Common Ownership Network and Mutual Fund Performance.

Shema F. Mitali*

January 2017

ABSTRACT

Using quarterly ownership data, I construct a network of actively managed U.S. domestic equity mutual funds. Cross-fund portfolio similarities are measured as the importance of each fund within the network. Empirically, I show that deviations from a common portfolio is associated with higher performance, conditional on fund characteristics. Theoretically, the trade-off between portfolio similarity and performance is consistent with a simple decision-making problem with negative network externalities.

JEL classification: G11, G23.

Keywords: Portfolio Similarity, Financial Networks, Mutual Funds.

*Warwick Business School, University of Warwick, Coventry, UK. `Shema.Mitali.15@mail.wbs.ac.uk`. I thank Constantinos Antoniou and Daniele Bianchi for very useful comments. I also thank Stanislao Gualdi, and Mikhail Simutin.

How similar is your fund manager compared to his benchmark? Previous work have derived several measures capturing the deviation from a benchmark and have shown that activeness positively affects performance. For instance, "tracking error" is a common measure of active management while "Active Share" (Cremers and Petajisto 2009) also measures activeness by addressing the drawbacks of the former. To further investigate this issue, the natural extension would be to examine how similar is your fund manager compared to other active managers. As in prior literature, a holdings-based measure relative to the other active funds could help answering that question. I argue that portfolio similarity is a key component driving performance among active mutual funds.

I propose a simple measure of portfolio similarity based on common ownership network. Using holdings data from 1980Q1 to 2014Q4, I connect actively managed U.S. domestic equity mutual funds through portfolio overlap, i.e. the number of positions they have in common. I derive a portfolio similarity measure as the importance of each fund within the network (its "centrality") at each quarter. Hence, a "central" fund has highly similar holdings with its peers, namely it holds a common portfolio.

To capture portfolio similarities at the fund level, I compute two simple metrics from the common ownership network: non-weighted degree and degree centrality. The first measure simply counts the number of funds that share similar positions, while the second centrality measure takes into account the weight (portfolio overlap) associated to the links. For example, if we have a network composed of three funds (A, B, and C) where the first two (fund A and B) have four positions in common while the third fund (C) does not share any position with the others; the non-weighted degree for fund A and B will be 1 since fund A shares similar positions only with fund B and vice versa. The non-weighted degree for fund C will be 0. However, the second measure (*weighted* degree centrality) will be 4 for the fund A and B since this is their overlap, i.e. the weight associated to their link.

In the main results, I find that an increase of one-standard-deviation in portfolio similarity (degree centrality) negatively impacts a fund's quarterly net return by 12 basis points ($t =$

11.75) in a regression approach and controlling for fund characteristics.

This measure provides a new dimension in assessing the originality of a fund's manager. Along with "Active Share" and other measures of activeness, portfolio similarity can tell if a fund is likely to beat not only the benchmark but also its peers. Indeed, a fund holding a common portfolio is less likely to outperform other funds having similar positions, unless relying on market timing strategies. In terms of stock picking strategy, a low portfolio similarity level is therefore necessary when looking for superior performance. This paper complements previous studies which argue that deviating from the norm tends to pay off. While I define the "norm" as being a common active portfolio, previous papers have used different settings (e.g. benchmark, systematic factors). For instance, Amihud and Goyenko (2013) find that mutual funds' R2 from a regression of its return on a multi-factor model predicts performance.[1]

This methodological approach is growing in popularity in the mutual fund literature. Recent studies have also built network of funds based on informational linkages and found that fund importance in the network is generally associated with higher performance (see e.g. Ozsoylev et al. 2014 or Rossi et al. 2015). I provide an alternative way of looking at fund centrality where this latter can be disadvantageous in the context of holdings ties among funds. On the side of securities, Anton and Polk (2014) look at the asset pricing implications of securities' ownership structure and show that shared ownership between stocks predicts cross-sectional variations in return correlation. On the fund level, Wahal and Wang (2011) look at the competition among funds in terms of fees and flows when attracting investors in the case of new entrants.[2] I show that competition among funds in delivering performance is also of interest. When following a given strategy, a set of funds will also be in competition when delivering returns to investors and this performance can be impacted if acting too

[1] Titman and Tiu (2011) draw similar conclusions about hedge funds.

[2] Other related papers include for instance Pareek (2012) which looks at the implications of mutual funds information network for stock returns and volatility and explains that important similar holdings is related to information linkages. Hochberg et al. (2007) and Hochberg et al. (2010) study venture capital firms (VCs) network and find that central VCs firms have significantly higher performance.

similar to the other funds.

Various explanations have been given as to why investors act similarly. Herding behavior has been proposed as a potential answer for instance but with relatively weak empirical support.[3] I use the argument of common information environment and strategy overlap as significant factors driving correlated trade directions. The dataset used for this study shows that the number of U.S. domestic equity mutual funds soared from approximately 200 in 1980 to almost 1,500 funds 30 years later.[4] With a number of potential strategies being much smaller and growing at a much smaller pace, strategy overlap naturally arises. Additionally, profit opportunities are in short supply and investors compete against each other for it as in Challet and Zhang (1997). Hence, it is reasonable to assume that competition and crowded positions have a negative impact on one's performance based on market efficiency.

Stein (2009) argues that arbitrageurs can face externalities related to crowded trades when following a given strategy in which they usually cannot tell the number of peers having the same strategy and entering the same trade. Therefore, these "unanchored" strategies result in negative externalities for arbitrageurs or more generally, sophisticated investors.[5] The crash in August 2007 is often cited as a notable example of the consequences of strategies overlap coupled with excessive leverage among institutional investors.[6] In a similar spirit, the motivation for this paper is to investigate the "crowded-trades" effect, as portfolio similarity gives a proxy of this latter, on mutual fund performance using network analysis.

The structure of this paper is as follows. First, I outline a simple market model based on Challet and Zhang (1997) to derive the main hypothesis in Section I. Second, I present the

[3] See, for example, Bikhchandani et al. (1992), Lakonishok et al. (1992), Grinblatt et al. (1995), Nofsinger and Sias (1999), and Wermers (1999).

[4] More generally, the asset management industry grew rapidly in the last decades relatively to the individual investors. As documented by French (2008), individual investors' holdings in the equity market went from 47.9% in 1980 to 21.5% in 2007. In the light of the increase in size of this industry, the position of institutional investors has become more and more important for regulatory bodies regarding financial institutions other than banks (see "Asset Management and Financial Stability", Office of Financial Research, Department of the Treasury, 2013).

[5] Unanchored strategies include positive feedback investment strategies for instance, momentum (De Long et al. 1990).

[6] Hedge funds in this particular case, see Khandani and Lo (2011).

dataset in Section II. Then, Section III presents the network analysis of the mutual funds' holdings database. Finally, I present the results in Section IV, and conclusion in Section V.

I. The Model

In this section, I present a simple generic model of decision-making with network externalities to rationalize the impact of connections among fund managers on performance. This model of collective behavior is a simplified version of the model in Arthur (1994) and the Minority Game by Challet and Zhang (1997) and helps understand the empirical results of this paper. In this model, bounded rationality plays an important role in studying the "myopic" behavior of fund managers and especially in the presence of crowded-trades effects. By observing prices movements, fund managers are unable to account for the actions of their numerous peers, i.e. it cannot be precisely *deducted.*

A. Players and payoffs

There is a population of N fund managers and a set of M assets. Each fund manager $i = \{1, ..., N\}$ needs to choose independently to buy or sell the asset $m = \{1, ..., M\}$. The choice is denoted $x_{i,m} = \{-1, +1\}$ ($x_{i,m} = +1$ means that manager i buys asset m and vice versa). Time is discrete and goes from $t = \{0, ..., T\}$.

Each manager gets a payoff related to his investment in asset m:

$$v_{i,m} = -x_{i,m} X_m \tag{1}$$

where $X_m = \sum_{i=1}^{N} x_{i,m}$ represents the aggregate demand for asset m. The actions of other fund managers affect agent i's payoff in the form of negative network externalities. The payoff for i across all assets is:

$$v_i = -\sum_{m=1}^{M} x_{i,m} X_m - c \qquad (2)$$

where c represents an exogenous cost of investing (e.g. search costs or transaction costs). The price of asset m will be determined as a function of X_m, i.e. the excess demand. Fund managers are price-takers, which means that they don't include their price impact when forming their optimal strategies (they take X as given).[7] This game is called Minority Game (Challet and Zhang 1997) because a manager taking an opposite action to the majority is in the winning side. For example, suppose 60 managers buy asset m while 40 managers sell it. The excess demand (X_m) for the asset m will be positive and therefore the price, set according to X_m by a zero-profit market maker, will be pushed up. Those 40 managers selling it will be on the winning side as they will sell it at a higher price. On the other hand, the 60 buying managers will pay more than what they would if they were a minority.

B. Information environment and inductive reasoning

The information set Ω is composed of a private signal (or predictors, Arthur 1994) and a public information set. There is a set $P = 2^{\tau}$ of predictors/strategies for each asset, where τ is the number of previous periods that the fund takes as input to form its strategy. For instance, if $\tau = 1$, the fund considers only the previous period t information (e.g. the price evolution of yesterday, through X_m) and will decide upon this information. Hence, it has 2 possible strategies. When $\tau = 2$, it will consider the previous two periods, and so on. Thus, τ refers to the *memory* of the fund managers. It is important to note that funds will not form their strategies by directly looking at what the other funds are doing but by looking at the price of each asset, which is a function of the excess demand (the actions of other funds).

[7] This assumption can be argued as practitioners take this issue seriously when submitting orders. However I am interested in negative externalities resulting from the overlap of strategies which is independent of the intelligence of the players of this game.

