trading horizon by the institutions. *Raw Return on Buy Principal* is defined as the ratio of *Raw Profit* to *Buy Principal*. *Raw Return on Maximum Investment* is defined as the ratio of *Raw Profit* to *Maximum Investment*. We also discount profit and investment amount back to the first day of the trading horizon using the buy-and-hold value-weighted return from the Fama and French 25 portfolios matched on size and book-to-market. For example, *Risk-adjusted Profit* is computed by discounting the raw profit back to the first day of the trading horizon using the benchmark return from the matched Fama-French 25 portfolios; and *Risk-adjusted Return on Buy Principal* equals *Risk-adjusted Profit* divided by *Risk-adjusted Buy Principal*. T-tests are performed on the profits and the returns, respectively. T-tests are also performed on the difference in the profits, the difference in the investment amount, and the difference in the returns. Statistical significance is indicated by *** for 1% level, ** for 5% level, and * for 10% level.

Panel A: Trading by all institutional investors

	Institutional trading over $[FQ1, FQ4]$					
	Panel A1: Split by OMSR Program Size			Panel A2: Split by Actual Repurchase		
	Below Median	Above Median	Diff (B-A)	Below Median	Above Median	Diff (B-A)
Number of Observations (Events)	1392	1392		1392	1392	
Raw Profit ($ thousands)	4488.44	-3085.63	7574.08	3590.33	-2187.52	5777.86
Risk-adjusted Raw Profit ($ thousands)	4479.27	-2838.73	7317.99	4901.88	-3261.33	8163.21*
Buy Principal ($ millions)	837.72	865.93	-28.21	674.67	1028.45	-353.78***
Risk-adjusted Buy Principal ($ millions)	840.05	880.86	-40.80	691.9	1028.56	-336.65***
Maximum Investment ($ millions)	582.48	601.87	-19.39	468.23	716.12	-247.89***

Risk-adjusted Maximum Investment ($ millions)	580.83	607.12	-26.29	474.75	713.21	-238.46***
Raw Return on Buy Principal (%)	0.54**	-0.36	0.90**	0.54*	-0.21	0.75**
Risk-adjusted Return on Buy Principal (%)	0.54**	-0.32*	0.86***	0.72***	-0.32*	1.03***
Raw Return on Maximum Investment (%)	0.77***	-0.51*	1.28***	0.77**	-0.31	1.07***
Risk-adjusted Return on Maximum Investment (%)	0.77***	-0.47*	1.24***	1.03***	-0.46**	1.49***

Panel B: Trading by hedge funds

	Hedge fund trading over $[FQ1, FQ4]$					
	Panel B1: Split by OMSR Program Size			Panel B2: Split by Actual Repurchase		
Number of Observations(Events)	1392	1392		1392	1392	
Raw Profit ($ thousands)	2626.12	595.81	2030.32	2154.55*	1068.61	1085.94
Risk-adjusted Raw Profit ($ thousands)	3588.88***	184.54	3404.33**	2738.66***	1036.68	1701.98
Buy Principal ($millions)	468.87	29.37	439.5***	193.67	304.26	-110.59***
Risk-adjusted Buy Principal ($millions)	472.63	28.96	443.66***	200.29	301	-100.71***
Maximum Investment ($millions)	306.01	23.61	282.41***	126.28	203.14	-76.87***
Risk-adjusted Maximum Investment ($millions)	307.89	23.23	284.66***	129.91	201.01	-71.11***
Raw Return on Buy Principal (%)	0.55*	1.90**	-1.35	1.10**	0.33	0.77
Risk-adjusted Return on Buy Principal (%)	0.75***	0.6	0.15	1.37***	0.33	1.04**
Raw Return on Maximum Investment (%)	0.90*	1.77*	-0.88	1.68***	0.51	1.17
Risk-adjusted Return on Maximum Investment (%)	1.17***	0.54	0.63	2.11***	0.49	1.63***

Table 2.6: **The Relation between Institutional Trading after OMSR Program Announcements and Actual Share Repurchases.** This table presents OLS regression analysis of the relationship between institutional trading after an OMSR announcement and the subsequent actual share repurchases by the firm. *Actual Share Repurchase*, the dependent variable, is the actual shares repurchased by the firm (normalized by the number of shares outstanding) during the second and third fiscal quarter after the OMSR program announcement (two fiscal quarters after the measurement period of *Net Buy*). In Models (1)-(5), *Net Buy* is aggregated net buying by all institutions (hedge funds) during the first month after the OMSR announcement in panel A (B); In Models (6)-(10), *Net Buy* is aggregated net buying by all institutions (hedge funds) during the first quarter after the OMSR announcement in panel A (B). *Log Total Assets* is the natural logarithm of total assets of firm at the most recent fiscal year end before the announcement; *Log B/M* is the natural logarithm of Book-to-Market ratio of firm at the most recent fiscal year end before the announcement; *Past Stock Return* is the market adjusted stock return during the one-year period before the announcement; *Cash* is firm's cash holdings normalized by total assets, at the most recent fiscal year end before the announcement; *CapEx* is firm's capital expenditure normalized by total assets; *R&D* is firm's R&D expenditure normalized by total assets; *Dividend Paying Dummy* is a dummy variable which equals one if firm pays out dividends during the most recent fiscal year before the announcement, and equals zero otherwise; *OMSR Program Size* is the value of the OMSR program (in dollar amount), normalized by the market capitalization of the firm as of the most recent month-end before the announcement date; *Actual Share Repurchase FQ1* is the actual share repurchased by the firm (normalized by the number of shares outstanding) during the first fiscal quarter after the OMSR program announcement; *Insider Net Buy* is the aggregate net buying from top level insiders of firm during the one-year period before the announcement; *Insider Holding* is the aggregate stock holding of top level insiders of firm at the beginning of the one-year period before the announcement. Industry (two-digit SIC code) and year fixed effects are included. t-statistics are in parentheses. ***, **, and * represent statistical significance at the 1, 5, and 10 percent levels.

Panel A: Trading by all institutional investors

	Net Buy (one month after OMSR program announcement)					*Net Buy* (one quarter after OMSR program announcement)				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Net Buy	0.0126	0.0074	0.0219	0.0162	0.0164	0.0309**	0.0290**	0.0369***	0.0346**	0.0346**
	(0.32)	(0.19)	(0.56)	(0.41)	(0.42)	(2.25)	(2.12)	(2.67)	(2.51)	(2.51)
Log Total Asset	0.1683***	0.1548***	0.1782***	0.1654***	0.1576***	0.1699***	0.1565***	0.1793***	0.1667***	0.1589***
	(4.98)	(4.57)	(5.00)	(4.64)	(4.29)	(5.03)	(4.62)	(5.04)	(4.68)	(4.33)
Log B/M	-0.2751***	-0.2020**	-0.2003**	-0.1148	-0.1150	-0.2715***	-0.1997**	-0.1958**	-0.1119	-0.1121
	(-3.48)	(-2.49)	(-2.41)	(-1.34)	(-1.34)	(-3.44)	(-2.47)	(-2.35)	(-1.31)	(-1.31)
Past Stock Return	-0.2189	-0.3660**	-0.2566	-0.4176**	-0.4168**	-0.2344	-0.3819**	-0.2785*	-0.4391**	-0.4384**

	(-1.31)	(-2.09)	(-1.55)	(-2.40)	(-2.39)	(-1.41)	(-2.18)	(-1.68)	(-2.52)	(-2.52)
Cash	0.3490	0.4052	0.0700	0.1147	0.1333	0.3624	0.4179	0.0782	0.1230	0.1416
	(0.90)	(1.04)	(0.17)	(0.28)	(0.32)	(0.93)	(1.08)	(0.19)	(0.30)	(0.34)
CapEx	-1.4664	-1.1464	-2.8027*	-2.3719	-2.3167	-1.5085	-1.1915	-2.8765*	-2.4452*	-2.3894
	(-1.26)	(-0.98)	(-1.89)	(-1.60)	(-1.57)	(-1.29)	(-1.02)	(-1.95)	(-1.66)	(-1.62)
R&D	-3.4815***	-3.0161**	-2.3372	-1.9301	-1.9621	-3.4672***	-3.0134**	-2.3180	-1.9259	-1.9581
	(-2.67)	(-2.31)	(-1.59)	(-1.31)	(-1.33)	(-2.66)	(-2.31)	(-1.58)	(-1.31)	(-1.33)
Dividend Paying Dummy	-0.5925***	-0.5934***	-0.3667***	-0.3566**	-0.3492**	-0.5896***	-0.5901***	-0.3604**	-0.3506**	-0.3431**
	(-4.45)	(-4.45)	(-2.62)	(-2.55)	(-2.49)	(-4.43)	(-4.43)	(-2.57)	(-2.51)	(-2.45)
OMSR Program Size	0.0782***	0.0771***	0.0694***	0.0679***	0.0678***	0.0781***	0.0770***	0.0692***	0.0677***	0.0675***
	(8.56)	(8.45)	(7.50)	(7.35)	(7.33)	(8.56)	(8.44)	(7.49)	(7.33)	(7.31)
Actual Repurchase FQ1	0.1896***	0.1861***	0.1675***	0.1633***	0.1627***	0.1940***	0.1906***	0.1719***	0.1678***	0.1672***
	(6.94)	(6.77)	(6.14)	(5.95)	(5.92)	(7.09)	(6.92)	(6.30)	(6.10)	(6.08)
Insider Net Buy				0.1309	0.1167				0.1121	0.0978
				(0.37)	(0.33)				(0.31)	(0.27)
Insider Holding					-0.0071					-0.0072
					(-0.87)					(-0.88)
Industry Fixed Effects	No	No	Yes	Yes	Yes	No	No	Yes	Yes	Yes
Year Fixed Effects	No	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes
Observations	2,857	2,857	2,857	2,857	2,857	2,857	2,857	2,857	2,857	2,857
R-squared	0.07	0.08	0.12	0.13	0.13	0.07	0.08	0.12	0.13	0.13

Panel B: Trading by hedge funds

	Net Buy (one month after OMSR program announcement)					*Net Buy* (one quarter after OMSR program announcement)				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Net Buy	0.0034**	0.0033**	0.0034**	0.0033**	0.0033**	0.0003	0.0003	0.0002	0.0002	0.0002

	(2.20)	(2.16)	(2.20)	(2.15)	(2.15)	(0.43)	(0.43)	(0.31)	(0.32)	(0.32)
Log Total Asset	0.1686***	0.1552***	0.1790***	0.1662***	0.1585***	0.1683***	0.1549***	0.1780***	0.1653***	0.1575***
	(4.99)	(4.58)	(5.03)	(4.66)	(4.32)	(4.98)	(4.57)	(5.00)	(4.64)	(4.29)
Log B/M	-0.2727***	-0.1998**	-0.1981**	-0.1129	-0.1131	-0.2752***	-0.2019**	-0.2011**	-0.1152	-0.1154
	(-3.45)	(-2.47)	(-2.38)	(-1.32)	(-1.32)	(-3.48)	(-2.49)	(-2.42)	(-1.34)	(-1.35)
Past Stock Return	-0.2152	-0.3607**	-0.2529	-0.4127**	-0.4120**	-0.2195	-0.3671**	-0.2572	-0.4191**	-0.4183**
	(-1.29)	(-2.06)	(-1.53)	(-2.37)	(-2.37)	(-1.32)	(-2.09)	(-1.55)	(-2.41)	(-2.40)
Cash	0.3383	0.3956	0.0589	0.1054	0.1237	0.3452	0.4034	0.0608	0.1082	0.1266
	(0.87)	(1.02)	(0.14)	(0.26)	(0.30)	(0.89)	(1.04)	(0.15)	(0.26)	(0.31)
CapEx	-1.5220	-1.2061	-2.8915*	-2.4676*	-2.4125	-1.4708	-1.1536	-2.7852*	-2.3606	-2.3053
	(-1.31)	(-1.04)	(-1.96)	(-1.67)	(-1.63)	(-1.26)	(-0.99)	(-1.88)	(-1.60)	(-1.56)
R&D	-3.4193***	-2.9613**	-2.2835	-1.8819	-1.9135	-3.4683***	-3.0082**	-2.3265	-1.9235	-1.9553
	(-2.62)	(-2.27)	(-1.55)	(-1.28)	(-1.30)	(-2.66)	(-2.30)	(-1.58)	(-1.31)	(-1.33)
Dividend Paying Dummy	-0.5922***	-0.5927***	-0.3646***	-0.3540**	-0.3466**	-0.5929***	-0.5937***	-0.3678***	-0.3574**	-0.3500**
	(-4.45)	(-4.45)	(-2.60)	(-2.53)	(-2.48)	(-4.45)	(-4.46)	(-2.62)	(-2.56)	(-2.50)
OMSR Program Size	0.0787***	0.0776***	0.0698***	0.0683***	0.0682***	0.0782***	0.0771***	0.0693***	0.0679***	0.0678***
	(8.62)	(8.51)	(7.55)	(7.40)	(7.38)	(8.56)	(8.45)	(7.50)	(7.35)	(7.33)
Actual Repurchase FQ1	0.1901***	0.1867***	0.1681***	0.1640***	0.1634***	0.1897***	0.1864***	0.1673***	0.1633***	0.1627***
	(6.97)	(6.80)	(6.18)	(5.98)	(5.95)	(6.95)	(6.78)	(6.14)	(5.94)	(5.92)
Insider Net Buy				0.1399	0.1258				0.1342	0.1201
				(0.39)	(0.35)				(0.38)	(0.34)
Insider Holding					-0.0070					-0.0071
					(-0.86)					(-0.87)
Industry Fixed Effects	No	No	Yes	Yes	Yes	No	No	Yes	Yes	Yes
Year Fixed Effects	No	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes

Observations	2,857	2,857	2,857	2,857	2,857	2,857	2,857	2,857	2,857	2,857
R-squared	0.07	0.08	0.12	0.13	0.13	0.07	0.08	0.12	0.13	0.13

Table 2.7: **The Relation between Institutional Trading after OMSR Program Announcements and Changes in Equity Market Information Asymmetry.**

This table presents OLS regression analysis of the effect institutional trading after an OMSR announcement has on the change in the firm's information asymmetry from before the OMSR announcement to after. The dependent variable is the difference in the measures of firm's information asymmetry from before the OMSR announcement to after. In Panel A, B, and C, we use three measures of information asymmetry based on I/B/E/S analyst earnings forecasts. For each firm announcing an OMSR program, we retrieve analyst earnings forecasts for the fiscal year end before the OMSR program announcement date (at least one year before the announcement date), and after the OMSR program announcement date (at least one year after the announcement date). In Panel A, the information asymmetry measure is the mean-squared error of analysts' forecasts (*MSE*). We measure forecast error as the absolute difference between the average forecasted earnings and the actual earnings per share divided by the price per share at the time of the forecast. In Panel B, the information asymmetry measure is the standard deviation of analyst forecasts (*Dispersion*). In Panel C, the information asymmetry measure is the coefficient of variation of analyst forecasts (*COV*), which is defined as the ratio of standard deviation to the absolute value of the average of analyst forecasts. In Panel D, we use bid-ask spread (*BidAskSpread*) as a measure of information asymmetry. We first calculate the daily bid-ask spread as the average of all quoted spread (the difference between the log ask price and log bid price) during normal trading hours on the day based on NYSE TAQ data, and we average the daily bid-ask spread over the one-year horizon before the OMSR program announcement, as well as the one-year horizon following the second quarter after the OMSR program announcement. *Log (Market Cap)* is the natural logarithm of market capitalization of the firm at the end of OMSR program announcement; *Book-to-Market Ratio* is the book-to-market ratio based on the accounting information at the end of the most recent fiscal year before the OMSR announcement; *Actual Repurchase* is the actual repurchased shares by the firm, as a percentage of total shares outstanding at OMSR program announcement, during the first two fiscal quarters after the OMSR announcement; *Past Stock Return* is the market adjusted stock return during the one-year period before the announcement; *Insider Net Buy* is the aggregate Net Buy from top level insiders of firm during the two fiscal quarters after the OMSR announcement; *Insider Holding* is the aggregate stock holding of top level insiders of firm at the end of the most recent fiscal year before the OMSR announcement. Industry (two-digit SIC code) and year fixed effects are included. t-statistics are in parentheses. ***, **, and * represent statistical significance at the 1, 5, and 10 percent levels.

Panel A: Using change in analyst forecast error as dependent variable

	Dependent Variable: $\Delta\log(MSE)$			
	(1)	(2)	(3)	(4)
Net Buy	-0.0571***	-0.0576***	-0.0517**	-0.0583***
	(-2.93)	(-3.07)	(-2.33)	(-2.69)
Log(Market Cap)	-0.0475	-0.0570	-0.0906*	-0.0801
	(-1.10)	(-1.30)	(-1.84)	(-1.58)
Book-to-Market Ratio	-0.0995	-0.1108	-0.1337	-0.2208

	(-0.51)	(-0.55)	(-0.59)	(-0.93)
Past Stock Return	-1.8606***	-1.5167***	-1.7081***	-1.3861***
	(-10.00)	(-7.80)	(-8.46)	(-6.56)
Actual Repurchase			0.0042	-0.0039
			(0.16)	(-0.15)
Insider Net Buy			0.6697***	0.1672
			(3.25)	(0.82)
Insider Holding			0.0288**	0.0278**
			(2.13)	(2.09)
Industry Fixed Effects	No	Yes	No	Yes
Year Fixed Effects	No	Yes	No	Yes
R^2	0.061	0.208	0.075	0.218
Observations	1821	1821	1394	1394

Panel B: Using change in analyst forecast dispersion as dependent variable

	Dependent Variable: $\Delta \log(Dispersion)$			
	(1)	(2)	(3)	(4)
Net Buy	-0.0106	-0.0158*	-0.0146	-0.0194**
	(-1.25)	(-1.85)	(-1.53)	(-1.99)
Log(Market Cap)	0.0036	-0.0065	0.0064	0.0044
	(0.19)	(-0.32)	(0.30)	(0.19)
Book-to-Market Ratio	-0.1344	-0.1878**	-0.2109**	-0.2731**
	(-1.57)	(-2.04)	(-2.14)	(-2.57)
Past Stock Return	-0.1818**	-0.1642*	-0.2076**	-0.1769*
	(-2.24)	(-1.86)	(-2.38)	(-1.87)
Actual Repurchase			0.0139	0.0121
			(1.21)	(1.03)
Insider Net Buy			-0.0090	-0.1396
			(-0.10)	(-1.52)
Insider Holding			0.0134**	0.0160***
			(2.21)	(2.58)

Industry Fixed Effects	No	Yes	No	Yes
Year Fixed Effects	No	Yes	No	Yes
R^2	0.005	0.091	0.014	0.100
Observations	1807	1807	1384	1384

Panel C: Using change in coefficient of variation of analyst forecasts as dependent variable

	Dependent Variable: $\Delta \log(COV)$			
	(1)	(2)	(3)	(4)
Net Buy	-0.0231**	-0.0268**	-0.0246**	-0.0281**
	(-2.18)	(-2.57)	(-2.06)	(-2.37)
Log(Market Cap)	-0.0245	-0.0283	-0.0249	-0.0155
	(-1.04)	(-1.15)	(-0.93)	(-0.56)
Book-to-Market Ratio	-0.0473	-0.0721	-0.1125	-0.1623
	(-0.44)	(-0.64)	(-0.92)	(-1.25)
Past Stock Return	-0.9062***	-0.7998***	-0.8458***	-0.7387***
	(-8.95)	(-7.39)	(-7.79)	(-6.40)
Actual Repurchase			0.0114	0.0130
			(0.79)	(0.91)
Insider Net Buy			0.2233**	-0.0237
			(2.00)	(-0.21)
Insider Holding			0.0196**	0.0216***
			(2.49)	(2.76)
Industry Fixed Effects	No	Yes	No	Yes
Year Fixed Effects	No	Yes	No	Yes
R^2	0.049	0.160	0.061	0.180
Observations	1805	1805	1383	1383

Panel D: Using change in average bid-ask spread (%) as dependent variable

	Dependent Variable: $\Delta BidAskSpread$			
	(1)	(2)	(3)	(4)
Net Buy	-0.0057***	-0.0037*	-0.0043**	-0.0036*
	(-2.76)	(-1.95)	(-2.11)	(-1.84)
Log(Market Cap)	-0.0594***	-0.0516***	-0.0472***	-0.0410***
	(-15.68)	(-13.65)	(-11.91)	(-10.13)
Book-to-Market Ratio	-0.0260	-0.0022	-0.0071	0.0020
	(-1.60)	(-0.14)	(-0.41)	(0.11)

Past Stock Return	-0.1248***	-0.0705***	-0.1350***	-0.0856***
	(-8.47)	(-4.87)	(-8.27)	(-5.17)
Actual Repurchase			0.0014	0.0018
			(0.59)	(0.80)
Insider Net Buy			0.0827***	0.0566***
			(6.56)	(4.61)
Insider Holding			0.0001	0.0011
			(0.07)	(1.04)
Industry Fixed Effects	No	Yes	No	Yes
Year Fixed Effects	No	Yes	No	Yes
R^2	0.125	0.297	0.147	0.291
Observations	2621	2621	1820	1820

Chapter 3

The Geography of Institutional Investors, Information Production, and Initial Public Offerings

3.1 Introduction

It is well-known that institutions play a key role in initial public offerings. On the one hand, it has been argued that IPO underwriters go out of their way to attract institutional participation in IPOs, possibly because, as argued by the bookbuilding literature, they wish to extract information from them about their valuation of the IPO firm's equity: see, e.g., Benveniste and Spindt (1989). On the other hand, the well-known underpricing model of Rock (1986) argues that institutions with private information may distort the IPO share-allocation process, bidding only on the equity of undervalued firms going public, thus leaving retail investors with a disproportionate share of overvalued IPO firm equity. The empirical evidence also suggests the notion that institutional investors have private information about the true long-run value of the shares of the firm going public: for example, Chemmanur, Hu, and Huang (2010), who show that institutions indeed have private information about IPOs, retain their information advantage in post-IPO trading, and are able to realize significant profits from their participation in IPOs (see also, Field and Lowry (2009)).[1] However, while the ability of institutions in general to produce information about IPOs has been well-documented, the ability of specific kinds of institutions to produce information about firms going public and the effects of such information production on the characteristics of the IPOs of these firms has not been studied. The objective of this paper is to fill this gap in the literature by analyzing for the first time how the geographical location of institutional investors affects their information production and investment in IPOs, and the relation between such information production by certain groups of institutions and the characteristics of the IPOs that they invest in.

The starting point of our analysis is an examination of how the geographical location of institutions affects the incentives to acquire information about IPO firms that they may be evaluating for possible investment. Hong, Kubik, and Stein (2005) explores (in a broader context) the hypothesis that investors exchange information and ideas about investing in stocks with each other directly through word-of-mouth communication, and argue that such communication may be easier if the institutions involved are in the same geographical location. However, it is our view that such ease of communication engendered by geographical proximity may be a double-edged sword. On the one hand, if several institutions that are geographically close together share information with each other, they may each have access to the signals available to each institution, thereby increasing the quality of the information available to each of them. On the other hand, if information production about IPO firms is costly, the ability of individual institutions to free-ride on each others' information may dampen their incentives to acquire independent information (signal) about the quality of the IPO firm, so that the precision of the information collectively held by a group of geographically proximate institutions may in fact be lower than that of the information produced by a group of geographically isolated institutions working independently.[2]

[1] A number of non-information related roles have also been postulated for institutions in the IPO process: see, e.g., Ritter and Zhang (2007) or Nimalendran, Ritter, and Zhang (2007).

[2] For a formal model that captures some of these ideas in a rational expectations framework, see the model by

We analyze the implication of the above idea for IPOs in this paper. We start by analyzing whether geographically proximate institutions tend to free-ride on each others' information when they choose IPOs to invest in. If this is indeed the case, the effect of neighboring institutions' IPO equity holdings on a given institution's IPO equity holdings will be greater than those of more distant institutions. This is the first hypothesis we test in this paper.[3]

We now turn to the analysis of how the differences in the incentive to produce information between geographically clustered versus geographically isolated institutions affect the characteristics of the IPOs dominated (in terms of investment) by the two kinds of institutions. We make use of measures of geographical dispersion of the institutions investing in a given IPO to conduct this analysis. If it is indeed the case that more isolated institutions, collectively, have more accurate information about firms going public, they are more likely to have more independent signals collectively. IPO underwriters are likely to extract this more precise information from institutions and use it to determine the final offer price of an IPO using the IPO book building process (see, e.g., Benveniste and Spindt (1989)). Institutions will invest in an IPO firm if the information produced by them is favorable and not invest in the firm when it is unfavorable. Therefore the IPO price revision (from the mid-point of the initial filing range to the offer price) will be increasing in the geographical dispersion of the institutions investing in an IPO, since the IPO offer price will reflect this more accurate favorable information held by institutions.[4] This also implies that the IPO valuation at the

Han and Yang (2013). They study a rational expectations equilibrium model of a competitive market in which traders can learn about a risky asset's payoff from three sources: the market price; costly information acquisition; and communication with other traders through a social network. When traders decide whether or not to acquire costly information, they take into consideration the expected learning through social communication. In equilibrium, information acquisition and asset prices are determined simultaneously. In the above setting, they show that, when information is exogenous, social communication improves market efficiency. However, social communication crowds out information production due to traders' incentive to "free ride" on informed friends and on a more informative price system. Overall, social communication hurts market efficiency when information is endogenous.

[3]While this "neighborhood effect" among mutual fund managers in a general investment setting has been studied in Hong, Kubik, and Stein (2005), the behavior of institutional investors facing a firm going public, where information asymmetry is especially high among institutional investors, is not documented.

[4]To generate this implication, we need two assumptions. First, we assume that underwriters set the mid-point of the initial filing range based on whatever information they have when they file the preliminary prospectus. In other words, this mid-point does not reflect any information generated from institutions. Second, institutions participating in an IPO are more likely to be those with favorable information about that IPO.

offer price will be increasing in geographical dispersion of institutions investing in an IPO.

We now turn to the relationship between geographical dispersion and secondary market valuation. Assume that it is common knowledge to all investors in the immediate secondary market that more isolated institutional investors have more accurate information about an IPO firm's value. Then, the immediate secondary market valuation of an IPO firm will be greater for IPOs dominated by equity holdings from geographically isolated institutions relative to IPOs dominated by geographically clustered institutions (due to a "certification effect"). This implies that the immediate secondary market valuation of an IPO will be increasing in measures of geographical dispersion of institutions investing in an IPO.

We now turn to analyzing the relationship between the geographical dispersion of the institutions investing in an IPO and IPO initial return. Clearly, the initial return on an IPO stock reflects the difference between its IPO valuation and its immediate secondary market (first trading day closing price) valuation. If the relationship between geographical dispersion and IPO valuation is stronger than the relationship between geographical dispersion and secondary market valuation, then the IPO initial return will be decreasing in measures of geographical dispersion of the institutions investing in an IPO. On the other hand, if the relationship between geographical dispersion and IPO valuation is weaker than the relationship between geographical dispersion and secondary market valuation, then the IPO initial return will be increasing in measures of geographical dispersion of the institutions investing in an IPO. Further, assume that all information produced by institutions is reflected in secondary market prices only gradually through time. Then, the long-run post-IPO stock return will be increasing in measures of the geographical dispersion of institutions investing in an IPO as well.

Finally, we study whether the information channel, i.e. clustered institutions are less likely to produce independent signals, compared to isolated institutions, is driving the above findings. We explore the information channel through three empirical exercises. First, we look at the information asymmetry facing a firm in the public equity market and study its relationship with the geographical dispersion of the institutional investors investing in the IPO firm's stock. Second, for each IPO firm, we classify the institutions investing in that IPO firm into geographically isolated and geographically clustered institutions. We then study whether the predictive power of institutional trading for the future stock return of that IPO firm is stronger for geographically isolated institutions compared to geographically clustered institutions, thus directly analyzing the information production argument that we discussed above. We measure institutional trading using the "Net Buy" by institutions, defined as the number of shares purchased by institutions minus the number of shares sold by institutions, normalized by the number of shares outstanding. Third, we look at the relationship between institutional trading and the surprise in the earnings announcements of the IPO firms above market expectations. We study whether trading by geographically isolated institutions is a stronger predictor of earnings surprises than trading by geographically clustered institutions, thereby providing further evidence of the information advantage collectively held by geographically isolated institutions.

We develop our empirical analysis of the implications of the above theoretical arguments making use of data on IPOs between January 1980 to December 2012 from the SDC Global New Issues database. We infer on institutional investments in IPOs using quarterly institutional holdings from Thomson Reuter's Institutional Holdings (13F) database. To obtain the geographical location for each institutions, we manually identify the location of institutional investors using the *Nelson's Directory of Investment Managers* and by searching the filings by institutional investors on the SEC Edgar website. We make use of the analysts earnings forecasts data from Thomson Reuter's Institutional Brokers' Estimate System (I/B/E/S) database and construct measures of information asymmetry facing the IPO firm in the secondary market.