The ratio P/N determines the commonality of the strategies across funds. When $P < N$ the number strategies is smaller than managers and therefore it is more likely that at least 2 managers will act similarly. On the other hand, if $P > N$, funds will be random players.[8] The timing of the model can be summarized in Figure 1. The strategies generate signals and the fund decides which strategy to pick and what action to take. There is no risk of no execution (perfect liquidity), a zero-profit market maker looks at all the orders and form a price. The price of asset m will go up if we have more buyers. Then, the logic of the minority game is that the sellers will have the advantage as they will sell at a higher price. The order is automatically executed ex-post the price formation. As I will apply this model to the mutual funds industry, if a manager wants to sell an asset, he must have bought it previously. Hence, at the start of the game, each manager is endowed a portfolio with a set of asset. The players have an inductive reasoning mechanism and follow their own rules such as: if price went up yesterday then it should go up today, i.e. any forecasting rule that maps price history into decision. However, it is not specified what these rules are (there are $P = 2^{\tau}$ of these). After receiving their payoff, players update the scores they keep on each strategies and select the strategy with the highest score (see Appendix A for more details).

Each manager selects a set of strategies S out of the entire set P, hence having limited abilities. A number of reasons can oblige funds to limit themselves (e.g. limited technology, geographical bias, cost of searching for strategies). This cost as exogenous and hence c in equation 2 includes the cost of monitoring these strategies but also switching from one to another.

C. Equilibrium definition

Definition 1. *The Nash equilibrium in the single-stage game will be determined by the optimal choice of agent i for a given asset $x^*_{i,m}$, and in fine, the number of agents across*

[8]It is also reasonable to assume that fund managers are not uniformly distributed across assets. Some funds might be specialized in some sectors or style (see e.g. Barberis and Shleifer (2003)) causing a second layer of crowding effect. I will focus here on the distribution of strategies across funds.

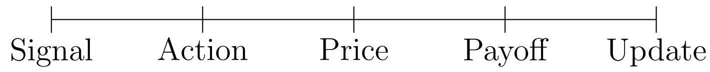

Figure 1.
Timeline of the decision making process. First the manager gets a signal. At the first period, it is random. Then the managers submit orders simultaneously (action). Then, the market forms a price and execute the orders. Afterwards, the managers compute their payoff, and finally update their scores on the strategies according to the payoff.

strategies, i.e. the optimal actions chosen by all the agents ($X_m = \sum_i x_{i,m}$). For a given asset m (partial equilibrium), assuming that agents are price takers and that supply can always satisfy the excess demand, we have:

$$v_{i,m}(x^*_{i,m}, X_m, \Omega) \geq v_{i,m}(x_{i,m}, X_m, \Omega) \tag{3}$$

Suppose N is even and $P < N$. Then the Nash equilibrium will be for each fund manager to choose $x_{i,m} = +1$ 50% of the time and $x_{i,m} = -1$ the rest of the time. Indeed, it is easy to see that no player would have any incentives to deviate as it would end up in the majority and hence have a negative payoff. Now, with N being odd, we would have a large number of equilibria. Suppose $N = 2k + 1$. $2k$ players are equally split in two groups ($x \in \{-1; +1\}$). The last player will be indifferent between $x = -1$ or $x = +1$ as he would end up in the majority and his payoff will be the same whatever action he chooses and this is the same for the $2k$ players as one would join a majority when switching action. Hence, without coordination among players, this is an incentive compatible Nash equilibrium and there are many combinations possible, more specifically $\binom{2k+1}{k}$.

D. Network game

Following Galeotti et al. (2010), I now define this non-cooperative game on an undirected network (N, g) where $g \in \{0, 1\}^{nxn}$ represents the associated adjacency matrix where $g_{ij} \neq 0$ denotes fund i being connected to j when having a similar investment position. We have M adjacency matrices, one for each asset. The sum of all matrices gives us a weighted matrix where each entry will denote the total number of similar investment positions between fund i and j. Let $N_i(g) = \{j|ij \in g\}$ be the set of neighbours of agent i in the network (the number of agents having bought the same asset). The degree of node i is simply $d_i(g) = |N_i|$. I am interested in seeing how the agent i's payoff is affected by $d_i(g)$.

The Nash equilibrium for this network game is as follows:

$$v_i(+1, g, \Omega) > v_i(-1, g, \Omega) \quad \text{if} \quad \frac{d_i(g)}{N} < \frac{1}{2} \tag{4}$$

Suppose we are considering the ownership network of asset m and fund i decides to buy this asset. As we can see, *ceteris paribus*, an increase in $d_i(g)$ will move the condition for a buy order to be optimal closer to the critical point of 1/2 and thus bring the payoff slowly toward zero. In other words, the probability to have a positive payoff decreases with $d_i(g)$. As I will investigate the case of mutual funds (excluding the actors involved in short selling activities), the case that apply is Equation 4, i.e. when staying long on a position is more advantageous than selling or closing the position.

PROPOSITION 1: *For any $P < N$, τ, S and adjacency matrix related to asset m, we have $v_i(+1, g, \Omega)$:*

1. *decreasing in $d_i(g)$*
2. *increasing in N*

This result is interesting as long as the hypothesis that financial markets are minority games hold. When markets are crowded, competition for the limited resource that are profit opportunities will be tougher. With a large number of potential strategies, funds can act randomly and avoid crowded-trades effects.[9] Suppose a fund has the minimum degree from the network (N, g), a minority game hypothesis assume that this fund will have a higher payoff compared to the other funds. The economic intuition behind this hypothesis could

[9]It is worth mentioning that strategies can also be thought of assets. Fewer assets would result in crowded situations and vice versa, i.e. the ratio M/N will act as P/N.

be that this fund has very good private information or good strategy to be one of the few to invest in a small set of assets. On the other hand, a fund with a very high degree has a lot of common holdings with its peers. This fund, which is in the majority, is investing in a highly similar fashion compared to the others funds. Its positions are close to the benchmark and is therefore more passive than active. This fund is less likely to outperform, at least his peers. This advocates for the present study of the position of the fund within the network, i.e. its centrality, based on the first item of Proposition 1.

II. Data

I collect the quarterly mutual holdings data from CDA/Spectrum S12 database during the period 1980Q1-2014Q4. Mutual funds are required to report to the SEC their holdings on a quarterly basis. The database also contains voluntary reports, sometimes in a higher frequency (monthly). The database provides two dates for the holdings: a report date and a filing data, which is often different. Following prior literature, I assume that there is no trading between the filing date and the report date in order to have a snapshot of the holdings of a mutual fund at a given date and adjusting for stock splits when necessary.

I obtain fund characteristics (total net assets, monthly net returns, expense ratios, turnover ratios, fund age) from the survivor-bias-free CRSP mutual fund database. I compute gross returns as the sum of net returns and 1/12 of the annual expense ratio. I aggregate the net and gross returns for multiple share classes to a get a weighted average and aggregate monthly returns to obtain quarterly returns. I use MFLinks to merge the two databases.

Since the focus is on domestic actively managed equity mutual funds, I restrict the investment objective code reported by CDA/Spectrum to be either aggressive growth, growth, growth and income, balanced, or unclassified following Lou (2012). Furthermore I require total net assets (TNA) reported by CRSP to be minimum \$1 million, the number of positions to be equal or greater than 10. and restrict the CRSP investment objective code to start with ED (equity domestic).[10] Finally, I exclude index funds using the CRSP index flag and their names. Table I reports summary statistics of the final database. Table I summarizes the dataset with the number of funds, the mean and median TNA and total equity holdings (TEH) and the number of assets held as of December of each year. I end up with a dataset at this stage of 113,030 fund-quarter observations and 3,585 distinct mutual funds. We can see a clear increase in the number of funds but also in asset under management over the years (see Figure 3).

A. Other Data

Following Franzoni and Schmalz (2014), I compute net fund flow of fund i at quarter t as:

$$flow_{i,t} = \frac{TNA_{i,t} - TNA_{i,t-1}(1 + RET_{i,t})}{TNA_{i,t-1}} \tag{5}$$

Due to the presence of extreme values in CRSP as mentioned in Elton et al. (2001) I winsorize the net fund flows at the 1% and 99% levels. This procedure slightly reduces the dataset.

I compute the total style drift for a given fund as in Wermers (2012):

$$TSD^{l}_{t'} = \sum_{j=1}^{N} (w_{j,t'} C^{l}_{j,t'} - w_{j,t'-1} C^{l}_{j,t'-1}) \tag{6}$$

where $w_{j,t}$ denotes the portfolio weight on asset j on year t' (June 30th of the year).[11] $C_{j,t'}$ is the characteristic-based benchmark returns of asset j in year t' and style dimension l (value, momentum, or book-to-market) as in Daniel et al. (1997).[12] Then, I average across the style

[10]The CRSP style objective code maps the existing investment objective codes (Wiesenberger, Strategic Insight, and Lipper Objective Codes) in order to have a continuous code. It is a four characters code that correspond to four level of granularity. For example, the code EDYG denotes an equity (E) domestic (D) with style (Y) defined as growth (G). I specify the code for each fund to begin with ED in order to have domestic equity funds only and exclude short, hedged, and option income funds using this mapping.

[11]The superscript on t denotes yearly data as opposed to the quarterly holdings data.

[12]The DGTW (Daniel et al. (1997)) benchmarks are available via http://www.smith.umd.edu/faculty/rwermers/ftpsite/Dgtw/coverpage.htm. The coverage for the DGTW benchmarks is 1975-2012. Therefore, any analysis performed using this measure stops in 2012.

dimensions l :

$$TSD_{t'} = \sum_{l=1}^{L} |TSD_{t'}^{l}| \tag{7}$$

III. Network analysis

Using mutual funds holdings I am able to perform a network analysis that will give insights related to the fund level variations in terms of portfolio similarities and its consequences on fund performance. More specifically I will use the concept of centrality that graph theory defines as the most important nodes in a network. In social network analysis, central nodes will be associated to influential agents. Since the following network analysis does not rely on social connections, a central fund will be defined as a fund that is highly similar in terms of holdings relatively to its peers.