The results of our empirical analysis can be summarized as follows. We start with our empirical analysis on the geographical proximity between institutions and their investments in IPOs. For every pair of institutions investing in the IPO considered, we classify them as "neighbors" if the geographical distance between the two institutions is within 50 miles. We find that a one-percentage-point increase in the aggregate investments in the IPO by neighboring institutions is associated with 1.17 basis points increase in the investment in the same IPO for the institution considered. On the other hand, a one-percentage-point increase in the aggregate investments in the IPO by distant institutions is associated with 0.80 basis points increase in the investment in the same IPO for the institution considered. This finding is consistent with our hypothesis that institutional investors' investment in IPOs is affected more by the investments made by neighboring institutions than those by distant institutions, implying that geographically proximate institutions are more likely to free-ride on each others' information when they choose IPOs to invest in.

Next, we construct a quantitative measure, *geographical dispersion*, which captures the extent to which the investments in IPOs are dominated by geographically isolated institutional investors. We make use of the *geographical dispersion* measure and study its relationship with various characteristics of the IPOs. First, we find that, consistent with our hypotheses, a one-standard-deviation increase in *geographical dispersion* is associated with a 2.3% upward IPO price revision by underwriter(s) and an increase of 0.21 in industry-adjusted Tobin's Q based on IPO offer price. It implies that underwriters extract more accurate (and favorable) signals from participation by geographically isolated institutional investors, compared to the participation by geographically clustered institutional investors, and use that information to revise the offer price upward and sell the IPO at a higher valuation at the offering. Second, we find that a one-standard-deviation increase in *geographical dispersion* is associated with an increase of 0.39 in industry-adjusted Tobin's Q based on first trading day closing price and an increase of 0.42 in industry-adjusted Tobin's Q at the first fiscal quarter end post-IPO. Lastly, we find that a one-standard-deviation increase in *geographical dispersion* is associated with an increase in IPO initial return by 3.2%. We also find that a one-standard-deviation increase in *geographical dispersion* is associated with an increase of one-year size and book-to-market adjusted buy-and-hold abnormal return by 4.9%. This is also consistent with our hypothesis that information produced by institutions is reflected in the secondary market prices gradually through time, and that IPOs dominated by geographically isolated institutional investors have higher initial returns and long-run abnormal stock returns, compared to IPOs dominated by geographically clustered institutions.

The result of our empirical test to analyze the differences in information production between clustered and isolated institutions can be summarized following. We find that higher *geographical dispersion* among the institutions investing in an IPO is associated with lower information asymmetry facing the IPO firm. This finding is robust to different measures of information asymmetry including the standard deviation of analysts forecasts, the analyst forecast error, and the coefficient of variation of analyst forecasts. Second, we find that aggregated net buying by geographically isolated institutions can predict future one-year abnormal holding period returns (adjusted for market returns or matched Fama-French 25 portfolio returns), while aggregated net buying by geographically clustered institutions does not exhibit such predictive power. Specifically, a one-standard-deviation increase in net buying by geographically isolated institutions predicts an increase of 2.9% in subsequent one-year abnormal buy-and-hold returns. Third, we find that aggregated net buying by geographically isolated institutions significantly predicts earnings surprises post-IPO, while aggregated net buying by geographically clustered institutions does not. In terms of economic magnitudes, a one-standard-deviation increase in net buying by geographically isolated institutions predicts an increase of 0.5% in standardized unexpected earnings. These findings further suggest that geographically isolated institutions produce more precise signals collectively, compared to geographically clustered institutions

The remainder of this paper is organized as follows. Section 3.2 briefly reviews the related literature and the contribution of our paper relative to the literature. Section 3.3 discusses the underlying theory and hypothesis for our empirical tests. Section 3.4 describes our data and sample selection procedures and presents summary statistics. Section 3.5 presents our main empirical tests and results. Section 3.6 presents additional tests of the relationship between geography and information production. Section 3.7 concludes.

3.2 Relation to the Existing Literature and Contribution

Our paper is related to several strands in the literature. One strand our paper is related to is the empirical literature on the role of institutional investors in IPOs. Chemmanur, Hu, and Huang (2010) show that institutional trading has predictive power for subsequent long-run IPO performance, even after controlling for publicly available information, suggesting that institutional investors possess private information about IPOs. Our finding regarding the predictive power of institutional trading for long-run IPO stock return is consistent with theirs (see also Boehmer, Boehmer, and Fishe (2006) and Field and Lowry (2009) for similar evidence). We extend their long-run post-IPO return predictability results by highlighting the effect of geographical concentration on institutional investors' incentives to produce information in IPOs.

The broader literature on the role of institutions in IPOs is also related to our paper. Aggarwal (2003) and Hanley and Wilhelm Jr. (1995) document that institutional investors receive significant allocations in underpriced IPOs. Aggarwal (2003) studies IPO allocation and immediate flipping over the first two days after the IPO. Boehmer, Boehmer, and Fishe (2006) study the relation between IPO allocation, flipping, and long-run IPO performance. Ellis, Michaely, and O'Hara (2000) and Ellis (2006) study aftermarket trading by market makers in IPOs.

The theoretical literature on information production by institutions and other investors around IPOs is also related to our paper. Rock (1986) argues that institutional investors with private information about the true long-run value of the shares of firms going public bid only on undervalued shares, leaving retail investors with a disproportionate share of overvalued IPOs. Benveniste and Spindt (1989) build on Rock (1986)'s assumption of informed institutional investors, and argue that the IPO bookbuilding process is a mechanism for underwriters to extract information from institutional investors in order to use it to price shares in the IPO at the appropriate level. Chemmanur (1993) views underpricing as a way of inducing information production by institutional and other investors about the firm going public. See Ritter and Welch (2002) for an excellent review of the related theoretical and empirical literature on IPOs.

Our paper is also connected to the literature on investor networks and information production. Han and Yang (2013) study a rational expectations equilibrium model of a competitive market in which traders can learn about a risky asset's payoff from three sources: the market price; costly information acquisition; and communication with other traders through a social network. They show that, when information acquisition is exogenous, social communication improves market efficiency. However, social communication crowds out information production due to traders' incentive to "free ride" on informed friends and on a more informative price system. In other words, social communication hurts market efficiency when information is endogenous. While our empirical analysis does not focus on social connections, the implications of the above theory broadly apply to our paper insofar as geographic proximity captures the ease of communication among investors.

Finally, our paper is related to the literature on geographical proximity and information sharing among investors. For example, Hong, Kubik, and Stein (2005) show that mutual fund managers located in the same city tend to make correlated investment decisions, suggesting that portfolio managers share investment ideas with each other through word-of-mouth communication. Building on Hong, Kubik, and Stein (2005), Pool, Stoffman, and Yonker (2015) find that the overlap in stock holdings and trades between funds whose managers living in the same neighborhood is considerably high than that of funds whose managers live in the same city but in different neighborhoods. In their investigation of stock trades by individual investors, Ivković and Weisbenner (2007) find strong evidence of correlated trades among individual investors in the same geographic location and attribute about one-quarter to one-half of the correlation between their trades to word-of-mouth communication.[5] Our paper contributes to this literature by showing that such ease of communication induced by geographical proximity can negatively affect information production.

[5] Shiller and Pound (1989) present survey evidence that both institutional and individual investors may be influenced by peer communications.

3.3 Theory and Hypotheses

The theoretical framework that we use to develop our testable hypotheses is adapted from the model of Han and Yang (2013). Han and Yang (2013) study a rational expectations equilibrium model of a competitive market in which traders can learn about a risky asset's payoff from three sources: the market price; costly information acquisition; and communication with other traders through a social network. We assume that institutional investors make use of two kinds of information to decide on the IPO firms to invest in as well as to value these IPO firms. First, each institution has its own (freely available) signal that can help in the above task, for example, the signal may be based on their prior experience with investing in IPO stocks. Second, institutions may produce an independent signal at a cost about each IPO firm that they are considering investing in, with the precision of this signal increasing in the amount of resources they devote to information production.

We assume that institutions are able to share each others' information. Further, each information sharing between any two institutions become easier when they are geographically closer to each other ("local information sharing" effect). On the other hand, if an institution has access to the information available to other institutions, it reduces its incentive to incur costs for producing information independently on its own, thereby reducing the precision of the information produced by each institution ("free riding on neighbors" effect). The above induces the following ambiguous relationship between the geographical dispersion among the institutions investing in an IPO and the precision of the aggregate amount of information available to them. To see this, let us consider two extremes. Consider first the case where the precision of the signal freely available to each institution is very high while at the same time, the cost to precision ratio of information production by institutions is also high. In this case, the precision of the aggregate amount of information available collectively to all institutions will be decreasing in their geographical dispersion, since the advantage to institutions of being able to more precisely share each other's information when institutions are geographically close to each other overcomes any disadvantage arising from the dampening of each institution's incentive to produce information (arising from the ability to free ride on the information signal of other institutions that are geographically close to them). Consider the other extreme case where the precision of the the signal freely available to each institution is very low, while the cost to precision ratio of information production by institutions is also low. In this case, the precision of the aggregate amount of information available collectively to all institutions will be increasing in their geographical dispersion, since the advantage to institutions of being able to more precisely share each other's information is not enough to overcome the disadvantage arising from the dampening of each institution's incentive to produce information when institutions are geographically close to each other.

In the following, we use the above theoretical framework to develop testable hypotheses to analyze the relationship between the geographical dispersion among institutions investing in an IPO and offer price revision; valuation of the IPO firm at offering; valuation of the IPO firm at the secondary market; IPO initial return; the long-run post-IPO stock return; and finally, the information asymmetry faced by the IPO firm in the secondary market. We also develop testable hypotheses for studying whether institutional trading by geographically clustered or geographically isolated institutions will have stronger predictive power for future stock returns and earnings surprises.

Our first hypothesis is related to the "word-of-mouth" effect documented in Hong, Kubik, and Stein (2005). Institutions that are closer together are more likely to free ride on each others' information about the IPO firms that they propose to invest in relative to those that are geographically more isolated. This implies that the effect of a neighboring institution's IPO equity holdings on a given institution's IPO equity holdings will be greater than those of more distant institutions (**H1**).

We now turn to developing testable hypotheses to analyze the relationship between the geographical dispersion among institutions investing in an IPO and various IPO characteristics. First, we look at the offer price revision and IPO valuation at the offer price. On the one hand, if the "free riding on neighbors" effect dominates the "local information sharing" effect, then more isolated institutions, collectively, will have more accurate information about firms going public (since they produce more precise signals). IPO underwriters are likely to extract this information from institutions and use it to determine the final offer price of an IPO using the IPO book building process (see, e.g., Benveniste and Spindt (1989)). Institutions will invest in an IPO firm if the information produced by them is favorable and not invest in the firm when it is unfavorable. Therefore the IPO price revision (from the mid-point of the initial filing range to the offer price) will be increasing in the geographical dispersion of the institutions investing in an IPO, since the IPO offer price will reflect this more accurate (and favorable) information held by geographically isolated institutions (**H2A**). Further, the IPO valuation at the offer price will also be increasing in geographical dispersion of institutions investing in an IPO (**H3A**). On the other hand, if the “local information sharing" effect dominates the “free riding on neighbors" effect, then more clustered institutions, collectively, will have more accurate (and favorable) information about firms going public since they have access to more signals. In this case, the IPO price revision will be decreasing in the geographical dispersion of the institutions investing in the IPO, since the IPO offer price will reflect this more accurate favorable information held by geographically clustered institutions (**H2B**). As a result, the IPO valuation at the offer price will be decreasing in the geographical dispersion of the institutions investing in an IPO (**H3B**).

In the secondary market, the geographical dispersion among institutions investing in the IPO becomes public knowledge to all investors, both institutional investors and retail investors. Similarly, if the “free riding on neighbors" effect dominates the “local information sharing" effect, then the geographically isolated institutions, collectively, will have more accurate information about firms going public, and the secondary market valuation of an IPO firm will be greater for IPOs dominated by equity holdings from geographically isolated institutions. Thus, the secondary market valuation of an IPO will be increasing in measures of geographical dispersion of institutions investing in an IPO (**H4A**). At the same time, the initial return on an IPO stock reflects the difference between IPO valuation and immediate secondary market (first trading day closing price) valuation. The favorable information contained in a geographically dispersed institutional shareholder base will attract further institutional and retail investors, and translate into higher valuation (relative to valuation at the offer price) for the IPO firm.[6] Thus the IPO initial return will be increasing in measures of geographical dispersion of institutions investing in an IPO (**H5A**) as well. Further, assuming the information produced by institutions is impounded into post-IPO stock prices only gradually through time, then the long-run post-IPO stock return of an IPO firm will be increasing in measures of geographical dispersion of the institutions investing in that firm's IPO (**H6A**). Further, as geographically isolated institutions collectively have more accurate information about firms' fundamentals, their participation in the IPO and subsequent trading activities in the secondary market reduces information asymmetry facing the IPO firm. Thus, the information asymmetry facing an IPO firm will be decreasing in measures of geographical dispersion of institutions investing in an IPO (**H7A**).

On the other hand, if the "local information sharing" effect dominates the "free riding on neighbors" effect, the secondary market valuation of an IPO will be decreasing in measures of geographical dispersion of institutions investing in an IPO (**H4B**); then the IPO initial return will be decreasing in measures of geographical dispersion of institutions investing in an IPO (**H5B**). Further, in this scenario, the long-run post-IPO stock return will be decreasing in measures of geographical dispersion of institutions investing in the firm's IPO (**H6B**). And finally, the information asymmetry facing an IPO firm will also be decreasing in measures of geographical dispersion of the institutions investing in that IPO (**H7B**).

We next examine which group of investors, geographically clustered or geographically isolated institutions, are more collectively informed about the firm going public. Assuming that institutions are able to generate private information about the intrinsic value and future performance of firms going public, secondary market trading (net buying) of institutions in the equity of the IPO firm will have predictive power for its subsequent long-run stock return performance and earnings surprises. Specifically, we study whether the secondary market trading by geographically clustered

[6]The relationship between breadth of investor base and asset valuation is theoretically modeled in Merton (1987).

or geographically isolated institutions has stronger predictive power for future stock returns and earnings surprises. If the “free riding on neighbors" effect dominates the “local information sharing" effect, then geographically isolated institutions will have more accurate information collectively about the IPO firm. Thus, the predictive power of secondary market trading by geographically isolated institutions for long-run IPO stock returns will be greater (**H8A**), and the predictive power of secondary market trading by geographically isolated institutions for IPO firm’s earnings surprises will also be greater (**H9A**). If the “local information sharing" effect dominates the “free riding on neighbors" effect, then geographically clustered institutions will have more accurate information collectively about the IPO firm. In this scenario, the predictive power of post-IPO secondary market trading by geographically clustered institutions for long-run IPO stock returns will be greater (**H8B**), and the predictive power of secondary market trading by geographically clustered institutions for IPO firm’s earnings surprises will also be greater (**H9B**).

3.4 Data and Summary Statistics

We first identify all IPOs conducted in the U.S. markets from January 1980 to December 2012 using the Thomson Financial’s Securities Data Company (SDC) Global New Issues database. We exclude certificates, ADRs, shares of beneficial interest, units, closed-end funds, REITs, IPOs with an offer price less than $5, and stocks that are not list on Center for Research in Security Prices (CRSP) within 5 days of SDC’s IPO date. We use CRSP to identify the exchange on which each stock first began trading, and retain only stocks that are traded on NYSE, AMEX, and NASDAQ. For each IPO firm, we collect the issue date, offer price, initial filing range, proceeds, underwriter name(s), SIC code, and whether the issue is backed by a venture capitalist from SDC. We use underwriter reputation rankings from Loughran and Ritter (2004) (based on earlier work by Carter and Manaster (1990)), which ranks each underwriter from zero to nine, with higher ranks representing higher reputation underwriters. We also collect data on firm age, i.e., the number of years since the company was founded, at the time of the IPO.[7]

Lacking public data on participation of institutional investors in IPOs, we use reported quarterly holdings of institutions from Thomson Reuter's Institutional (13F) Holdings database to construct proxies for institutional investors' participation in IPOs. We use the first reported holdings within three months of the offer date for each IPO as our proxy for initial IPO participation.[8] To obtain the geographical location for each institution, we manually identify the location (zip code) of the headquarters of the institutional investors using the *Nelson's Directory of Investment Managers* and by searching the filings by institutional investors on the SEC Edgar website. We exclude institutions without valid location information. The headquarter location of the IPO firm comes from Compustat.

Our initial sample consists of 5,590 IPOs from January 1980 to December 2012. We present summary statistics of these IPOs in Table 3.1. The mean *Initial Return*, defined as the difference between the offer price and the first-day closing price divided by the offer price, is 19.6%. The mean *Price Revision*, measured as the percentage difference between offer price and the mid-point of the initial filing range, is 1.1%. The average *Age* of the firm at the time of the offering is 16 years. The mean *Total Proceeds* of the IPO is 88.1 million. Table 1 also reports valuation of the IPO firm at the offer price (*QOPAdj*) and in the immediate secondary market (*QFTDAdj*, and *QFQAdj*). The valuation measure we use is Tobin's Q, which is the ratio of the market value of assets over the book value of assets. We calculate the market value of assets as the book value of assets minus the book value of equity plus the product of the number of shares outstanding and share price. We calculate industry-adjusted Tobin's Q as the raw Tobin's Q minus the median Tobin's Q in the 2-digit SIC industry. We measure the secondary market valuation using the first trading day closing price as the share price in the above definition (*QFTDAdj*), and the share price at the end of the first post-IPO fiscal quarter (*QFQAdj*). The book value of assets and the book value of equity are measured as of

[7]We thank Jay Ritter for making the data on firms' founding dates and underwriter reputation rankings available on his website.

[8]Reported holdings are potentially different from initial allocations. For a discussion of the rationales and issues with using reported holdings as proxies for IPO allocations , see, e.g., Ritter and Zhang (2007).

the first post-IPO quarter.

3.5 Empirical Tests and Results

3.5.1 Neighbor Effect on Participation of Institutions in the IPO

As discussed in our hypothesis **H1**, if geographically proximate institutions tend to free-ride on each others' information when they choose IPOs to invest in, then we would observe the effects of neighboring institutions' investments in the IPO be greater than those of more distant institutions. To test this hypothesis, we regress individual institution's holding in an IPO on the total holdings of neighboring institutions (*Neighbor Holdings*), and total holdings of all distant institutions (*Non-neighbor Holdings*). The neighboring institutions are defined as those institutions headquartered within 50 miles of the institution considered, while the distant institutions are defined as those institutions headquartered beyond 50 miles of the institution considered.[9]

To control for local bias in institutional investments (Coval and Moskowitz (1999)), we include in our regression a *Local* dummy, which equals 1 if the institution is headquartered within 50 miles of the IPO firm, and 0 otherwise. We also control for the size of the institution, defined as the natural logarithm of the total net assets of the institution, *Log(TNA)*, since larger institutions are more likely to participate in an IPO and receive more allocations. In addition, we control for year fixed effects, industry fixed effects and institution fixed effects.

The first column of Table 3.3 reports the baseline result, where institutions are defined as "neighbor" if they are headquartered within 50 miles of each other. We find that both the coefficient of *Neighbor Holdings* and that of *Non-neighbor Holdings* are positive and significant, indicating that institutions' participating in an IPO is affected by both types of institutions. Importantly, the coefficient of *Neighbor Holdings* is about 50% larger than that of *Non-neighbor Holdings*, and

[9] In our baseline result, two institutional investors are defined as "neighbor" if they are headquartered within 50 miles of each other. We also vary the definition of "neighbors" using 100 miles and 200 miles, and obtain similar results.

a Wald-test examining the equality of the two coefficient yields a p-value less than 0.001. This suggests that the effect of neighboring institutions' participation in an IPO is much stronger than that of distant institutions. In terms of economic magnitudes, a one-percentage-point increase in the holdings of the IPO stock by the neighboring institutions is associated with 1.17 basis points increase in the holding of the same stock by the institution considered while a one-percentage-point increase in the holdings of the IPO stock by the distant institutions is associated with only 0.80 basis points increase in the holding by the institution considered.

In models (2) and (3) of Table 3.3, we vary the definition of "neighbor" using 100 miles and 200 miles as the cut-off points and re-estimate the regression. We find qualitatively similar results as those in the baseline regression in model (1). To test whether our findings are driven by cities with concentrated institutions, we further exclude institutions located in New York and Boston metropolitan statistical areas (MSAs). The results, reported in model (4), continue to show that neighboring institutions' participation in an IPO has a greater impact on the institution's participation in the IPO compared to that of distant institutions, and the economic magnitude becomes greater. Specifically, a one-percentage-point increase in the holdings of the IPO stock by the neighboring institutions is associated with 3.09 basis points increase in the holding of the same stock by the institution considered while a one-percentage-point increase in the holdings of the IPO stock by the non-neighbor institutions is associated with only 1.18 basis points increase. Consistent with the local information advantage story (Coval and Moskowitz (1999)), Table 2 also suggests that institutions are more likely to participate in the IPOs of local firms. Also, larger institutions tend to make larger investments in IPO stocks.

3.5.2 Geographical Dispersion and IPO Characteristics

In this section, we study how the differential incentives to produce information by geographically clustered versus isolated institutions are related to the characteristics of the IPO they participated. Specifically, we construct a measure of geographical dispersion among institutional investors investing in a given IPO, and empirically analyze the relationship between this geographical dispersion measure and a list of IPO characteristics including offer price revision, IPO valuation at the offer price, IPO valuation in the secondary market, IPO initial return (i.e. IPO underpricing), and IPO long-run stock returns.

3.5.2.1 Measuring the Geographical Dispersion among Institutional Shareholders

To construct the geographic dispersion measure, we calculate the weighted-average geographic distance among institutional shareholders of a firm. In particular, for each institutional shareholder of the firm, we calculate the average geographic distance between the institution and all institutions in the firm, weighted by their respective fractional holdings in the firm. This measure captures the average distance between an institutional shareholder and its peers. To measure geographic distance between a pair of institutions, we define an indicator that equals one if the two institutions are headquartered more than 50 miles away from each other, and zero otherwise. We then calculate a weighted-average of the geographic distance across all institutional shareholders of the firm, again weighted by their fractional holdings. This weighting scheme ensures that institutions that are likely to be more influential, i.e., those with larger holdings in the firm, receive greater weights in determining geographic dispersion among shareholders. Specifically, geographical dispersion among institutional investors of IPO firm c is defined as,

$$G_c = \sum_{i \in S} w_{i,c} \sum_{j \in S} w_{i,c} I(Dist_{ij} > 50) \tag{3.1}$$

where $w_{i,c}$ is the holdings of institution i in IPO firm c as a fraction of total institutional holdings within three months of the IPO date; S is the set of institutional shareholders in firm c at first calendar quarter ending after the IPO; $I(Dist_{ij} > 50)$ is an indicator variable for whether the geographical distance between institutions i and j is more than 50 miles.

Table 3.2 presents summary statistics for the geographical dispersion measure. The average geographical dispersion for the IPO firms is 0.700 and there is a fair degree of cross-sectional variation across IPO firms. Table 3.2 also presents summary statistics for other institutional shareholder characteristics. The average number of institutions holding the equity of the IPO firm is 26 and the mean *Inst. ownership* is 23.0%.[10] We define *Inst. ownership concentration* as the Herfindahl Index of institutional ownership concentration based on each institution's holding as a percentage of total holdings of 13F institutions. The average *Inst. ownership concentration* is 0.188. *Inst. ownership* and *Inst. ownership concentration* are two important control variables of institutions' participation in our study. *Inst. ownership* reflect the aggregate "level" of institutional participation, while *Inst. ownership concentration* reflects the "concentration" of institutional participation.

3.5.2.2 Geographical Dispersion and IPO Offer Price Revision

In this subsection, we study the relationship between geographical dispersion of institutional investors and the IPO offer price revision. We estimate the following OLS regression,

$$y_j = \alpha + \beta G_j + \sum \phi' Z_j + \epsilon_j \tag{3.2}$$

The dependent variable y is offer price revision (*Price Revision*) and firm valuation at the offer price (*QOPAdj*). The definitions for both variables are detailed in Section 3.4. The main independent variable of interest is *Geographic dispersion* (G). The control variables Z include *Inst. ownership*, *Inst. ownership concentration*, and IPO offering and firm characteristics. Specifically, we control for *Log(Reputation)*, defined as the natural logarithm of underwriter reputation ranking. Underwriter reputation has been shown in the literature to be an important determinant of various IPO characteristics. We also control for IPO offer size *Log(Proceeds)*, which is the natural logarithm of IPO total proceeds. Further, we include *Log(Age+1)*, the natural logarithm of firm age plus one, as a control variable since there is less uncertainty associated with older firms. We use two dummies for hi-tech (*High-Tech*, equals to one if the IPO firm is in high-tech industry; see Loughran and

[10] The average institutional ownership we report here is consistent with that reported in Field and Lowry (2009).

Ritter (2004) for details) and VC-backed (*VC backed*) firms. High-tech and VC-backed firms tend to be younger, higher growth companies and therefore, are expected to have higher price revision during the book-building process. In addition, we include a dummy for IPOs issued during the bubble periods (*Bubble*, equals to one for IPOs in 1990 and 2000). IPOs issued during bubble periods are likely to have higher price revision and valuation at offering. We also control for market movement prior to the issue date of the IPO (*Prior market return*, defined as absolute return on the CRSP value-weighted index one month before the IPO issue date) since market movement before IPO issue date affects the investors' demand for IPO shares, and therefore the eventual IPO offer price. Finally, *Lockup* is a dummy variable that equals to one if the IPO has a lock up provision and *Financial* is a dummy variable that equals to one if the IPO firm is in the financial industry (with the first-two digits of SIC code being 60-63 or 67).

We present the results in Table 3.4. Model (1) is our baseline regression and model (2) includes industry fixed effects and year fixed effects. In both regressions, the coefficient of *Geographic dispersion* is positive and statistically significant, suggesting that greater geographical dispersion of institutional investors participating in an IPO is associated with greater offer price revision. The findings provide support for our hypothesis **H2A**, instead of hypothesis **H2B**. It indicates that geographically isolated institutions collectively have more accurate information about the firm going public, and IPO underwriters extract this more precise (and favorable) information from institutions to determine the final offer price of an IPO. Therefore, IPO underwriters are more likely to revise the offer price up during the book-building process. In terms of economic magnitudes, a one-standard-deviation increase in *Geographic dispersion* is associated with an upward price revision of approximately 2.3%.

Our regressions in Table 3.4 also show that the IPO offer price revision increases with the size of the offering and prior one-month stock market return and it decreases with institutional ownership, institutional ownership concentration, the reputation of the underwriters, and age of the IPO firm; IPO valuation at the offer price increases with institutional ownership concentration, proceeds of the offer, and reputation of the underwriters, and it decreases with institutional ownership, the age of the IPO firm, and prior one-month stock market return. Further, IPO offer price revision is are higher for VC-backed and high-tech firms and during the IPO bubble years.

3.5.2.3 Geographical Dispersion and IPO Valuation

In this subsection, we study the effect of geographical dispersion among institutional investors on the IPO valuation at offer price. Similar to regression model in Eq. (3.2), we regress IPO valuation at offer price on geographical dispersion among institutional investors and other controls. As described in Section 3.4, we measure IPO valuation using industry-adjusted Tobin's Q (*QOPAdj*).

We control for IPO offer size *Log(Proceeds)*, which is the natural logarithm of IPO total proceeds. Further, we include *Log(Age+1)*, the natural logarithm of firm age plus one, as a control variable since there is less uncertainty associated with older firms. We also use two dummies for hi-tech (*High-Tech*, equals to one if the IPO firm is in high-tech industry; see Loughran and Ritter (2004) for details) and VC-backed (*VC backed*) firms. High-tech and VC-backed firms tend to be younger, higher growth companies and therefore, are expected to have higher valuation at offering. In addition, we include a dummy for IPOs issued during the bubble periods (*Bubble*, equals to one for IPOs in 1990 and 2000). IPOs issued during bubble periods are likely to have higher valuation at offering. We also control for market movement prior to the issue date of the IPO (*Prior market return*, defined as absolute return on the CRSP value-weighted index one month before the IPO issue date) since market movement before IPO issue date affects the investors' demand for IPO shares, and therefore the eventual IPO valuation. Finally, *Lockup* is a dummy variable that equals to one if the IPO has a lock up provision and *Financial* is a dummy variable that equals to one if the IPO firm is in the financial industry (with the first-two digits of SIC code being 60-63 or 67).