In this section, I would like to start from an adjacency matrix. This square matrix $g \in \{0,1\}^{nxn}$ is the matrix representation of a given graph such as:

$$g = \begin{bmatrix} 0 & 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 \end{bmatrix}$$

Figure 2 shows the graph representation of the adjacency matrix g. First, we can see that the matrix g is symmetric and have zeros on the diagonal. Second, the graph in Figure 2 shows easily the difference in terms of centrality. Node 2 is more central in terms of degrees/connections with the other nodes (4 edges compared to the other nodes that have 3 edges). This is the first measure I will compute across funds.

The first challenge when building an ownership network for mutual funds from the data is that I cannot obtain directly a square ownership matrix. Let's define $W_{i,j,t}$ as the bipartite

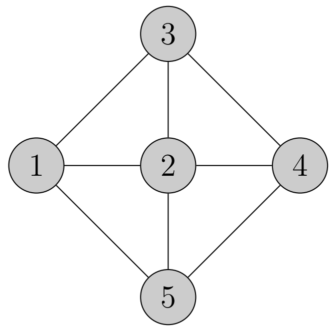

Figure 2.
Graph representation of matrix g.

matrix of holdings of fund i in asset j at quarter t in dollar value. Following Gualdi et al. (2016), I map this matrix $W_{i,j,t}$ into a binary matrix $A_{i,j,t} = sign(W_{i,j,t})$ and project it into a monopartite adjacency matrix:

$$O_t = A_{i,j,t}A'_{i,j,t} - I_n \sum_j A'_{i,j,t} \tag{8}$$

where I_n is the identity matrix of size n (the total number of funds at quarter t). This standard projection method allows us to obtain a fund level portfolio overlap measure. Indeed, O_t is a square matrix where the entries represent the number of securities that a given fund i has in common with any other fund $-i$ but itself (i.e. zeros on the diagonal). For instance, if $O_{1,4} = 20$, it means that fund 1 and fund 4 share 20 positions in common (among the M assets). This will give us a weighted undirected network where the links will have an intensity/weight represented by the overlap between two nodes/funds.[13]

The graph I obtain is close to a complete one. Hence, no visual representation would be interesting. However, we can see in Figure 4, the empirical densities show a significant proportion of nodes with few connections that grows over the years. Figure 5 shows the empirical counter-cumulative distribution function (CCDF) in a log-log plot. As in Acemoglu et al. (2012), the empirical CCDF can be approximated with a distribution in the power law family

[13] I don't use the matrix $W_{i,j,t}$ as this projection method wouldn't give the same meaningful overlap measure.

(here an exponential distribution with scale of 10). This confirms the importance of the tails in the distribution of the degrees and the presence of funds with extreme degrees.

A. *Style Cluster*

The previous subsection presented the network of all actively managed mutual funds. However, as noted by Hoberg et al. (2015), it is relevant to investigate the performance of a given fund relative to its rivals in terms of style dimension. I refine the previous network by connecting funds to their style-peers. This gives style clusters within the complete network.

I classify funds within their style categories using their holdings. With the characteristic-based benchmarks of Daniel et al. (1997), we have a list of stocks annually ranked by size, book-to-market, and momentum. I focus here only on small-cap and large-cap funds. I identify small-cap funds by requiring them to have at least 75% of their holdings invested in the lowest quintile of size (and highest quintile for large-cap funds). Style could also be identified using their names but a holdings-based measure is more precise as names cannot prevent funds from significantly deviating from their advertised style category.

With this style cluster within the whole network I expect a different picture of portfolio similarities within investment style categories. While portfolio similarities for all actively managed mutual funds are also informative, investors often group stocks in different style categories and mutual funds focus, specialize, and advertise in these groups (Barberis and Shleifer 2003). Hence, this style cluster analysis complements the network analysis of all mutual funds.

B. *Centrality*

The next step in the analysis is to compute summary measures at the fund-level. The centrality measure will identify funds that invest in crowded positions. It measures a node's importance in the network. A highly central fund has many positions overlapping with its peers and hence faces tougher competition for profit. The degree centrality represents the number of links a given fund has. In a simple network g with N nodes, the degree of the node i is denoted $|N_i| = \sum_j^N sign(o_{ij})$. This is the simplest centrality measure.[14] In the weighted network, the degree centrality of a mutual fund will be its number of edges weighted by the respective overlaps with each neighbour. More formally, we have:

$$d_i^w = \sum_j^N o_{ij} \tag{9}$$

where o_{ij} is the overlap or weight (i.e. the overlap of fund i with fund $j \neq i$).

Other measures I don't consider include eigenvector, betweenness or closeness centrality for instance. I compute the two aforementioned measures for the complete network and the style clusters network. Table II shows the summary statistics of the centrality measures. [15] Both measures exhibit high variations by looking at the standard deviations with respect to the mean and median. As expected, the non-weighted degree measure is much smaller compared to the degree one.

When looking at which funds are on the extremes positions in terms of centrality measures we can see that style matters. Table III shows the top 5 and bottom 5 funds for the years 1980 and 2010. If relying at this stage on the name of the funds as a potential indication of the style category Table III shows that small-cap funds more likely in the bottom of centrality ranked funds in 2010 while very few funds were categorised as small-cap in 1980. Growth funds are more present in the top centrality ranked funds. Although I only report the top 5 and bottom 5 funds, this table shows already that, on average, we find higher net returns among low central funds.

[14] I denote it in the next section as $Degree_{i,t}^{nw}$ where nw represents the *non-weighted* measure.
[15] Those computed within the style clusters are denoted with a *Style* before.

IV. Results

A. Determinants of Centrality

I investigate in this subsection the driving factors of the centrality measures. Using a regression approach, I use the centrality measures as independent variables and select fund flows, fund age, the number of positions, fund size, past return, and fund turnover as potential explanatory variables.

As mentioned in the literature, mutual funds tend to increase their current positions following inflows (Lou 2012). Hence, fund flows to play an important role in the fund's position in the network as inflows will drive a fund to increase its positions and thus its weighted degree measures. I find that fund flows are significantly and negatively affecting centrality. A one-standard-deviation increase in fund flows will lead to a decrease of degree of approximately 722 ($t = -8.46$) (Table IV). Previous work (Gruber 1996) shows that flows positively predict performance ("smart money effect"). The negative relationship between fund flows and centrality is consistent with this latter effect.

Additionally, a young fund will be more likely to have fewer positions and therefore its age will influence its centrality as well as the number of positions. Along with age, centrality is capturing a size effect with bigger funds being highly central funds while small funds will invest in a smaller number of positions. Table IV shows that the number of positions is the main determinant of centrality. An increase of one-standard-deviation of the natural logarithm of a fund's number of positions leads to an increase in non-weighted degree of approximately 120 ($t = 20.91$). The number of positions a fund has is directly related to the fund's position in the common ownership network. Being directly in the manager's control, it shows that a high number of positions is not necessarily for the best as centrality negatively affects net return and tracking error. Moreover I find that age and size positively and significantly affect centrality.

Then, I investigate the predictive power of past return on centrality. Table IV shows that high return leads to lower centrality. This comforts the hypothesis that centrality and fund performance are negatively related. I formally test this prediction in the next subsection.

Finally, I use fund turnover as a potential determinant of centrality. As it is also under the manager's control, a high turnover will have a more fluctuating position within the network and therefore should affect its centrality. Interestingly, turnover positively affects non-weighted degree centrality while negatively affecting degree. I believe the last measure to be more concluding as it includes the importance of the linkages in the network. A high turnover being associated to lower centrality is an indicator that aggressive managers are better performing.

B. Centrality and Fund Performance

In this subsection, I test the predictive power of centrality measures using a regression approach. Controlling for fund characteristics, I run the following predictive regression:

$$ret_{i,t} = \beta_0 + \beta_1 Centrality_{i,t-1} + \gamma Controls_{i,t-1} + \epsilon_{i,t-1} \tag{10}$$

$Centrality_{i,t-1}$ is the main independent variable. I consider the two aforementioned centrality measures for this predictive regression, i.e. non-weighted degree, and degree. The control variables include fund's Carhart four factor alpha, style drift as in Wermers (2012), fund flow, fund age, number of stocks in the portfolio, fund size, expense ratio, past return, and portfolio turnover.

Table V presents the results using a pooled OLS estimation approach with time fixed effects and standard errors clustered by funds. This methodology account for potential time series and cross-sectional dependence in the residuals as documented by Petersen (2009). I find that all centrality measures are negative and significant at the 1% level in the main specifications. For instance, Table V shows a coefficient of -0.116 ($t = -11.25$) for degree when controlling for fund flow, fund age, number of stocks in the portfolio, fund size, expense ratio, past return, and portfolio turnover. This means that, holding other variables constant, a one-standard-deviation increase in the weighted degree leads to a reduction of the fund (net) quarterly return of 12 basis points.[16] The main coefficient gives similar conclusions for the non-weighted degree measure. The other results are consistent with prior literature. For instance, a higher number of stocks, flows, and age tend to be positively associated to fund return, while size and expenses are negatively associated with performance. In the last specifications, I include a style dummy using the CRSP style investment objective code which maps the Wiesenberg, Strategic Insight, and Lipper objective codes. Considering this style fixed effects reduces the statistical significance of centrality measures, yet the two degree measures (non-weighted and weighted) are still statistically significant (at the 10 and 1% level respectively). On the other hand, if I control for potential deviations for the style objective by including the Style Drift measure (Wermers 2012), I find that centrality still negatively and significantly affects performance no matter how the fund deviates from its style objective.

These results confirm my hypothesis. Higher centrality leads to lower returns. The economic rational being that a given fund is less likely to perform or outperform if it owns (in relatively high proportions) the same positions as everyone else. On the other hand, a fund with low centrality, located in the edge of the network can extract more profits for himself. This can be an indication that this low central fund has some piece of information that the others don't have and/or that it is good at picking stocks.

C. Long Term Effects and Timing

In this section I investigate the long term effects of centrality on performance. More precisely, I test the effects of portfolio similarities in the previous 6, and 9 months as well as the change of centrality, i.e. the difference between centrality in the last quarter and the

[16]The results still hold using other performance measure as dependent variable (gross return, tracking error, Carhart four factor alpha). When using Carhart four factor alpha as the dependent variable, I find that only degree is negatively and significantly affecting the fund's abnormal return.

centrality 6, and 9 months ago.

Table VI shows that the effect of centrality tends to last beyond one quarter. The centrality of a given fund 6 months ago has a negative and statistically significant effect on net return. A increase of one-standard-deviation in centrality 6 months ago leads a reduction of net return today of 5 basis points for the degree measure. The effect of $Centrality_{i,t-2}$ tends to disappear if considering $Centrality_{i,t-3}$ and controlling for funds' characteristics which confirms a high level of persistence. The centrality measure in the previous quarter along with the centrality 9 months ago are negative and significant when controlling for fund characteristics.

Moreover, the change of centrality can be determinant for the fund's performance. For instance, a fund pursuing a momentum strategy can benefit from a high portfolio similarity as long as he entered before his peers. Hence, a positive change of centrality should lead to higher returns. This follows Stein (2009) with the crowding effect being relevant when considering unanchored strategies such as momentum. Therefore, if pursuing a given long position, it is better to enter as early as possible while other people buying subsequently can only benefit this long position. Table VI shows that, for the non-weighted degree, a change of one-standard-deviation in the difference between the lagged 3 months and lagged 9 months leads to a significant positive change in performance of 2 basis points. This change of centrality effect is however absent the non-weighted degree.

D. Centrality and Style

In this subsection I contrast the previous results obtained from the full network with the style clusters network analysis. I define small (large) cap funds by requiring each fund to have at least 75% of its positions invested in companies in the lowest (highest) size quintile.

Using a regression approach with time fixed effects and standard errors clustered by funds, I test the significance of *style* centrality measures on performance. I find in Table VII Panel A that a positive change of one-standard-deviation in non-weighted degree leads a reduction of quarterly net return of 5 basis points ($t = -1.98$) and 3 basis points ($t = -1.77$) for the degree centrality measure. Hence, after controlling for style effects, either as a fixed effects as in Table V or beforehand when computing the centrality measures as in this subsection, the predicted result that low centrality is beneficial for mutual funds' performance is still present. A fund that is too similar to its style peers will have a lower performance.

E. Centrality and Active Share

Centrality measures in the context of common ownership network capture the degree to which a fund a similar to its (style, when specified) peers. To some extent, it captures a deviation from the norm in a similar fashion as the Active Share measure does (Cremers and Petajisto 2009). While this latter shows how much a fund deviates from the benchmark, centrality measures presented in this paper go beyond this scope. It can tell how much a fund differs from the other active funds. A fund with high Active Share can tell how much a fund is active but a fund with a low degree can tell how much a fund is good at identifying good positions relative to its fund competitors. Suppose a given fund has a high Active Share and high degree. This will tell us that this fund is actively deviating from the benchmark but the centrality measure will tell us that its peers have done the same, i.e. they have deviated from the benchmark and invested in the same positions. Therefore it is unlikely that a fund with a high degree will be able to outperform its peers, while it is clear that a highly active fund in terms of Active Share can lead to better performance relative to the benchmark.

Table VIII (Panel A to C) shows various measures of correlation between centrality measures and Active Share. A low ranked fund in terms of non-weighted degree will be highly ranked in terms of Active Share one third of the time (Kendall's rank correlation of 34%). This correlation is higher for degree (57%). Other measures of correlation confirm the negative relationship between centrality and Active Share, with the strongest relationship for the degree centrality measure .

I investigate the robustness of the main regression specification from Table V to the Active Share measure in a regression approach (Panel D of Table **??**). This table shows that centrality is still statistically significant when controlling for Active Share confirming the hypothesis that centrality in common ownership network captures other relevant information missing in the Active Share measure. When controlling for Active Share, a one-standard-deviation increase in degree leads a reduction of quarterly net return of 6 basis points ($t = -3.87$). As in the previous tests (Table V), when including a style fixed effect, results still hold for degree centrality but not non-weighted degree.

V. Conclusion

In this paper, I build a fund level centrality measure that captures portfolio similarities in crowded markets. Specifically, the more central a fund manager is, the more crowded trades he has. Assuming that profits are scare resources, the more central is a fund, the lower its profits should be. I document this relationship and find that centrality is significantly associated to lower returns and is a significant predictive measure of performance for actively managed U.S. equity mutual funds from 1980Q1 to 2014Q4, controlling for fund's characteristics. The main results hold across three portfolio similarity measures: non-weighted degree, and degree where the latest is the most robust. These findings are consistent with prior theoretical literature (Stein 2009). While the theoretical framework builds on unanchored strategy such as momentum, I show in this respect that a change of centrality leads to a higher performance. Hence, investing earlier than its peers in a given position, a fund can be better off. When accounting for potential style effects, the main findings are still significant whether it is in the form of a style investment objective code or a centrality measure constructed in a style cluster network .

Besides from the crowded-trade effect, we can also rationalize these findings by arguing that a low central fund is more likely to outperform (its peers) when having few similar positions. Indeed, deviating from the norm, being its peers or simply the benchmark (Cremers and Petajisto 2009), is beneficial as documented in previous studies. I show that this new portfolio similarity measure captures some aspects in common with other measures of deviations however they have their differences. The empirical findings presented in this paper offer a novel angle in assessing mutual funds. Previous measures show how active a fund is relative to the benchmark for instance, the centrality measure built in this work can tell an investor how different a fund is from its peers or how original it is.

This paper complements the growing literature in financial networks of common ownership. The main findings give a different perspective compared to recent studies. In social networks, centrality has often been associated to higher performance, where connections in a network allow for better information flow between the participants. The contribution is a new case where centrality is negatively associated to performance. In the present context, I exclude potential benefits from informational linkages as the common ownership network includes as much as 1,400 funds (see Figure **??**).

The common ownership network framework has offered multiple directions for research and still has potential for future works. Several papers have focused on the importance of the ownership structure and the asset pricing implications (e.g. Anton and Polk 2014). The methodology in this paper can offer multiple extensions such as studying the underlying dynamic of the position of the fund in the network or the time evolution of the links among the active mutual funds.

Appendix A. Appendix

Appendix A. Computer experiment

To further investigate the properties of the Minority Game (Challet and Zhang 1997), I implement a computer simulation and compare its results with the dataset. More precisely, I simulate a set of funds competing as described in the model from Section I and observe the impact of the aggregate demand and network structure on the profit of a single fund that I assume having a long position during the entire length of the game ($t = 0...T$).

At the beginning of the game the fund (or its manager) selects a strategy s at random. To update its information, the fund will assign a virtual score, denoted $U_{i,s}$ where $s \in S$, for each strategy s which reflects its cumulated payoff. All strategies have a score of zero at the beginning of the game. At each time step the fund will pick the strategy with the highest score. The virtual score for each strategy s is as follows.

$$U_{i,s}(t+1) = U_{i,s}(t) - x_i^s \frac{X}{N} \tag{A1}$$

The probabilistic choice model associated to the strategies follows a Logit model (see Cavagna et al. (1999); McFadden (1980)).

$$Prob\{s_i(t) = s\} = \frac{e^{\Gamma U_{i,s}(t)}}{\sum_{s'} e^{\Gamma U_{i,s'}(t)}} \tag{A2}$$

where $\Gamma < \infty$ measures the learning rate. Each agent will choose the optimal strategy (the one with the highest score) and will solve:

$$s_i = \text{argmax}_s U_{i,s} \tag{A3}$$

I define the aggregate demand for a particular asset m at a period t as:

$$X = N_+ - N_- \tag{A4}$$

where N_+ is the number of funds having chosen action $x = +1$ and N_- is the number of fund selling this asset at a given period t. Since we have $N = N_+ + N_-$ we can easily get $N_+ = \frac{X+N}{2}$. Now for the case of the fund being long during the entire simulation of the game, we can see that N_+ is equivalent to $d_i(g)$ which represents the degree of fund i, i.e. the number of funds in its neighborhood or having the same position. Then we can relate $d_i(g)$ to the payoff of fund i which is given by equation1. To have a complete picture of the game for M assets, I simulate the same game M times and aggregate it in order to have a weighted matrix which shows the intensity of the connections. As we can see in Figure 6, degree is associated to a lower payoff as expected by the rules of the game.

REFERENCES

Acemoglu, Daron, Vasco M Carvalho, Asuman Ozdaglar, and Alireza Tahbaz-Salehi, 2012, The network origins of aggregate fluctuations, *Econometrica* 80, 1977–2016.

Amihud, Yakov, and Ruslan Goyenko, 2013, Mutual fund's r2 as predictor of performance, *Review of Financial Studies* 26, 667–694.

Anton, Miguel, and Christopher Polk, 2014, Connected stocks, *The Journal of Finance* 69, 1099–1127.

Arthur, W Brian, 1994, Inductive reasoning and bounded rationality, *The American economic review* 84, 406–411.

Barberis, Nicholas, and Andrei Shleifer, 2003, Style investing, *Journal of financial Economics* 68, 161–199.

Bikhchandani, Sushil, David Hirshleifer, and Ivo Welch, 1992, A theory of fads, fashion, custom, and cultural change as informational cascades, *Journal of political Economy* 992–1026.

Cavagna, Andrea, Juan P Garrahan, Irene Giardina, and David Sherrington, 1999, Thermal model for adaptive competition in a market, *Physical Review Letters* 83, 4429.