We present the results in Table 3.5. Model (1) is our baseline regression and model (2) includes industry fixed effects and year fixed effects. In both regressions, the coefficient of *Geographic dispersion* is positive and statistically significant, suggesting that greater geographical dispersion of institutional investors participating in an IPO is associated with higher IPO valuation. The findings provide support for our hypotheses **H3A**, instead of hypothesis **H3B**.. It indicates that geographically isolated institutions collectively have more accurate information about the firm going public, and IPO underwriters assign a higher valuation to the IPO firm. In terms of economic magnitudes, a one-standard-deviation increase in *Geographic dispersion* is associated with an increase of approximately 0.21 in the industry-adjusted Tobin's Q of the IPO firm at the offer price.

Table 3.5 also shows that the IPO valuation increases with the size of the offering and the reputation of its underwriter(s). Also IPO valuation decreases with the age of the IPO firm. Further, IPO valuation is higher for VC-backed firms, high-tech firms, and firms went public during the IPO bubble years.

3.5.2.4 Geographical Dispersion and Secondary Market Valuation

In this subsection, we study the effect of geographical dispersion among institutional investors on the immediate secondary market valuation of the IPO firms. Similar to the regression model in Eq. (3.2), we regress secondary market valuation on geographical dispersion among institutional investors and other controls. As described in Section 3.4, we measure secondary market valuation using industry-adjusted Tobin's Q (*QFTDAdj* and *QFQAdj*).

We control for underwriter reputation, *Log(Reputation)*, as we expect firms underwritten by higher reputation underwriters to receive higher valuations. We also control for the age of the firm, *Log(Age+1)*, since younger firms are likely to have more growth opportunities and thus higher valuations. We also include *VC backed* and *High-Tech* in our regressions since VC-backed and hi-tech firms are expected to have more growth options and higher valuations. In addition, we control for IPO offer size, *Log(Proceeds)*, and a list of dummies including *Lockup*, *Financial*, and *Bubble*.

We report the results in Table 3.6. The dependent variable in models (1) and (2) is *QFTDAdj* and that in models (3) and (4) is *QFQAdj*. Model (1) and (3) include industry fixed effects and year fixed effects. The coefficient on *Geographic dispersion* is positive and statistically significant at the 1% level in all four specifications, suggesting that greater geographical dispersion among institutional investors participating in an IPO is associated with higher IPO valuation in the immediate secondary market. The findings provide support for our hypothesis **H4A**, , instead of hypothesis **H4B**. It indicates that secondary market valuation also reflects the more accurate (and favourable) information held collectively by geographically isolated institutions. In terms of economic magnitudes, a one-standard-deviation increase in *Geographic dispersion* leads to an increase of 0.39-0.42 in industry-adjusted Tobin's Q.

The results in Table 3.6 also show that younger firms, high-tech firms, VC-backed firms, firms offered during the bubble years, as well as those underwritten by higher reputation underwriters and have higher institutional ownership concentration receive higher secondary market valuations. On the other hand, firms with higher institutional ownership and firms have lock-up provisions receive lower secondary market valuations.

3.5.2.5 Geographical Dispersion and IPO Initial Return

We study the effect of geographical dispersion among institutional investors on IPO initial returns by regressing *Initial Return*, also commonly referred as "underpricing" in the literature, on *Geographic dispersion* and controls. We control for underwriter reputation and IPO offer size since prior literature shows that these variables are significant predictors of IPO underpricing.[11] Sherman and Titman (2002) predict greater underpricing when the cost of investors' information acquisition is greater, for example, due to increased uncertainty about the IPO firm. We control for such uncertainty by including firm age and dummy variables for hi-tech and VC-backed firms. We further control for market movement in the pre-IPO period using the return on the CRSP value-weighted

[11] Loughran and Ritter (2004) docuemtn a negative relationship between underpricing and IPO offer size in the 1980s and the beginning of the 2000s but a positive relationship in the 1990s. Carter and Manaster (1990) document a negative relationship between underwriter reputation and underpricing using data from 1980s. They argue that prestigious underwriters are associated with lower risk offerings, and consequently lower underpricing. However, in later studies using data from 1990s and 2000s, Aggarwal, Krigman, and Womack (2002) document a positive relationship between underwriter reputation and underpricing.

index over the one month period prior to the IPO (*Prior market return*) to account for the flow of new information to the equity market prior to the IPO.

The results, presented in Table 3.7, show that the coefficient on *Geographic dispersion* is significant and positive in both specifications. This finding is consistent with hypothesis **H5A**, but not **H5B**. The positive relationship between geographical dispersion among institutions participating in the IPO and IPO initial return indicates that geographically isolated institutions hold more accurate (and favourable) information. This favourable information embedded in the participation of more geographically isolated institutions further attracts higher demand from other institutional and retail investors during the first day trading in the public market, and leads to greater IPO initial return. In terms of economic magnitudes, a one-standard-deviation increase in *Geographic dispersion* is associated with an increase of approximately 2.8% in IPO initial returns, which is a 14% increase relative to the sample mean of 19.8 percentage points. We also find younger firms, VC-backed firms, hi-tech firms, and IPOs issued during bubble years are associated with higher initial returns.

3.5.2.6 Geographical Dispersion and Post-IPO Stock Return Performance

In this subsection, we study the effect of geographical dispersion of institutional investors on the post-issue stock return performance of IPO firms. We regress market-adjusted one-year holding period return (*1YrHPRAdjMM*) and book-to-market and size adjusted one-year holding period return (*1YrHPRAdjFF25*) on *Geographic dispersion* and controls. *1YrHPRAdjMM* is the IPO firms' one-year holding period return calculated by compounding daily returns over 252 trading days after the IPO (excluding the first trading day's return) minus the CRSP value-weighted market return during the same period; *1YrHPRAdjFF25* is the IPO firms' one-year holding period return calculated by compounding daily returns over 252 trading days after the IPO (excluding the first trading day's return) minus the matched Fama/French 25 size and book-to-market portfolio buy-and-hold value-weighted return during the same period. If an IPO firm is delisted before the end of the one-year period, the returns of the IPO firm and benchmark returns are compounded until the delisting date.

Similar to the regression in Eq. (3.2), we control for *Inst. ownership* and *Inst. ownership concentration*. We also control for underwriter reputation, offer size, firm age, IPO initial return (i.e., underpricing), and indicators for whether the IPO has a lockup provision and whether the IPO is backed by VCs.[12]

Table 3.8 presents the regression results. The dependent variable in models (1) and (2) is *1YrHPRAdjFF25*, and that in models (3) and (4) is *1YrHPRAdjMM*. The coefficient for *Geographic dispersion* is positive and statistically significant across all four specifications. The economic magnitude is also significant: for example, model (2) shows that a a one-standard-deviation increase in *Geographic dispersion* is associated with an increase of approximately 2.9% in book-to-market and size adjusted one-year holding period return. The result is consistent with our hypothesis **H6A**, instead of hypothesis **H6B**, indicating that while geographically isolated institutions have more precise and favorable information about the IPO firm, the information produced by institutions is reflected in the secondary market only gradually over time. Hence the long-run post-IPO stock return increases with the geographically dispersion of the institutions participating in the IPO.

Table 3.8 also shows that larger IPOs, higher IPO initial returns, and the use of lock-up provisions are associated with lower post-IPO long-run abnormal stock returns. Moreover, firm age is weakly positively correlated with market-adjusted one-year holding period returns. We do not find a statistically significant relationship between underwriter reputation and post-IPO long-run abnormal stock returns.

[12]Using a sample of IPOs from 1975 to 1984, Ritter (1991) documents a positive effect of offer size and firm age on the post-issue long-run stock returns of IPO firms, and a negative effect of underpricing on the same long-run stock returns.

3.6 Tests of the Relation between Geography and Information Production

In the previous section, we show that higher geographical dispersion of the institutions investing in an IPO is associated with higher IPO price revisions, higher IPO valuation and secondary market firm valuation, larger IPO initial returns, as well as greater long-run post-IPO stock returns. These results indicate that the “free riding on neighbors" effect dominates the “local information sharing" effect, i.e. geographically isolated institutions hold more precise information, compared to geographically clustered institutions, as the disadvantage arising from the dampening of each institution’s incentive to produce information overrides any advantage to institutions of being able to more precisely share each other information. In this section, we test the information advantage of geographically isolated institutions directly in a variety of ways. First, we study relationship between geographical dispersion of the institutions investing in an IPO and the information asymmetry facing the firm in the secondary market. Second, we study the relative predictive power of geographically isolated versus geographically clustered institutions’ trading for future abnormal stock returns and earnings surprises of the IPO firm.

3.6.1 Information Asymmetry Facing an IPO Firm

If more isolated institutional investors have more accurate information about an IPO firm’s value, the extent of information asymmetry facing an IPO firm will be smaller if its equity is predominantly held by such institutions. As a result, we will expect that analysts produce higher quality research about IPO firms dominated by geographically isolated institutions.

To test the relationship between information asymmetry facing the IPO firm and the geographical dispersion of institutional shareholders, we retrieve analyst earnings forecast data from I/B/E/S. Specifically, for each IPO firm, we retrieve sell-side analyst earnings forecasts (within 90 days of the announcement date of actual earnings) for the first fiscal year post-IPO. We employ four measures for analyst forecasts. The first measure is the mean-squared error of analysts' forecasts (*MSE*). We measure forecast error as the absolute difference between the average earnings forecast and the actual earnings per share divided by the price per share at the time of the forecast. The second measure is the standard deviation of analysts' forecasts (*Dispersion*). The third measure is the coefficient of variation of analyst forecasts (*COV*), which is defined as the ratio of standard deviation to the absolute value of the average of analyst forecasts. The fourth measure is the number of analysts following the firm. We take the log of each measure to reduce skewness.

We regress each of the four measures of information asymmetry facing the IPO firm on the geographical dispersion measure and controls. We control for institutional ownership, ownership concentration, underwriter reputation, total offer size, firm age, and a list of dummies including *Lockup*, *VC backed*, *High-Tech*, *Financial*, and *Bubble*. These control variables are associated with various IPO characteristics and are potentially significant determinants of the information environment facing the IPO firm in the secondary market. Table 3.9 reports the regression results. The coefficient for *Geographic dispersion* is positive and statistically significant for three out of four information asymmetry measures including *Log(MSE)*, *Log(Dispersion)*, and *Log(COV)*. On the other hand, we do not find statistically significant relationship between *Geographic dispersion* and *Log(# of Analysts)*. These findings indicate that while IPOs dominated by geographically isolated institutions, compared to those dominated by geographically clustered institutions, are followed by a similar number of analysts after they become public, the analysts covering the former tend to produce earnings information that is of higher quality. The results are consistent with hypothesis **H7A**, instead of hypothesis **H7B**, indicating that information asymmetry facing the IPO firm in the secondary market is lower for IPO firms dominated by geographically isolated institutions.

The results presented in Table 3.9 also show that information asymmetry in the secondary market is generally lower for older firms, high-tech firms, and financial firms. On the other hand, information asymmetry in the secondary market is generally higher for firms with higher institutional ownership concentration.

3.6.2 Predictive Power of Institutional Trading

Chemmanur, Hu, and Huang (2010) show that institutional trading has predictive power for long-run IPO performance. In this section, we study whether geographically isolated institutions produce more accurate information about the intrinsic value and future performance of the IPO firm, compared to geographically clustered institutions. Specifically, we first study whether trading by geographically isolated institutions is a stronger predictor of future stock returns post-IPO. Second, we study whether trading by geographically isolated institutions is a stronger predictor of earning surprises post-IPO.

We measure institutional trading (*Net Buy*) as the change in the fractional ownership of the IPO stock by institutional investors between the first and the second full quarter following the IPO. For an institution in an IPO stock, we first calculate the average geographical distance between the institution and all other institutional shareholders of the stock.[13] We classify institutions with the average distance below the median as “clustered" institutions and those above the median as “isolated" institutions. To capture the aggregate information held by geographically clustered institutions and geographically isolated institutions, we sum up trading by “clustered" institutions (denoted as *Net Buy Clustered*) and that by “isolated" institutions (denoted as *Net Buy Isolated*) separately. Table 3.2 reports the summary statistics for *Net Buy Clustered* and *Net Buy Isolated*. On average, geographically isolated (clustered) institutions sell approximately 1.1% (1.3%) of the total number of shares outstanding of the IPO firm in the first full quarter post-IPO. We now turn to examining the predictive power of trading by these two types of institutions for long-run stock returns and earnings surprises.

[13]In untabulated results, we also rank the institutional shareholders according the number of “neighbors" each institutions have and obtain quantitatively similar results. A pair of institutions are defined as “neighbor" when they are headquartered within 50 miles of each other.

3.6.2.1 Predictive Power of Institutional Trading For Subsequent Stock Returns

To examine the predictive power of institutional trading for subsequent returns, we regress stock returns during the 12 months immediately following the first two full quarter post-IPO on institutional trading during the second full quarter post-IPO.[14] Similar to Section 3.5.2.6, we adjust our raw holding period returns using contemporaneous market returns and the returns on Fama-French 25 portfolios. Specifically, *HPRAdjMM* is the IPO firm's one-year market-adjusted holding period return, defined as the buy-and-hold returns of the stock during the 12 months minus the buy-and-hold CRSP value-weighted market returns; *HPRAdjFF25* is the IPO firm's one-year size and book-to-market adjusted holding period return, defined as the buy-and-hold returns of the stock during the 12 months minus the buy-and-hold returns of the matched Fama-French 25 portfolios. We regress *HPRAdjMM* and *HPRAdjFF25* on *Net Buy Clustered*, *Net Buy Isolated*, and control variables. We use the same set of controls as in Eq. (3.2). Table 3.10 presents the regression results. The dependent variable in models (1) and (2) is *HPRAdjFF25*, and that in models (3) and (4) is *HPRAdjMM*. The coefficient for *Net Buy Isolated* is positive and statistically significant across four models, while that for Net Buy Clustered is statistically indistinguishable from zero. In model (1), the coefficient for *Net Buy Isolated* is 0.721, indicating that a one-standard-deviation increase in *Net Buy Isolated* is associated with an increase of approximately 2.9% in subsequent 12-month size and book-to-market adjusted abnormal holding period returns. This finding is consistent with hypothesis **H8A**, instead of hypothesis **H8B**, that geographically isolated institutions have more accurate information collectively.