Challet, Damien, and Yi-Cheng Zhang, 1997, Emergence of cooperation and organization in an evolutionary game, *arXiv preprint adap-org/9708006* .

Cremers, KJ 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.

De Long, J Bradford, Andrei Shleifer, Lawrence H Summers, and Robert J Waldmann, 1990, Positive feedback investment strategies and destabilizing rational speculation, *the Journal of Finance* 45, 379–395.

Elton, Edwin J, Martin J Gruber, and Christopher R Blake, 2001, A first look at the accuracy of the crsp mutual fund database and a comparison of the crsp and morningstar mutual fund databases, *The Journal of Finance* 56, 2415–2430.

Franzoni, Francesco A, and Martin C Schmalz, 2014, Performance measurement with uncertain risk loadings, *Swiss Finance Institute research paper* .

French, Kenneth R, 2008, Presidential address: The cost of active investing, *The Journal of Finance* 63, 1537–1573.

Galeotti, Andrea, Sanjeev Goyal, Matthew O Jackson, Fernando Vega-Redondo, and Leeat Yariv, 2010, Network games, *The review of economic studies* 77, 218–244.

Grinblatt, Mark, Sheridan Titman, and Russ Wermers, 1995, Momentum investment strategies, portfolio performance, and herding: A study of mutual fund behavior, *The American economic review* 1088–1105.

Gruber, Martin J, 1996, Another puzzle: The growth in actively managed mutual funds, *The journal of finance* 51, 783–810.

Gualdi, Stanislao, Giulio Cimini, Kevin Primicerio, Riccardo Di Clemente, and Damien Challet, 2016, Statistically similar portfolios and systemic risk, *arXiv preprint arXiv:1603.05914* .

Hoberg, Gerard, Nitin Kumar, and Nagpurnanand Prabhala, 2015, Mutual fund competition, managerial skill, and alpha persistence .

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.

Hochberg, Yael V, Alexander Ljungqvist, and Yang Lu, 2010, Networking as a barrier to entry and the competitive supply of venture capital, *The Journal of Finance* 65, 829–859.

Khandani, Amir E, and Andrew W Lo, 2011, What happened to the quants in august 2007? evidence from factors and transactions data, *Journal of Financial Markets* 14, 1–46.

Lakonishok, Josef, Andrei Shleifer, and Robert W Vishny, 1992, The impact of institutional trading on stock prices, *Journal of financial economics* 32, 23–43.

Lou, Dong, 2012, A flow-based explanation for return predictability, *Review of financial studies* 25, 3457–3489.

McFadden, Daniel, 1980, Econometric models for probabilistic choice among products, *Journal of Business* S13–S29.

Nofsinger, John R, and Richard W Sias, 1999, Herding and feedback trading by institutional and individual investors, *The Journal of finance* 54, 2263–2295.

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.

Pareek, Ankur, 2012, Information networks: Implications for mutual fund trading behavior and stock returns, in *AFA 2010 Atlanta Meetings Paper*.

Petersen, Mitchell A, 2009, Estimating standard errors in finance panel data sets: Comparing approaches, *Review of financial studies* 22, 435–480.

Rossi, Alberto G, David P Blake, Allan G Timmermann, Ian Tonks, and Russ Wermers, 2015, Network centrality and pension fund performance .

Stein, Jeremy C, 2009, Presidential address: Sophisticated investors and market efficiency, *The Journal of Finance* 64, 1517–1548.

Titman, Sheridan, and Cristian Tiu, 2011, Do the best hedge funds hedge?, *Review of Financial Studies* 24, 123–168.

Wahal, Sunil, and Albert Yan Wang, 2011, Competition among mutual funds, *Journal of Financial Economics* 99, 40–59.

Wermers, Russ, 1999, Mutual fund herding and the impact on stock prices, *The Journal of Finance* 54, 581–622.

Wermers, Russ, 2012, Matter of style: The causes and consequences of style drift in institutional portfolios, *Available at SSRN 2024259* .

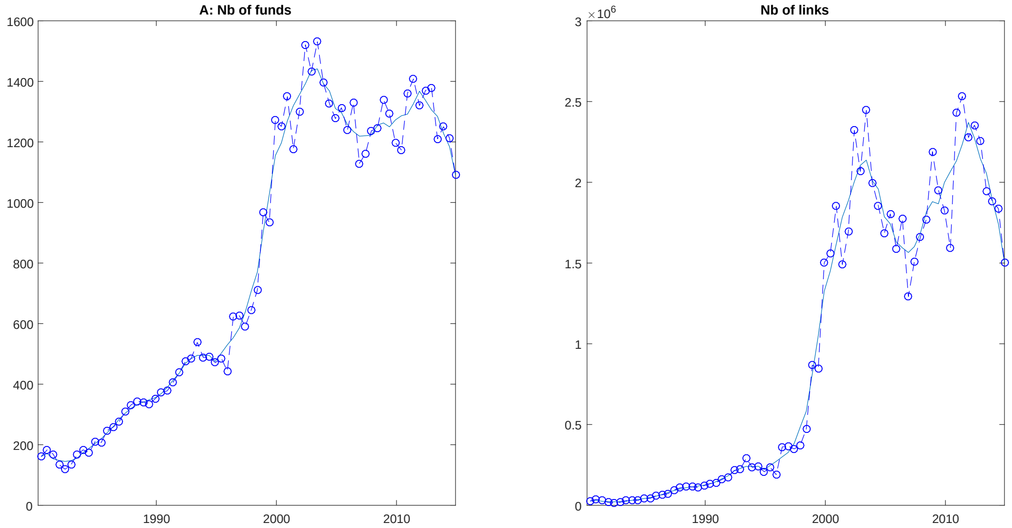

Figure 3.
Evolution of number of funds and degrees/number of links: This figure represents the time evolution on the left of the number of actively managed U.S. equity mutual funds in the sample from 1980Q1 to 2014Q4. On the right is represented the evolution of the total number of links or the sum of the non-weighted degrees for all the funds for each quarter.

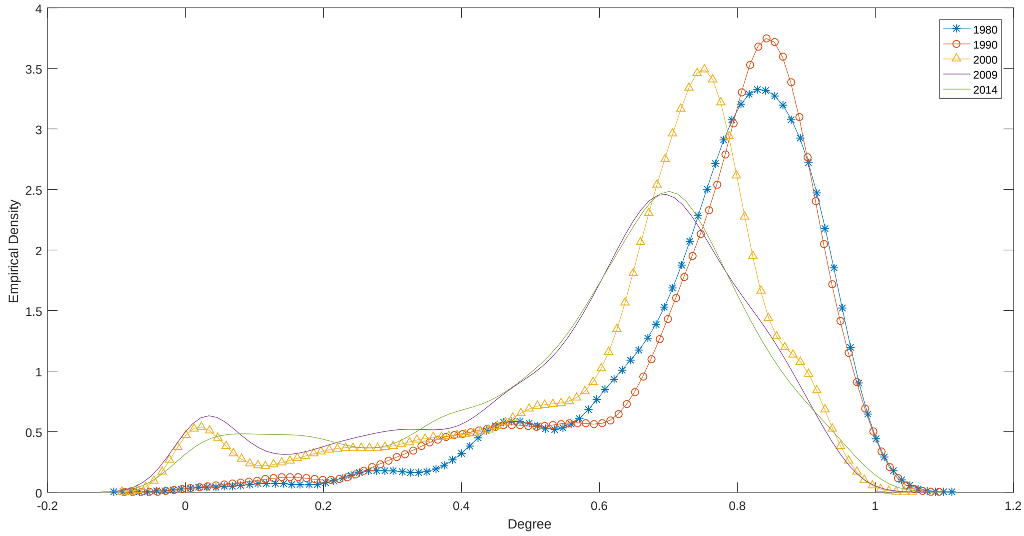

Figure 4.
Empirical densities of network degrees: This figure represents the empirical probability density function of the degrees obtained from the common ownership network. The estimation is based on a non-parametric kernel density estimation.

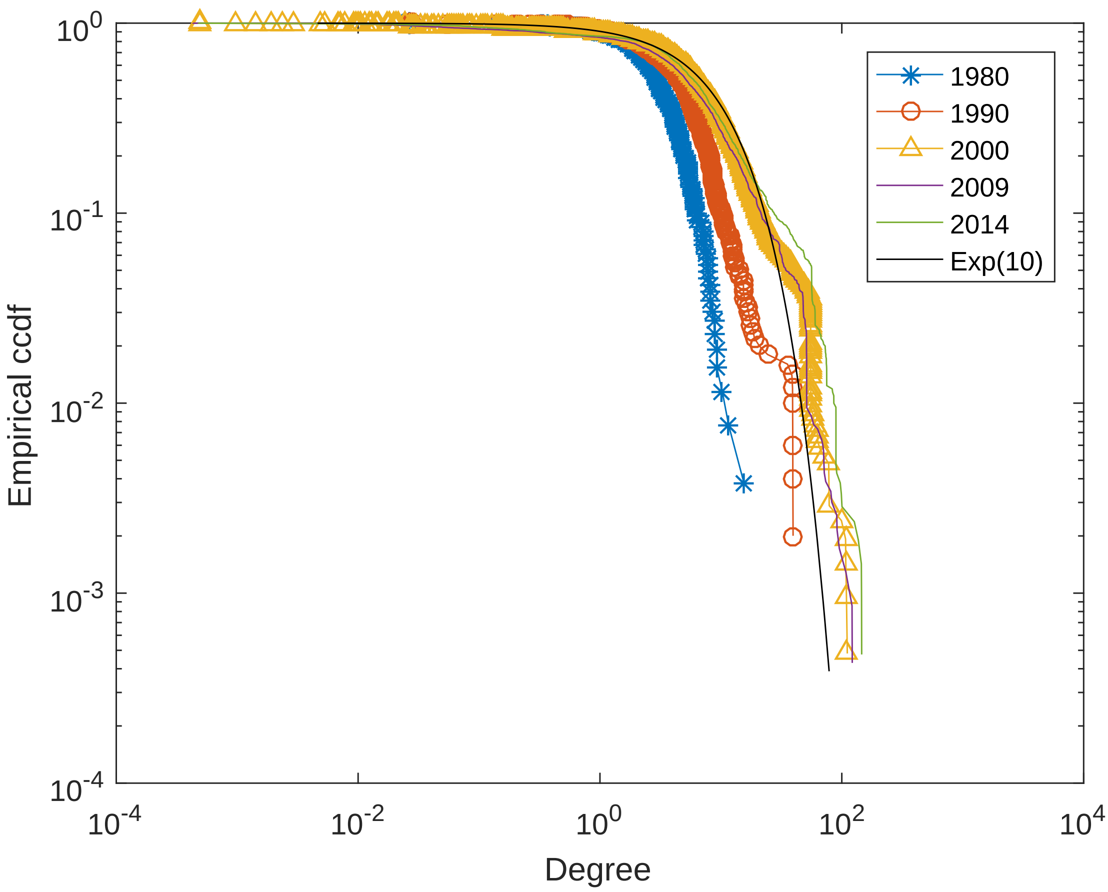

Figure 5.