3.6.2.2 Predictive Power of Institutional Trading For Earnings Surprises

To examine the source of the return predictability of trading by geographical isolated institutions, we test whether geographical isolated institutions possess superior ability to predict earnings

[14]We measure institutional trading as the change in the fractional ownership of the IPO stock by institutional investors between the first and the second full quarter following the IPO.

surprises of the IPO firm. Following Livnat and Mendenhall (2006), we define standardized unexpected earning (SUE) as analyst forecast error scaled by stock price. More specifically,

$$SUE = \frac{EPS_A - EPS_E}{P_0} \tag{3.3}$$

where EPS_A is the actual earnings per share (EPS) reported in the I/B/E/S database, and EPS_E is the median EPS forecast by analysts during the period from the IPO date to the end of the first full calendar quarter post-IPO. If an analyst makes multiple forecasts during this period, we retain only the last forecast. P_0 is the stock price at the end of the first full calendar quarter post-IPO. We regress *SUE* on *Net Buy Clustered*, *Net Buy Isolated*, and the same set of control variables used in Eq. 3.2. It should be noted that institutional trading is measured before the release of quarterly earnings, which typically occurs 30-40 days after a quarter-end.

The results, reported in Table 3.11, show that the coefficient for *Net Buy Isolated* is positive and highly statistically significant, whereas that for *Net Buy Clustered* is negative and statistically indistinguishable from zero. In terms of economic magnitudes, a one-standard-deviation increase in *Net Buy Isolated* predicts an increase of 0.5% in standardized unexpected earnings. The result suggests that, compared to geographically clustered institutions, geographically isolated institutions possess superior information about IPO firms' fundamentals. This finding is consistent with hypothesis **H9A**, instead of hypothesis **H9B**, that geographically isolated institutions, collectively, have more accurate information regarding the IPO firm's fundamentals.

3.7 Conclusion

In this paper, we analyze how the geographical locations of institutions affect their investments in IPOs and various characteristics of the IPOs that they invest in. We argue that institutions geographically close to each other may influence each other's investment decisions in IPOs. Further, they may also free-ride on each other's information when evaluating IPOs, resulting in IPOs dominated by geographically clustered institutions reflecting less accurate information signals compared to those dominated by geographically dispersed institutions.

We test the implications of the above hypotheses using a measure of institutions' geographical dispersion. Our empirical results can be summarized as follows. First, the equity holdings of institutions in IPOs are influenced more by the investments made by neighboring institutions than by those of distant institutions. Second, an increase in the geographical dispersion of the institutions investing in an IPO is associated with higher IPO price revisions, higher IPO and immediate secondary market firm valuations, larger IPO initial returns, and greater long-run post-IPO stock returns. Further, consistent with an information channel driving the above results, we find that the extent of information asymmetry facing an IPO firm is decreasing in the geographical dispersion of institutions investing in its IPO. Finally, the predictive power of institutional trading post-IPO for subsequent long-run stock returns and earnings surprises for the first fiscal-year end after the IPO is greater for geographically isolated institutions compared to those that are geographically clustered.

Table 3.1: **Summary Statistics for the IPO Sample**

This table reports the summary statitics of the characteristics of the IPO sample. The sample consists of IPOs conducted in 1980-2012. *Initial Return* is the percentage difference between the first trading day closing price and IPO offer price. *Revision* is the percentage difference between the IPO offer price and the midpoint of original filing range. *1YearHPR* is the IPO firms' one-year holding period return calculated by compounding daily returns over 252 trading days after the IPO (excluding the first trading day's return). *QOPAdj*, *QFTDAdj*, and *QFQAdj* are the industry-adjusted Tobin's Q ratio calculated using the IPO offer price, the first trading day closing price, and the price at the end of the first post-IPO fiscal quarter, respectively. Tobin's Q is the ratio of the market value of assets to the book value of assets, where the market value of assets is equal to the book value of assets minus the book value of common equity plus the number of shares outstanding times the share price. *Firm age* is the number of years from IPO firm founding year to the IPO issue year. *Total proceeds* is the total proceeds raised in the IPO in millions (USD).

	N	Mean	Std.	Median	10th	90th
Initial Return	5590	0.196	0.424	0.075	-0.017	0.477
Price Revision	5156	0.011	0.250	0.000	-0.250	0.250
1YrHPR	5590	0.110	0.858	-0.043	-0.675	0.959
Firm age	5512	17	22	8	2	45
Total proceeds ($ mil)	5590	96.8	396.8	40.0	11.7	168.2
Total assets before IPO ($ mil)	4303	749.2	7878.3	46.5	7.7	675.1
QOPAdj	5558	0.96	2.75	0.46	-0.39	2.47
QFTDAdj	5558	1.65	4.20	0.65	-0.29	3.88
QFQAdj	5558	2.04	6.03	0.65	-0.35	4.68

Table 3.2: **Summary Statistics of the Institutions' Investments in IPOs**
This table reports the summary statitics of the institutional shareholders characteristics of the IPO firms. The sample consists of IPOs conducted in 1980-2012. *Number of institutions* is the number of institutional shareholders in the first fiscal quarter after IPO reported in 13F filings. *Inst. geographic proximity* is the weighted-average geographic proximity among institutional shareholders in the first fiscal quarter after IPO reported in 13F filings. *Inst. ownership* is the fraction of shares outstanding held by institutional investors. *Inst. ownership concentration* is calculated as a Herfindahl Index of institutional ownership concentration based on the percentages of institutional holdings by all 13F institutions. For each IPO, we rank its institutional shareholders using their average geographical distance to all other shareholders holding the same stock. The institutional shareholders in the lower half rank are classified as clustered institutions, while the institutional shareholders in the upper half rank are classified as isolated institutions. We aggregate the net buying (change of holdings as a percentage of share outstanding) by clustered institutions and denote it as *Net Buy Clustered*; we aggregate the net buying by isolated institutions and denote it as *Net Buy Isolated*.

	N	Mean	Std.	Median	10th	90th
Number of institutions	5590	26	24	20	5	53
Geographical dispersion	5590	0.700	0.168	0.749	0.480	0.852
Inst. ownership	5590	0.230	0.195	0.181	0.053	0.460
Inst. ownership concentration	5590	0.188	0.152	0.138	0.064	0.389
Net Buy Isolated	5590	-0.011	0.041	-0.006	-0.047	0.020
Net Buy Clustered	5590	-0.013	0.041	-0.008	-0.048	0.016

Table 3.3: **The Effect of Neighboring Institutions on INstitutions' Investment in IPOs**
This table presents an institution-IPO level regressional analysis of the effect of neighbors on the participation of the institutions in the IPO. The dependent variable is the institution's holding as a percentange of total share outstanding reported in 13F filings the first calendar quarter after the IPO. *Neighbor Holdings* is the total holdings of neighboring institutions as a percentange of total share outstanding. *Non-neighbor Holdings* is the total holdings of non-neighboring institutions as a percentange of total share outstanding. In columns (1) and (4), two institutional investors are defined as *neighbor* if they are headquartered within 50 miles of each other. In columns (2) and (3), we vary the definition of *neighbor* using 100 miles and 200 miles as the cut-off boundary. *Local* is a dummy variable, which equals 1 if the particular institution is headquartered within 50 miles of IPO firm. *Log(TNA)* is the natural logarithm of total net assets of the institution. All standard errors are clustered at industry level. t-statistics are in parentheses. ***, **, and * indicate significance at the 1, 5, and 10 percent levels, respectively. We also report the p-value of the Wald-test for the null hypothesis that the coefficient for *Neighbor Holdings* equals to *Non-neighbor Holdings*.

	(1) Less than 50 miles	(2) Less than 100 miles	(3) Less than 200 miles	(4) Excluding Institutions from NYC or Boston
Neighbor Holdings	0.0117***	0.0104***	0.0116***	0.0309***
	(9.43)	(8.80)	(13.59)	(5.84)
Non-neighbor Holdings	0.0080***	0.0081***	0.0072***	0.0118***
	(10.32)	(10.14)	(8.29)	(6.97)
Local	0.0003	0.0004**	0.0005***	0.0006***
	(1.05)	(2.21)	(3.21)	(3.02)
Log(TNA)	0.0015***	0.0015***	0.0015***	0.0015***
	(12.27)	(12.31)	(12.29)	(13.49)
Year Fixed Effects	Yes	Yes	Yes	Yes
Industry Fixed Effects	Yes	Yes	Yes	Yes
Institution Fixed Effects	Yes	Yes	Yes	Yes
N	138,424	138,424	138,424	83,139

Adjusted R^2	0.198	0.198	0.198	0.219
Wald-test p-value	0.001	0.049	0.000	0.001

Table 3.4: **Geographical Dispersion Among Institutions and Price Revision**
This table presents a regressional analysis of IPO offer price revision on the geographical dispersion of institutional shareholders. *Revision* is the percentage difference between the IPO offer price and the midpoint of original filing range. *Inst. geographic dispersion* is the weighted-average geographic dispersion among institutional shareholders in the first fiscal quarter after IPO reported in 13F filings. *Inst. ownership* is the fraction of shares outstanding held by institutional investors. *Inst. ownership concentration* is calculated as a Herfindahl Index of institutional ownership concentration based on the percentages of institutional holdings by all 13F institutions. *Log(Reputation)* is the natural logarithm of the lead underwriter reputation ranking. The rankings are obtained from Jay Ritter's website.The maximum ranking is used when there are multiple lead underwriters. *Log(Age+1)* is the natural logarithm of the IPO firm age plus one, where age is IPO year minus company founding year. *Log(Proceeds)* is the natural logarithm of the IPO offering proceeds. *Prior market return* is the return on the CRSP value-weighted index over the 1 month period prior to the IPO. *HighTech* equals one if the IPO firm is in high-tech industries, and zero otherwise. *Financial* equals one if the IPO firm is in the financials industry, and zero otherwise. *VC backed* equals one if the IPO has venture capital backing, and zero otherwise. *Lockup* equals one if the IPO has a lockup provision, and zero otherwise. All standard errors are clustered at industry level. t-statistics are in parentheses. ***, **, and * indicate significance at the 1, 5, and 10 percent levels, respectively.

	(1) Price Revision	(2) Price Revision
Geographical dispersion	0.111***	0.137***
	(4.03)	(4.07)
Inst. ownership	-0.111***	-0.134***
	(3.51)	(3.97)
Inst. ownership concentration	-0.047	-0.056*
	(1.29)	(1.71)
Log(Reputation)	-0.011**	-0.009*
	(2.18)	(1.87)
Log(Proceeds)	0.073***	0.050***
	(4.30)	(4.32)
Log(Age+1)	-0.022***	-0.018***
	(4.41)	(3.49)
Prior market return	0.991***	1.071***
	(6.59)	(7.77)
Lockup	-0.019	-0.051***

	(1.31)	(4.32)
VC backed	0.031**	0.021
	(2.16)	(0.93)
HighTech		0.074***
		(3.32)
Financial		0.001
		(0.07)
Bubble		0.079***
		(11.28)
Year Fixed Effects	Yes	No
Industry Fixed Effects	Yes	No
Observations	5096	5096
Adjusted R^2	0.32	0.16

Table 3.5: **Geographical Dispersion Among Institutions and IPO Valuation**
This table presents a regressional analysis of IPO valuation on the geographical dispersion of institutional shareholders. *QOPAdj* is the industry-adjusted Tobin's Q ratio calculated using the IPO offer price. *Inst. geographic dispersion* is the weighted-average geographic dispersion among institutional shareholders in the first fiscal quarter after IPO reported in 13F filings. *Inst. ownership* is the fraction of shares outstanding held by institutional investors. *Inst. ownership concentration* is calculated as a Herfindahl Index of institutional ownership concentration based on the percentages of institutional holdings by all 13F institutions. *Log(Reputation)* is the natural logarithm of the lead underwriter reputation ranking. The rankings are obtained from Jay Ritter's website.The maximum ranking is used when there are multiple lead underwriters. *Log(Age+1)* is the natural logarithm of the IPO firm age plus one, where age is IPO year minus company founding year. *Log(Proceeds)* is the natural logarithm of the IPO offering proceeds. *Prior market return* is the return on the CRSP value-weighted index over the 1 month period prior to the IPO. *HighTech* equals one if the IPO firm is in high-tech industries, and zero otherwise. *Financial* equals one if the IPO firm is in the financials industry, and zero otherwise. *VC backed* equals one if the IPO has venture capital backing, and zero otherwise. *Lockup* equals one if the IPO has a lockup provision, and zero otherwise. All standard errors are clustered at industry level. t-statistics are in parentheses. ***, **, and * indicate significance at the 1, 5, and 10 percent levels, respectively.

	(1) QOPAdj	(2) QOPAdj
Geographical dispersion	1.270***	1.257***
	(6.28)	(5.74)
Inst. ownership	-2.933***	-2.846***
	(9.82)	(9.01)
Inst. ownership concentration	1.053***	1.204***
	(5.42)	(5.39)
Log(Reputation)	0.069*	0.066**
	(1.81)	(2.33)
Log(Proceeds)	0.207***	0.278***
	(3.90)	(4.63)
Log(Age+1)	-0.274***	-0.227***
	(5.82)	(4.76)
Prior market return	-2.063***	-2.214***
	(3.33)	(3.07)
Lockup	-0.213	-0.097

	(1.65)	(1.11)
VC backed	0.328**	0.321**
	(2.61)	(2.44)
HighTech		0.405***
		(4.02)
Financial		-0.220*
		(1.88)
Bubble		0.250***
		(2.79)
Year Fixed Effects	Yes	No
Industry Fixed Effects	Yes	No
Observations	5481	5481
Adjusted R^2	0.07	0.07

Table 3.6: **Geographical Dispersion Among Institutions and Secondary Market Valuation**
This table presents a regressional analysis of IPO valuation at offering and secondary market on the geographical dispersion of institutional shareholders. *QFTDAdj* and *QFQAdj* are the industry-adjusted Tobin's Q ratio calculated using the first trading day closing price, and the price at the end of the first post-IPO fiscal quarter, respectively. *Inst. geographic dispersion* is the weighted-average geographic dispersion among institutional shareholders in the first fiscal quarter after IPO reported in 13F filings. *Inst. ownership* is the fraction of shares outstanding held by institutional investors. *Inst. ownership concentration* is calculated as a Herfindahl Index of institutional ownership concentration based on the percentages of institutional holdings by all 13F institutions. *Log(Reputation)* is the natural logarithm of the lead underwriter reputation ranking. The rankings are obtained from Jay Ritter's website.The maximum ranking is used when there are multiple lead underwriters. *Log(Age+1)* is the natural logarithm of the IPO firm age plus one, where age is IPO year minus company founding year. *Log(Proceeds)* is the natural logarithm of the IPO offering proceeds. *High-Tech* equals one if the IPO firm is in high-tech industries, and zero otherwise. *Financial* equals one if the IPO firm is in the financials industry, and zero otherwise. *VC backed* equals one if the IPO has venture capital backing, and zero otherwise. *Lockup* equals one if the IPO has a lockup provision, and zero otherwise. All standard errors are clustered at industry level. t-statistics are in parentheses. ***, **, and * indicate significance at the 1, 5, and 10 percent levels, respectively.