Empirical counter-cumulative distribution function of network weighted degrees: This figure represents the empirical counter-cumulative distribution function (one minus the cumulative distribution function) in a log-log scale of the degrees of the common ownership network. The estimation is based on a non-parametric kernel density estimation. I fit the empirical CCDF with an Exponential distribution with a parameter of 10.

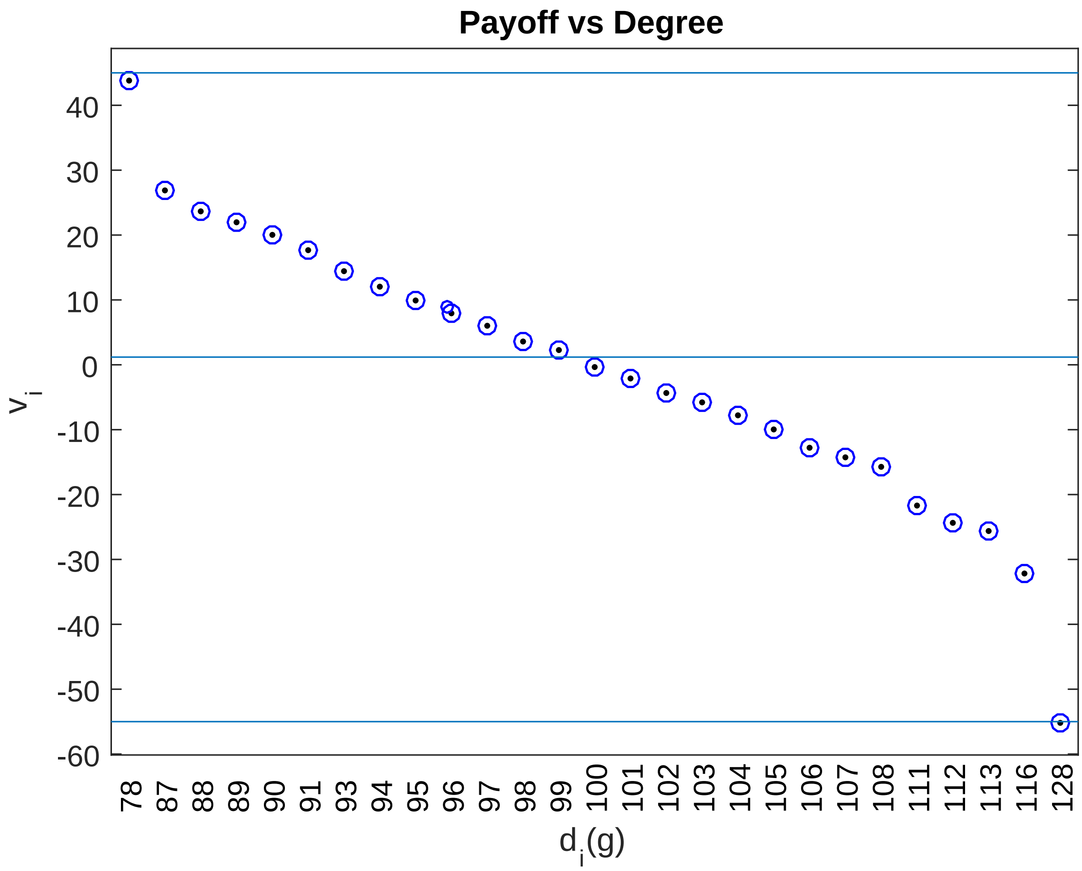

Figure 6.
Computer experiment: This figure represents the result of the computer experiment. I plot the payoff of each fund against the respective degrees using boxplots. I run a computer simulation with the rules of the Minority Game (Challet and Zhang (1997)) with the followings parameters: $M = 100$, $N = 200$, $P = 50$, $T = 150$.

Table I

Sample Summary Statistics: This table presents summary statistics of the mutual funds sample from 1980Q4 to 2014Q4 as of December of each year. The sample includes only U.S. equity as denoted by the Thompson Reuters' investment codes (aggressive growth, growth, growth and income, balanced, unclassified, or missing) with assets above \$1mio. We exclude international, fixed income, and metals funds. Funds' holdings data are obtained from the Thompson Financials CDA/Spectrum database. Funds' characteristics are obtained from the CRSP survivorship-bias-free mutual fund database. The databases are merged using MFLinks. *TNA* represents the total net assets and *TEH* represents the total equity holdings (both in millions \$).

Year	Nb Funds	TNA (\$ M) Median	TNA (\$ M) Mean	TEH (\$ M) Median	TEH (\$ M)Mean	Nb Stocks
1980	189	62.58	173.46	53.76	152.76	2198
1981	136	58.38	146.95	47.01	125.55	2106
1982	137	101.67	202.25	71.73	168.96	2289
1983	185	116.73	255.36	102.17	217.88	3151
1984	212	89.42	239.09	79.35	198.45	3308
1985	250	116.71	308.78	98.28	247.18	3663
1986	279	111.42	339.53	90.99	269.84	3451
1987	335	87.22	313.93	75.50	258.35	3848
1988	344	88.06	296.79	74.14	242.10	3980
1989	355	102.50	370.87	81.74	296.11	3958
1990	381	87.58	326.25	67.20	260.37	3711
1991	440	92.99	405.20	82.42	329.75	3947
1992	488	153.75	575.63	128.67	453.22	4144
1993	496	137.11	548.81	107.17	439.41	5467
1994	481	115.40	494.33	97.63	399.04	5769
1995	459	147.87	685.29	125.44	563.70	6246
1996	670	124.50	615.55	117.17	532.41	6173
1997	664	164.44	944.84	154.39	808.82	6156
1998	1030	153.55	975.14	145.15	884.87	6150
1999	1335	182.20	1097.54	171.57	984.02	6713
2000	1453	170.10	968.77	152.65	855.41	6598
2001	1366	126.70	634.58	118.14	591.80	6036
2002	1481	109.70	655.44	99.45	583.20	5640
2003	1453	139.90	812.73	131.51	731.61	5445
2004	1324	154.25	1084.78	140.42	949.84	5157
2005	1269	193	1253.30	182.35	1140.14	5473
2006	1144	253.35	1583.51	230.97	1404.48	5544
2007	1250	241.15	1489.41	225.88	1313.73	5420
2008	1452	139.10	847.60	131.77	729.35	5194
2009	1308	212.15	1245	192.01	1094.34	4939
2010	1493	269.10	1405.19	248.44	1243.40	5027
2011	1431	260.80	1304.23	240.69	1173.82	4789
2012	1493	349.40	1455.25	319.54	1316.68	4620
2013	1340	476.65	1994.48	439.32	1879.77	4654
2014	1167	568.60	2393.83	516.50	2135.78	4781

Table II

Variables summary statistics: This table presents the summary statistics for centrality measures and control variables for the period 1980Q4-2014Q4. $Degree^{nw}$ represents the non-weighted degree measure while $Degree$ is the weighted-degree centrality measure. The style drift measure stops in 2012 due to the availability of the data.

Panel A: Centrality measure statistics						
-	Mean	Std. Dev.	Median	Min	Max	Nb
$Degree^{nw}$	1,180.44	569.19	1,313	13	2,783	107,155
$Degree$	14,496.24	17,234.08	9,799	14	264,367	107,155
$StyleDegree^{nw}$	530.56	207.09	575	0	917	26,876
$StyleDegree$	8,106.84	5,569.37	7,287	0	44,928	26,876

Panel B: Controls statistics						
-	Mean	Std. Dev.	Median	Min	Max	Nb
TNA ($ M)	1,089.22	4,748.58	186	1.01	195,806.90	107,155
Nb stocks	95.26	119.15	66	11	2,336	107,155
Fund flow	-1.29e-02	9.75e-01	-2.01e-02	-0.34	0.45	95,816
Expense ratio	1.35e-02	6.91e-03	1.25e-02	1.00e-04	1.02	107,155
Turnover ratio	0.89	1.32	0.63	0	60.7	107,155
Fund age (months)	179.43	171.25	127	1	1,085	107,155
Style drift	0.77	0.62	0.61	1.49e-03	10.60	75,013
Active Share	0.81	0.15	0.84	1.50e-02	1	49,136

Table III

Top-bottom central funds: This table presents the top 5 and bottom 5 funds for each centrality measures (percentage) as of December 1980 and 2010 and on average, along with the average net quarterly return (Ret.) for each year and for the whole sample in Panel C.

Panel A: Top and bottom central funds as of Dec. 1980

-	Top 5	$Degree^{nw}$	Ret.	Bottom 5	$Degree^{nw}$	Ret.
1	Dreyfus Fund Incorporated	100 %	2.28%	Lord Abbett Developing Growth Fund	32.02%	2.65%
2	Wells Fargo Premier Large Company Growth Fund	99.60%	1.78%	Scudder Development Fund	35.18%	4.78%
3	Massachusetts Investors Growth Stock	99.21 %	2.94 %	Vanguard Explorer Fund	41.90%	3.81%
4	Elfun Trusts	99.21%	3.91%	Guggenheim Large Cap Value Fund	41.90 %	2.47%
5	Keystone Mid-Cap Growth Fund	98.81%	0.03%	Concord Fund	43.87%	2.34%
-	Top 5	Degree	Ret.	Bottom 5	Degree	Ret.
1	Wells Fargo Premier Large Company Growth Fund	100 %	1.78%	Lord Abbett Developing Growth Fund	3.72 %	2.65%
2	Evergreen Blue Chip Fund	74.93 %	1.43%	Guggenheim Large Cap Value Fund	3.99 %	2.47%
3	Massachusetts Investors Growth Stock	59.85%	2.94 %	Scudder Development Fund	4.36 %	4.78%
4	AB Growth & Income Fund	59.77%	1.78%	Unified Growth Fund	4.43%	3.22%
5	Fidelity Hastings Street	56.95%	2.50%	Concord Fund	4.63 %	2.34%

Panel B: Top and bottom central funds as of Dec. 2010

-	Top 5	$Degree^{nw}$	Ret.	Bottom 5	$Degree^{nw}$	Ret.
1	USAA Cornerstone Moderately Aggressive Fund	100%	1.87%	New Century Opportunistic Portfolio	3.89%	3.70%
2	Reynolds Blue Chip Growth Fund	99.89 %	0.87%	Calvert VP Lifestyle Aggressive Portfolio	4.96%	3.04%
3	American Century Strategic Allocation Aggressive Fund	99.81%	2.59%	Perritt Ultra MicroCap Fund	9.23%	4.75%
4	Fidelity Charles Street Trust Avisor Asset Manager 85%	99.54%	1.33%	Bertolet Capital Trust Pinnacle Value Fund	12.82%	1.14%
5	Wells Fargo Diversified Equity Fund	99.31%	3.64%	Catalyst Small-Cap Insider Buying Fund	15.53 %	1.61%
-	Top 5	Degree	Ret.	Bottom 5	Degree	Ret.
1	Wells Fargo Diversified Equity Fund	100 %	3.64%	New Century Opportunistic Portfolio	0.06 %	3.70%
2	SA US Core Market Fund	99.35 %	0.51%	Calvert VP Lifestyle Aggressive Portfolio	0.08 %	3.04%
3	Wilmington Large-Cap Strategy Fund	84.04%	0.23%	Bertolet Capital Trust Pinnacle Value Fund	0.35%	1.14%
4	SEI Institutional Tax.Managed Large Cap Fund	80.96 %	0.26 %	Federated Prudent Bear Fund	0.39 %	-2.82%
5	Reynolds Blue Chip Growth Fund	75.66 %	0.87%	Forward Salient Select Income Fund	0.43 %	1.54 %

Panel C: Top and bottom central funds on average						
-	Top 5	$Degree^{nw}$		Bottom 5	$Degree^{nw}$	
1	Voya Partners Contrafund Portfolio	100%	3.93%	Fidelity Strat. Advisers US Opportunity Fund	1.63 %	1.89 %
2	Russell US Value Fund	96.81%	0.37%	Sherman Dean Fund	2.12%	-0.45%
3	American Century Strategic Allocation Aggressive Fund	96.36 %	0.49%	Deutsche Multi-Asset Conservative Allocation Fund	2.24%	0.40%
4	Voya Partners Equity-Income Portfolio	96.04%	2.93%	Deutsche Multi-Asset Moderate Allocation Fund	2.35%	1.01%
5	Eaton Vance Tax-Managed Equity Asset Allocation Fund	95.43%	1.11%	PMFM Managed Portfolio Trust	3.04%	2.66%
-	Top 5	Degree		Bottom 5	Degree	
1	SA US Core Market Fund	100%	0.76%	Fidelity Strat. Advisers US Opportunity Fund	0.03%	1.89%
2	Wilmington Large-Cap Strategy Fund	89.32%	1.29%	Sherman Dean Fund	0.03%	-0.45%
3	Wells Fargo Diversified Equity Fund	81.41%	0.44%	PMFM Managed Portfolio Trust	0.04 %	2.66%
4	SEI Large Cap Diversified Alpha Fund	75.26%	-0.29%	Deutsche Multi-Asset Conservative Allocation Fund	0.05%	0.40%
5	American Century Strategic Allocation Aggressive Fund	71.88%	0.38%	Deutsche Multi-Asset Global Allocation Fund	0.06%	0.98%

Table IV

Determinants of Centrality: This table reports results of pooled OLS estimation of mutual funds quarterly centrality measures on fund-level characteristics. The sample period is 1980Q1-2014Q4. The dependent variables are the centrality measures (degree non-weighted $Degree^{nw}_{i,t}$, and degree $Degree_{i,t}$) of fund i at time t. The independent variables include past net return of fund i (previous quarter), the logarithm of the number of stocks hold in the portfolio, the logarithm of total net assets (TNA), the logarithm of fund age since inception, fund flow, and turnover ratio of the previous quarter. All specifications include a quarter fixed effect. The standard errors are clustered by funds. The *t-statistics* are in parentheses. All the independent variables are rescaled by their standard deviation for interpretation. Significance at the 1, 5, 10% levels are indicated with *,**,*** respectively.

-	$\mathbf{Degree^{nw}_{i,t}}$	$\mathbf{Degree_{i,t}/100}$
-	[1]	[2]
$Intercept$	355.18***	-526.02***
	(9.05)	(-17.56)
$Flow_{i,t-1}$	-12.49***	-7.22***
	(-5.57)	(-8.46)
$log(Age_{i,t-1})$	24.25***	3.22*
	(5.11)	(1.68)
$log(numStocks_{i,t})$	119.96***	109.21***
	(20.91)	(19.81)
$log(TNA_{i,t-1})$	24.33***	2.06
	(4.35)	(0.89)
$Return_{i,t-1}$	-22.19***	-11.28***
	(-9.02)	(-12.25)
$Turnover_{i,t-1}$	7.21**	-7.67***
	(2.06)	(-5.15)
Fixed effects	Quart.	Quart.
Adj-R^2	68.92%	51.44%
Nb Obs.	92,448	92,448

Table V

Mutual Fund Performance Regressions: This table reports results of pooled OLS estimation of mutual funds quarterly net returns on fund-level centrality measures. The sample period is 1980Q1-2014Q4. The dependent variable is the net return of fund i at time t in percentage and the independent variables include the centrality measures (degree non-weighted $Degree^{nw}_{i,t-1}$, and degree $Degree_{i,t-1}$) of fund i at time $t-1$, past net return of fund i and Carhart four factors alpha (previous quarter), the fund-level total style drift measure (*Styledrift*) as calculated in Wermers (2012) (which stops in 2012Q2), and controls (the logarithm of the number of stocks hold in the portfolio, the logarithm of total net assets (TNA), the logarithm of fund age since inception, fund flow, expense ratio of the previous year, turnover ratio of the previous quarter). All specifications include a quarter fixed effect (Quart.) and a style dummy when specified (style, investment objective code). The standard errors are clustered by fund. The *t-statistics* are in parentheses. All the independent variables are rescaled by their standard deviation for interpretation. Significance at the 1, 5, 10% levels are indicated with *,**,*** respectively.

-	[1]	[2]	[4]	[5]	[7]	[8]	[10]	[11]
$Intercept$	0.550***	0.114*	0.557***	0.120*	0.663***	0.121	0.597***	0.360***
	(9.45)	(1.77)	(12.32)	(1.82)	(9.88)	(1.59)	(11.60)	(5.66)
$\mathbf{Degree^{nw}_{i,t-1}}$	-0.133***		-0.137***		-0.165***		-0.033*	
	(-11.39)		(-11.54)		(-11.90)		(-1.69)	
$\mathbf{Degree_{i,t-1}}$		-0.116***		-0.116***		-0.138***		-0.066***
		(-11.25)		(-11.18)		(-10.98)		(-7.09)
$Alpha_{i,t-1}$			0.810	0.698				
			(1.17)	(1.00)				
$Styledrift_{i,t-1}$					-0.015**	-0.015**		
					(-2.09)	(-2.01)		
$Expenses_{i,t-1}$	-0.054***	-0.056***	-0.050***	-0.052***	-0.046***	-0.048***	-0.051***	-0.052***
	(-7.04)	(-7.12)	(-6.75)	(-6.89)	(-5.21)	(-5.28)	(-7.25)	(-7.38)
$Flow_{i,t-1}$	0.107***	0.106***	0.103***	0.102***	0.075***	0.074***	0.092***	0.091***
	(8.54)	(8.50)	(8.11)	(8.07)	(5.26)	(5.18)	(8.84)	(8.74)
$log(Age_{i,t-1})$	0.044***	0.040***	0.044***	0.040***	0.026***	0.023***	0.016**	0.015**
	(5.40)	(4.97)	(5.32)	(4.87)	(2.93)	(2.58)	(2.44)	(2.23)
$log(numStocks_{i,t})$	0.073***	0.114***	0.073***	0.114***	0.086***	0.135***	0.034***	0.072***
	(11.05)	(12.33)	(10.84)	(12.09)	(11.05)	(12.67)	(4.21)	(6.86)
$log(TNA_{i,t-1})$	-0.070***	-0.074***	-0.068***	-0.074***	-0.063***	-0.070****	-0.043***	-0.045***
	(-8.57)	(-9.25)	(-8.46)	(-9.18)	(-6.90)	(-7.60)	(-6.33)	(-6.65)
$Return_{i,t-1}$	0.263***	0.263***	0.261***	0.261***	0.168***	0.169***	0.236***	0.235***
	(13.61)	(13.65)	(13.26)	(13.31)	(9.48)	(9.56)	(15.52)	(15.45)
$Turnover_{i,t-1}$	-0.024**	-0.031***	-0.023**	-0.030**	-0.045***	-0.053***	-0.017*	-0.019*
	(-2.10)	(-2.66)	(-1.97)	(-2.51)	(-3.21)	(-3.82)	(-1.66)	(-1.89)
Fixed effects	Quart.	Quart.	Quart.	Quart.	Quart.	Quart.	Quart.	Quart.
Fixed effects							Style	Style
Adj-R^2	76.54%	76.55%	76.56%	76.57%	77.77%	77.78%	83.91%	83.92%
Nb Obs.	92,397	92,397	89,853	89,853	63,318	63,318	92,397	92,397