	(1) QFTDAdj	(2) QFTDAdj	(3) QFQAdj	(4) QFQAdj
Geographical dispersion	2.146***	2.371***	2.081***	2.509***
	(6.96)	(6.62)	(5.93)	(5.18)
Inst. ownership	-4.154***	-4.085***	-4.113***	-4.062***
	(7.56)	(6.93)	(6.26)	(5.78)
Inst. ownership concentration	1.798***	1.968***	0.877***	1.218***
	(3.64)	(4.12)	(2.99)	(3.43)
Log(Reputation)	0.091*	0.096**	0.150**	0.144**
	(1.73)	(2.37)	(2.05)	(2.32)
Log(Proceeds)	0.473**	0.554***	0.256**	0.478***
	(2.61)	(3.34)	(2.24)	(3.07)
Log(Age+1)	-0.497***	-0.423***	-0.579***	-0.546***
	(4.69)	(3.93)	(3.65)	(3.00)
Lockup	-0.825***	-0.428**	-1.901**	-0.864**
	(2.99)	(2.46)	(2.61)	(2.11)
VC backed	0.771***	0.690***	0.835**	0.708**

	(3.14)	(2.66)	(2.51)	(2.42)
HighTech		0.878***		1.346***
		(4.89)		(5.72)
Financial		-0.364**		-0.404**
		(2.35)		(2.42)
Bubble		2.442***		3.257***
		(9.55)		(6.25)
Year Fixed Effects	Yes	No	Yes	No
Industry Fixed Effects	Yes	No	Yes	No
Observations	5481	5481	5481	5481
Adjusted R^2	0.16	0.16	0.14	0.12

Table 3.7: **Geographical Dispersion Among Institutions and IPO Initial Return**
This table presents a regressional analysis of IPO initial return on the geographical dispersion of institutional shareholders. *Initial Return* is the percentage difference between the first trading day closing price and IPO offer price. *Inst. geographic dispersion* is the weighted-average geographic dispersion among institutional shareholders in the first fiscal quarter after IPO reported in 13F filings. *Inst. ownership* is the fraction of shares outstanding held by institutional investors. *Inst. ownership concentration* is calculated as a Herfindahl Index of institutional ownership concentration based on the percentages of institutional holdings by all 13F institutions. *Firm-institution distance* is the natural logarithm of weighted-average geographical distance between the institutional shareholders and the IPO firm. *Log(Reputation)* is the natural logarithm of the lead underwriter reputation ranking. The rankings are obtained from Jay Ritter's website.The maximum ranking is used when there are multiple lead underwriters. *Log(Age+1)* is the natural logarithm of the IPO firm age plus one, where age is IPO year minus company founding year. *Log(Proceeds)* is the natural logarithm of the IPO offering proceeds. *Prior market return* is the return on the CRSP value-weighted index over the 1 month period prior to the IPO. *High-Tech* equals one if the IPO firm is in high-tech industries, and zero otherwise. *Financial* equals one if the IPO firm is in the financials industry, and zero otherwise. *VC backed* equals one if the IPO has venture capital backing, and zero otherwise. *Lockup* equals one if the IPO has a lockup provision, and zero otherwise. All standard errors are clustered at industry level. t-statistics are in parentheses. ***, **, and * indicate significance at the 1, 5, and 10 percent levels, respectively.

	(1) Initial Return	(2) Initial Return
Geographical dispersion	0.106**	0.165***
	(2.50)	(3.53)
Inst. ownership	-0.086***	-0.092***
	(4.18)	(3.72)
Inst. ownership concentration	-0.041	-0.030
	(0.95)	(0.79)
Log(Reputation)	-0.002	0.002
	(0.42)	(0.46)
Log(Proceeds)	0.032*	0.034**
	(1.86)	(2.39)
Log(Age+1)	-0.041***	-0.038***
	(3.91)	(3.16)
Prior market return	1.464***	1.458***

	(3.33)	(3.40)
Lockup	-0.098***	-0.046*
	(3.12)	(1.99)
VC backed	0.079***	0.059***
	(4.98)	(2.87)
HighTech		0.099***
		(3.54)
Financial		-0.009
		(0.43)
Bubble		0.447***
		(9.33)
Year Fixed Effects	Yes	No
Industry Fixed Effects	Yes	No
Observations	5509	5509
Adjusted R^2	0.25	0.24

Table 3.8: **Geographical Dispersion Among Institutions and Long-run Post-IPO Stock Performance**

This table presents a regressional analysis of post-IPO stock return performance on the geographical dispersion of institutional shareholders. *1YrHPRAdjFF25* is the IPO firms' one-year holding period return calculated by compounding daily returns over 252 trading days after the IPO (excluding the first trading day's return) adjusted for (minus) the matched Fama/French 25 size and book-to-market portfolio buy-and-hold value-weighted return. *1YrHPRAdjMM* is the IPO firms' one-year holding period return calculated by compounding daily returns over 252 trading days after the IPO (excluding the first trading day's return) adjusted for (minus) the CRSP value-weighted market return. *Inst. geographic dispersion* is the weighted-average geographic dispersion among institutional shareholders in the first fiscal quarter after IPO reported in 13F filings. *Inst. ownership* is the fraction of shares outstanding held by institutional investors. *Inst. ownership concentration* is calculated as a Herfindahl Index of institutional ownership concentration based on the percentages of institutional holdings by all 13F institutions. *Log(Reputation)* is the natural logarithm of the lead underwriter reputation ranking. The rankings are obtained from Jay Ritter's website.The maximum ranking is used when there are multiple lead underwriters. *Log(Age+1)* is the natural logarithm of the IPO firm age plus one, where age is IPO year minus company founding year. *Log(Proceeds)* is the natural logarithm of the IPO offering proceeds. *VC backed* equals one if the IPO has venture capital backing, and zero otherwise. *Lockup* equals one if the IPO has a lockup provision, and zero otherwise. All standard errors are clustered at industry level. t-statistics are in parentheses. ***, **, and * indicate significance at the 1, 5, and 10 percent levels, respectively.

	(1) 1YrHPRAdjFF25	(2) 1YrHPRAdjFF25	(3) 1YrHPRAdjMM	(4) 1YrHPRAdjMM
Geographical dispersion	0.292***	0.171**	0.265***	0.137*
	(3.57)	(2.28)	(3.27)	(1.89)
Inst. ownership	0.016	0.026	0.060*	0.072*
	(0.40)	(0.64)	(1.67)	(1.90)
Inst. ownership concentration		-0.204**		-0.217**
		(2.39)		(2.42)
Log(Reputation)	0.006	0.004	0.007	0.005
	(0.47)	(0.29)	(0.58)	(0.39)
Log(Proceeds)	-0.047***	-0.053***	-0.036**	-0.042**
	(3.08)	(3.36)	(2.33)	(2.62)
Log(Age+1)	0.018	0.018	0.021*	0.021*
	(1.58)	(1.57)	(1.81)	(1.79)
Initial Return	-0.061*	-0.061*	-0.076**	-0.077**

	(1.92)	(1.90)	(2.21)	(2.18)
Lockup	-0.112**	-0.114**	-0.134**	-0.135**
	(2.10)	(2.11)	(2.17)	(2.18)
VC backed	-0.026	-0.027	-0.030	-0.032
	(0.52)	(0.55)	(0.62)	(0.65)
Year Fixed Effects	Yes	Yes	Yes	Yes
Industry Fixed Effects	Yes	Yes	Yes	Yes
Observations	5509	5509	5509	5509
Adjusted R^2	0.03	0.03	0.03	0.03

Table 3.9: **Geographical Dispersion Among Institutions and Information Asymmetry**

This table presents a regressional analysis of secondary-market information asymmetry measures on the geographical dispersion of institutional shareholders. For each IPO sample, we retrieve analyst earnings forecasts from I/B/E/S for the first fiscal year post-IPO. *Log(MSE)* is the natural logarithm of mean-squared error in the earnings forecast. We measure forecast errors as the absolute difference between the average forecasted earnings and the actual earnings per share divided by the price per share at the time of the forecast. *Log(Dispersion)* is the natural logarithm of standard deviation of analyst forecasts. *Log(# of Analysts)* is the natural logarithm of the number of analysts following the firm *Log(COV)* is the natural logarithm of the coefficient of variation (COV) of analyst forecasts. COV is defined as the ratio of standard deviation to the absolute value of the average of analyst forecasts. *Inst. geographic dispersion* is the weighted-average geographic dispersion among institutional shareholders in the first fiscal quarter after IPO reported in 13F filings. *Inst. ownership* is the fraction of shares outstanding held by institutional investors. *Inst. ownership concentration* is calculated as a Herfindahl Index of institutional ownership concentration based on the percentages of institutional holdings by all 13F institutions. *Log(Reputation)* is the natural logarithm of the lead underwriter reputation ranking. The rankings are obtained from Jay Ritter's website.The maximum ranking is used when there are multiple lead underwriters. *Log(Age+1)* is the natural logarithm of the IPO firm age plus one, where age is IPO year minus company founding year. *Log(Proceeds)* is the natural logarithm of the IPO offering proceeds. *Prior market return* is the return on the CRSP value-weighted index over the 1 month period prior to the IPO. *High-Tech* equals one if the IPO firm is in high-tech industries, and zero otherwise. *Financial* equals one if the IPO firm is in the financials industry, and zero otherwise. *VC backed* equals one if the IPO has venture capital backing, and zero otherwise. *Lockup* equals one if the IPO has a lockup provision, and zero otherwise. All standard erros are clustered at industry level. t-statistics are in parentheses. ***, **, and * indicate significance at the 1, 5, and 10 percent levels, respectively.

	(1) Log(MSE)	(2) Log(MSE)	(3) Log(Dispersion)	(4) Log(Dispersion)	(5) Log(COV)	(6) Log(COV)	(7) Log(# of Analysts)	(8) Log(# of Analysts)
Geographical dispersion	-2.016***	-2.952***	-0.672**	-1.192***	-0.307	-0.664**	0.048	-0.003
	(3.72)	(7.14)	(2.54)	(4.41)	(1.11)	(2.17)	(0.40)	(0.02)
Inst. ownership	0.228	0.471	0.137	0.285	0.045	0.134	-0.035	-0.029
	(0.72)	(1.43)	(0.83)	(1.43)	(0.23)	(0.69)	(0.49)	(0.40)
Inst. ownership concentration	1.726***	2.296***	0.635**	0.803***	0.615**	0.752***	0.033	0.027
	(3.44)	(5.12)	(2.53)	(3.21)	(2.18)	(2.82)	(0.31)	(0.25)
Log(Reputation)	0.094	-0.083	0.061	-0.010	0.104	-0.001	0.015	0.008
	(0.74)	(0.72)	(1.07)	(0.26)	(1.48)	(0.02)	(0.77)	(0.42)
Log(Proceeds)	-0.250**	0.205***	0.126***	0.295***	0.018	0.168***	0.287***	0.296***

	(2.33)	(3.28)	(3.26)	(10.07)	(0.39)	(5.12)	(17.69)	(19.99)
Log(Age+1)	-0.326***	-0.404***	-0.130***	-0.197***	-0.092***	-0.158***	-0.009	-0.005
	(5.11)	(4.75)	(2.99)	(3.59)	(3.09)	(4.31)	(0.54)	(0.29)
Lockup	0.410	0.550**	0.069	0.116	0.086	0.049	-0.129***	-0.124***
	(1.66)	(2.57)	(0.77)	(1.34)	(1.02)	(0.94)	(4.98)	(4.76)
VC backed	0.236	0.567	0.142*	0.319**	0.313***	0.393***	0.115***	0.109***
	(1.20)	(1.63)	(1.89)	(2.42)	(4.86)	(7.51)	(3.57)	(2.84)
HighTech		-1.077**		-0.392*		-0.038		0.088**
		(2.21)		(1.94)		(0.34)		(2.55)
Financial		-0.816***		-0.330**		-0.593***		0.102*
		(2.66)		(2.18)		(4.73)		(1.88)
Bubble		1.594***		0.376*		-0.284***		-0.133***
		(4.25)		(1.90)		(3.34)		(4.39)
Year Fixed Effects	Yes	No	Yes	No	Yes	No	Yes	No
Industry Fixed Effects	Yes	No	Yes	No	Yes	No	Yes	No
Observations	3724	3724	2314	2314	2311	2311	3898	3898
Adjusted R^2	0.13	0.07	0.13	0.07	0.12	0.06	0.25	0.22

Table 3.10: **Predictive Power of Institutional Trading for Subsequent Long-run IPO Abnormal Performance**

This table presents the regressional analysis whether trading by geographically isolated institutions have stronger predictive power for future IPO stock returns, compared to geographically clustered institutions. The dependent variable is market-adjusted or book-to-market and size adjusted holding period return during the 12 months immediately following the first two quarters post-IPO. For each IPO, we rank its institutional shareholders using their average geographical distance to all other shareholders holding the same stock. The institutional shareholders in the lower half rank are classified as clustered institutions, while the institutional shareholders in the upper half rank are classified as isolated institutions. We aggregate the net buying (change of holdings as a percentage of share outstanding) between the first and the second quarter post-IPO by clustered institutions and denote it as *Net Buy Clustered*; we aggregate the net buying by isolated institutions between the first and the second quarter post-IPO and denote it as *Net Buy Isolated*. *Inst. ownership* is the fraction of shares outstanding held by institutional investors. *Inst. ownership concentration* is calculated as a Herfindahl Index of institutional ownership concentration based on the percentages of institutional holdings by all 13F institutions. *Log(Reputation)* is the natural logarithm of the lead underwriter reputation ranking. The rankings are obtained from Jay Ritter's website.The maximum ranking is used when there are multiple lead underwriters. *Log(Age+1)* is the natural logarithm of the IPO firm age plus one, where age is IPO year minus company founding year. *Log(Proceeds)* is the natural logarithm of the IPO offering proceeds. *Prior market return* is the return on the CRSP value-weighted index over the 1 month period prior to the IPO. *High-Tech* equals one if the IPO firm is in high-tech industries, and zero otherwise. *Financial* equals one if the IPO firm is in the financials industry, and zero otherwise. *VC backed* equals one if the IPO has venture capital backing, and zero otherwise. *Lockup* equals one if the IPO has a lockup provision, and zero otherwise. All standard errors are clustered at industry level. t-statistics are in parentheses. ***, **, and * indicate significance at the 1, 5, and 10 percent levels, respectively.

	(1) HPRAdjFF25	(2) HPRAdjFF25	(3) HPRAdjMM	(4) HPRAdjMM
Net Buy Isolated	0.721**	0.751**	0.723**	0.766**
	(2.27)	(2.47)	(2.42)	(2.55)
Net Buy Clustered	0.062	0.019	0.088	0.018
	(0.15)	(0.04)	(0.20)	(0.04)
Inst. ownership	-0.007	-0.002	0.023	0.049
	(0.14)	(0.04)	(0.43)	(0.85)
Inst. ownership concentration	-0.059	-0.048	-0.047	0.004
	(0.95)	(0.80)	(0.72)	(0.07)
Log(Reputation)	-0.002	0.011*	0.001	0.010

	(0.21)	(1.77)	(0.13)	(1.67)
Log(Proceeds)	-0.009	-0.005	0.003	0.022*
	(0.62)	(0.47)	(0.19)	(1.84)
Log(Age+1)	0.026***	0.020*	0.031***	0.028**
	(2.72)	(1.87)	(3.14)	(2.58)
Initial Return	-0.077***	-0.072**	-0.105***	-0.098***
	(2.84)	(2.63)	(3.10)	(2.89)
Lockup	-0.076***	-0.052**	-0.089***	-0.033*
	(2.99)	(2.52)	(3.25)	(1.67)
VC backed	-0.014	-0.002	-0.014	0.002
	(0.45)	(0.06)	(0.46)	(0.07)
HighTech		0.047		0.023
		(1.66)		(0.79)
Financial		0.020		0.058**
		(0.51)		(2.01)
Bubble		-0.332***		-0.188**
		(4.29)		(2.01)
Year Fixed Effects	Yes	No	Yes	No
Industry Fixed Effects	Yes	No	Yes	No
Observations	5324	5324	5324	5324
Adjusted R^2	0.03	0.03	0.04	0.02

Table 3.11: **Predictive Power of Institutional Trading for Subsequent Earnings Surprises**
This table presents the regressional analysis whether trading by geographically isolated institutions have stronger predictive power for firms' earnings surprises, compared to geographically clustered institutions. The dependent variable is standardized unexpected earnings (SUE), defined as I/B/E/S analyst forecast error scaled by stock price. For each IPO, we rank its institutional shareholders using their average geographical distance to all other shareholders holding the same stock. The institutional shareholders in the lower half rank are classified as clustered institutions, while the institutional shareholders in the upper half rank are classified as isolated institutions. We aggregate the net buying (change of holdings as a percentage of shares outstanding) by clustered institutions and denote it as *Net Buy Clustered*; we aggregate the net buying by isolated institutions and denote it as *Net Buy Isolated*. *Inst. ownership* is the fraction of shares outstanding held by institutional investors. *Inst. ownership concentration* is calculated as the Herfindahl Index of institutional ownership concentration based on the percentages of institutional holdings by all 13F institutions. *Log(Reputation)* is the natural logarithm of the lead underwriter reputation ranking. The rankings are obtained from Jay Ritter's website.The maximum ranking is used when there are multiple lead underwriters. *Log(Age+1)* is the natural logarithm of the IPO firm age plus one, where age is IPO year minus company founding year. *Log(Proceeds)* is the natural logarithm of the IPO offering proceeds. *Prior market return* is the return on the CRSP value-weighted index over the 1 month period prior to the IPO. *High-Tech* equals one if the IPO firm is in high-tech industries, and zero otherwise. *Financial* equals one if the IPO firm is in the financials industry, and zero otherwise. *VC backed* equals one if the IPO has venture capital backing, and zero otherwise. *Lockup* equals one if the IPO has a lockup provision, and zero otherwise. All standard errors are clustered at industry level. t-statistics are in parentheses. ***, **, and * indicate significance at the 1, 5, and 10 percent levels, respectively.

	(1) SUE	(2) SUE
Net Buy Isolated	0.134***	0.123***
	(4.52)	(4.28)
Net Buy Clustered	-0.005	-0.004
	(0.21)	(0.17)
Inst. ownership	0.010	0.011
	(1.22)	(1.41)
Inst. ownership concentration	-0.026**	-0.025**
	(2.20)	(2.17)
Log(Reputation)	0.002	0.002
	(0.93)	(1.17)
Log(Proceeds)	0.004***	0.002

	(3.16)	(1.45)
Log(Age+1)	0.002	0.003^{**}
	(1.39)	(2.07)
Initial Return	0.010^{***}	0.010^{***}
	(4.50)	(4.72)
Lockup	-0.001	-0.002
	(0.58)	(1.11)
VC backed	-0.001	-0.001
	(0.41)	(0.40)
HighTech		0.005^{**}
		(2.26)
Financial		0.008^{*}
		(1.94)
Bubble		-0.009^{**}
		(2.06)
Year Fixed Effects	Yes	No
Industry Fixed Effects	Yes	No
Observations	2986	2986
Adjusted R^2	0.01	0.02

Bibliography

Aggarwal, Reena, 2003, Allocation of initial public offerings and flipping activity, *Journal of Financial Economics* 68, 111–135.

Aggarwal, Rajesh K., Laurie Krigman, and Kent L. Womack, 2002, Strategic IPO underpricing, information momentum, and lockup expiration selling, *Journal of Financial Economics* 66, 105–137.

Allen, Franklin, Antonio E. Bernardo, and Ivo Welch, 2000, A theory of dividends based on tax clienteles, *The Journal of Finance* 55, 2499–2536.

Allen, Franklin, and Roni Michaely, 2003, Chapter 7 - payout policy, *Handbook of the Economics of Finance*.

Babenko, Ilona, Yuri Tserlukevich, and Alexander Vedrashko, 2012, The credibility of open market share repurchase signaling, *Journal of Financial and Quantitative Analysis* 47 (5), 1059–1088.

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.

Banyi, Monica L., Edward A. Dyl, and Kathleen M. Kahle, 2008, Errors in estimating share repurchases, *Journal of Corporate Finance* 14 (4), 460–474.

Bayar, Onur, Thomas J. Chemmanur, and Mark H. Liu, 2014, Dividends versus stock repurchases and long-run stock returns under heterogeneous beliefs, *Working Paper*.

Ben-Rephael, Azi, Jacob Oded, and Avi Wohl, 2014, Do Firms Buy Their Stock at Bargain Prices? Evidence from Actual Stock Repurchase Disclosures, *Review of Finance* 18, 1299–1340.

Benveniste, Lawrence M., and Paul A. Spindt, 1989, How investment bankers determine the offer price and allocation of new issues, *Journal of Financial Economics* 24, 343–361.

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

Boehmer, Beatrice, Ekkehart Boehmer, and Raymond P. H. Fishe, 2006, Do Institutions Receive Favorable Allocations in IPOs with Better Long-Run Returns?, *Journal of Financial and Quantitative Analysis* 41, 809–828.

Brennan, Michael J., and Anjan V. Thakor, 1990, Shareholder preferences and dividend policy, *Journal of Finance* 45 (4), 993–1018.

Brunnermeier, Markus, and Stefan Nagel, 2004, Hedge Funds and the Technology Bubble, *The Journal of Finance* 59, 2013–2040.

Busch, Pascal, and Stefan Obernberger, 2017, Actual Share Repurchases, Price Efficiency, and the Information Content of Stock Prices, *The Review of Financial Studies* 30, 324–362.

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.

Carter, Richard, and Steven Manaster, 1990, Initial Public Offerings and Underwriter Reputation, *The Journal of Finance* 45, 1045–1067.

Chemmanur, Thomas J., 1993, The Pricing of Initial Public Offerings: A Dynamic Model with Information Production, *The Journal of Finance* 48, 285–304.

——— , Shan He, and Gang Hu, 2009, The role of institutional investors in seasoned equity offerings, *Journal of Financial Economics* 94 (3), 384–411.

Chemmanur, Thomas J., Gang Hu, and Jiekun Huang, 2010, The Role of Institutional Investors in Initial Public Offerings, *Review of Financial Studies* 23, 4496–4540.

Chemmanur, Thomas J., and Yawen Jiao, 2011, Institutional trading, information production, and the SEO discount: A model of seasoned equity offerings, *Journal of Economics and Management Strategy* 20 (1), 299–338.

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.

Comment, Robert, and Gregg A. Jarrell, 1991, The relative signalling power of dutch-auction and fixed-price self-tender offers and open-market share repurchases, *Journal of Finance* 46 (4), 1243–1271.

Conrad, Jennifer S., Kevin M. Johnson, and Sunil Wahal, 2001, Institutional trading and soft dollars, *Journal of Finance* 56 (1), 397–416.

Constantinides, George M., and Bruce D. Grundy, 1989, Optimal investment with stock repurchase and financing as signals, *Review of Financial Studies* 2 (4), 445–465.

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.

Dann, Larry Y., 1981, Common stock repurchases: An analysis of returns to bondholders and stockholders, *Journal of Financial Economics* 9 (2), 113–138.

DeLisle, R. Jared, J.D. Morscheck, and John R. Nofsinger, 2014, Share repurchases and institutional supply, *Journal of Corporate Finance* 27, 216 – 230.

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.

Ellis, Katrina, 2006, Who trades IPOs? A close look at the first days of trading, *Journal of Financial Economics* 79, 339–363.

———, Roni Michaely, and Maureen O'Hara, 2000, When the Underwriter Is the Market Maker: An Examination of Trading in the IPO Aftermarket, *The Journal of Finance* 55, 1039–1074.

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, 2001, Disappearing dividends: changing firm characteristics or lower propensity to pay?, *Journal of Financial Economics* 60 (1), 3–43.

———, 2010, Luck versus Skill in the Cross-Section of Mutual Fund Returns, *The Journal of Finance* 65, 1915–1947.

Field, Laura Casares, and Michelle Lowry, 2009, Institutional versus Individual Investment in IPOs: The Importance of Firm Fundamentals, *Journal of Financial and Quantitative Analysis* 44, 489–516.

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.

Gibson, Scott, Assem Safieddine, and Ramana Sonti, 2004, Smart investments by smart money: Evidence from seasoned equity offerings, *Journal of Financial Economics* 72 (3), 581–604.

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.

Griffin, John M., and Jin Xu, 2009, How Smart Are the Smart Guys? A Unique View from Hedge Fund Stock Holdings, *Review of Financial Studies* 22, 2531–2570.

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

Grullon, Gustavo, and Roni Michaely, 2002, Dividends, share repurchases, and the substitution hypothesis, *Journal of Finance* 57 (4), 1649–1684.

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

Hanley, Kathleen Weiss, and William J. Wilhelm Jr., 1995, Evidence on the strategic allocation of initial public offerings, *Journal of Financial Economics* 37, 239–257.

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.

Huang, Sheng, and Joe Zhang, 2013, How do institutional investors trade when firms are buying back shares?, *Working Paper*.

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.

Ikenberry, David, Josef Lakonishok, and Theo Vermaelen, 1995, Market underreaction to open market share repurchases, *Journal of Financial Economics* 39 (2-3), 181–208.

Ikenberry, David L., and Theo Vermaelen, 1996, The option to repurchase stock, *Financial Management* 25, 9–24.

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.

Jones, Charles M, and Marc L Lipson, 2001, Sixteenths: direct evidence on institutional execution costs, *Journal of Financial Economics* 59 (2), 253–278.

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

Keim, Donald B., and Ananth Madhavan, 1995, Anatomy of the trading process empirical evidence on the behavior of institutional traders, *Journal of Financial Economics* 37 (3), 371–398.

Kostovetsky, Leonard, and Jerold B. Warner, 2015, You're 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.

Kyle, Albert S., 1985, Continuous auctions and insider trading, *Econometrica* 53 (6), 1315–1335.

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

Leland, Hayne E., and David H. Pyle, 1977, Informational asymmetries, financial structure, and financial intermediation, *The Journal of Finance* 32, 371–387.

Livnat, Joshua, and Richard R. Mendenhall, 2006, Comparing the Post–Earnings Announcement Drift for Surprises Calculated from Analyst and Time Series Forecasts, *Journal of Accounting Research* 44, 177–205.

Loughran, Tim, and Jay Ritter, 2004, Why Has IPO Underpricing Changed over Time?, *Financial Management* 33, 5–37.

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

McNally, William J., 1999, Open market stock repurchase signaling, *Financial Management* 28, 55–67.

Nimalendran, M., Jay R. Ritter, and Donghang Zhang, 2007, Do today's trades affect tomorrow's IPO allocations?, *Journal of Financial Economics* 84, 87–109.

Oded, Jacob, 2005, Why do firms announce open-market repurchase programs?, *Review of Financial Studies* 18 (1), 271–300.

Ofer, Aharon R., and Anjan V. Thakor, 1987, A theory of stock price responses to alternative corporate cash disbursement methods: Stock repurchases and dividends, *Journal of Finance* 42 (2), 365–394.

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.

Persons, Jonh C., 1994, Signaling and takeover deterrence with stock repurchases: Dutch auctions versus fixed price tender offers, *The Journal of Finance* 49, 1373–1402.

Petajisto, Antti, 2013, Active Share and Mutual Fund Performance, *Financial Analysts Journal* 69, 73–93.

Peyer, Urs, and Theo Vermaelen, 2009, The nature and persistence of buyback anomalies, *Review of Financial Studies* 22 (4), 1693–1745.

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.

Ritter, Jay R., 1991, The Long-Run Performance of initial Public Offerings, *The Journal of Finance* 46, 3–27.

——— , and Ivo Welch, 2002, A Review of IPO Activity, Pricing, and Allocations, *The Journal of Finance* 57, 1795–1828.

Ritter, Jay R., and Donghang Zhang, 2007, Affiliated mutual funds and the allocation of initial public offerings, *Journal of Financial Economics* 86, 337–368.

Rock, Kevin, 1986, Why new issues are underpriced, *Journal of Financial Economics* 15, 187–212.

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.

Sherman, Ann E., and Sheridan Titman, 2002, Building the IPO order book: underpricing and participation limits with costly information, *Journal of Financial Economics* 65, 3–29.

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*.

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

Stephens, Clifford P., and Michael S. Weisbach, 1998, Actual share reacquisitions in open-market repurchase programs, *Journal of Finance* 53 (1), 313–333.

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.

Vermaelen, Theo, 1981, Common stock repurchases and market signalling: An empirical study, *Journal of Financial Economics* 9 (2), 139–183.

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.