Table VI

Lagged centrality effect: This table reports results of pooled OLS estimation of mutual funds quarterly net returns on fund-level centrality measures. The sample period is 1980Q1-2014Q4. The dependent variable is the net return of fund i at time t and the independent variables include the centrality measures (degree non-weighted $Degree^{nw}_{i,t-1}$, and degree $Degree_{i,t-1}$) of fund i at time $t-1$ (previous quarter), $t-2$, and $t-3$. Controls include past net return of fund i (previous quarter), the logarithm of the number of stocks hold in the portfolio, the logarithm of total net assets (TNA), the logarithm of fund age since inception, fund flow, expense ratio of the previous year, turnover ratio of the previous quarter. All specifications include a quarter fixed effect. The standard errors are clustered by funds. The *t-statistics* are in parentheses. All the independent variables are rescaled by their standard deviation for interpretation. Significance at the 1, 5, 10% levels are indicated with *,**,*** respectively.

-	[1]	[2]	[3]	[4]	[5]	[6]
$Intercept$	0.870***	0.896***	0.919***	0.743***	0.441***	0.466***
	(32.40)	(32.21)	(32.11)	(12.42)	(7.59)	(7.99)
$\mathbf{Degree^{nw}_{i,t-1}}$	-0.107***	0.008	0.058**	-0.049**		
	(-8.81)	(0.34)	(2.13)	(-2.06)		
$\mathbf{Degree^{nw}_{i,t-2}}$		-0.130***	-0.063**	0.008		
		(-5.16)	(-2.37)	(0.36)		
$\mathbf{Degree^{nw}_{i,t-3}}$			-0.127***	-0.110***		
			(-4.99)	(-5.03)		
$\mathbf{\Delta Degree^{nw}_{i,t-1:t-2}}$					-0.004	
					(-0.51)	
$\mathbf{\Delta Degree^{nw}_{i,t-1:t-3}}$						0.020**
						(2.37)
$Expenses_{i,t-1}$				-0.054***	-0.049***	-0.049***
				(-6.66)	(-6.21)	(-6.23)
$Flow_{i,t-1}$				0.108***	0.111***	0.112***
				(8.30)	(8.84)	(8.92)
$log(Age_{i,t-1})$				0.021***	0.038***	0.035***
				(2.60)	(4.61)	(4.22)
$log(numStocks_{i,t})$				0.068***	0.042***	0.040***
				(10.42)	(6.40)	(6.19)
$log(TNA_{i,t-1})$				-0.059***	-0.073***	-0.073***
				(-7.47)	(-9.03)	(-9.02)
$Return_{i,t-1}$				0.246***	0.270***	0.270***
				(13.15)	(14.01)	(14.02)
$Turnover_{i,t-1}$				-0.032***	-0.026**	-0.026**
				(-2.81)	(-2.25)	(-2.25)
Adj-R^2	75.80%	76.13%	76.60%	77.26%	76.49%	76.50%
Nb Obs.	103,460	100,216	97,046	86,013	92,396	92,395

-	[7]	[8]	[9]	[10]	[11]	[12]
	0.685***	0.686***	0.694***	0.278***	0.437***	0.441***
	(81.75)	(80.90)	(80.80)	(4.17)	(7.53)	(7.59)
$\mathbf{Degree_{i,t-1}}$	-0.041***	0.010	0.005	-0.114***		
	(-6.80)	(0.65)	(0.32)	(-6.27)		
$\mathbf{Degree_{i,t-2}}$		-0.054***	-0.022	0.030		
		(-3.31)	(-1.16)	(1.51)		
$\mathbf{Degree_{i,t-3}}$			-0.030**	-0.034**		
			(-2.11)	(-2.30)		
$\mathbf{\Delta Degree_{i,t-1:t-2}}$					-0.007	
					(-1.31)	
$\mathbf{\Delta Degree_{i,t-1:t-3}}$						-0.002
						(-0.37)
				-0.055***	-0.049***	-0.049***
				(-6.62)	(-6.21)	(-6.22)
				0.108***	0.111***	0.111***
				(8.32)	(8.87)	(8.88)
				0.017**	0.038***	0.038***
				(2.04)	(4.62)	(4.60)
				0.109***	0.042***	0.042***
				(11.56)	(6.49)	(6.37)
				-0.065***	-0.074***	-0.073***
				(-8.25)	(-9.04)	(-9.03)
				0.247***	0.270***	0.270***
				(13.17)	(14.01)	(14.00)
				-0.040***	-0.026**	-0.026**
				(-3.57)	(-2.26)	(-2.26)
	75.78%	76.11%	76.57%	77.25%	76.49%	76.49%
	103,460	100,216	97,046	86,013	92,396	92,395

Table VII
Style Cluster Centrality: This table reports results of pooled OLS estimation of mutual funds quarterly net returns on fund-level centrality measures. The sample period is 1980Q1-2012Q4. The dependent variable is the net return of fund i at time t and the independent variables include the centrality measures (degree non-weighted $StyleDegree^{nw}_{i,t-1}$, and degree $StyleDegree_{i,t-1}$, computed within style cluster networks) of fund i at time $t-1$, past net return of fund i (previous quarter), and controls (the logarithm of the number of stocks hold in the portfolio, the logarithm of total net assets (TNA), the logarithm of fund age since inception, fund flow, expense ratio of the previous year, turnover ratio of the previous quarter). All specifications include a quarter fixed effect. The standard errors are clustered by funds. The *t-statistics* are in parentheses. All the independent variables are rescaled by their standard deviation for interpretation. Significance at the 1, 5, 10% levels are indicated with *,**,*** respectively.

-	[1]	[2]
$Intercept$	0.144	-0.032
	(1.08)	(-0.31)
$\mathbf{StyleDegree^{nw}_{i,t-1}}$	-0.053**	
	(-1.98)	
$\mathbf{StyleDegree_{i,t-1}}$		-0.026*
		(-1.77)
$Expenses_{i,t-1}$	-0.049***	-0.047***
	(-3.83)	(-3.72)
$Flow_{i,t-1}$	0.094***	0.094***
	(4.90)	(4.90)
$log(Age_{i,t-1})$	0.059***	0.059***
	(4.49)	(4.53)
$log(numStocks_{i,t})$	0.042***	0.052***
	(3.74)	(4.51)
$log(TNA_{i,t-1})$	-0.069***	-0.071***
	(-5.22)	(-5.37)
$Return_{i,t-1}$	0.261***	0.262***
	(9.20)	(9.23)
$Turnover_{i,t-1}$	-0.079***	-0.080***
	(-2.79)	(-2.81)
Fixed effects	Quart.	Quart.
Adj-R^2	82.11%	82.10%
Nb Obs.	25,165	25,165

Table VIII

Centrality and Active Share: This table reports results of correlation measures (Pearson, Spearman, and Kendall's rank) for centrality measures and the Active Share measure for the sample 1980Q1-2009Q4 in Panel A to C. In Panel D, I run a pooled OLS regression of mutual funds quarterly net returns on fund level centrality measures controlling for Active Share. The sample period is 1980Q1-2014Q4. The dependent variable is the net return of fund i at time t and the independent variables include the centrality measures (degree non-weighted $Degree^{nw}_{i,t-1}$, and degree $Degree_{i,t-1}$) of fund i at time $t-1$, past net return of fund i (previous quarter), *ActiveShare* as in Cremers and Petajisto (2009) (which stops in 2009Q4), and controls (the logarithm of the number of stocks hold in the portfolio, the logarithm of total net assets (TNA), the logarithm of fund age since inception, fund flow, expense ratio of the previous year, turnover ratio of the previous quarter). All specifications include a quarter fixed effect and we include a style dummy according to the fund's investment objective code when specified. The standard errors are clustered by fund. The *t-statistics* are in parentheses. All the independent variables are rescaled by their standard deviation for interpretation. Significance at the 1, 5, 10% levels are indicated with *,**,*** respectively.

Panel A: Pearson's correlation

-	$Degree^{nw}$	$Degree$	$ActiveShare$
$Degree^{nw}$	1		
$Degree$	0.58	1	
$ActiveShare$	-0.42	-0.75	1

Panel B: Spearman's correlation

-	$Degree^{nw}$	$Degree$	$ActiveShare$
$Degree^{nw}$	1		
$Degree$	0.89	1	
$ActiveShare$	-0.49	-0.76	1

Panel C: Kendall's rank correlation

-	$Degree^{nw}$	$Degree$	$ActiveShare$
$Degree^{nw}$	1		
$Degree$	0.72	1	
$ActiveShare$	-0.34	-0.57	1

Panel D	[1]	[2]	[4]	[5]
Intercept	-0.559***	-0.553***	-0.502***	-0.521***
	(-5.45)	(-6.20)	(-4.49)	(-4.59)
$\mathbf{Degree^{nw}_{i,t-1}}$	-0.020		-0.007	
	(-1.07)		(-0.25)	
$\mathbf{Degree_{i,t-1}}$		-0.055***		-0.039***
		(-3.87)		(-3.23)
$\mathbf{ActiveShare_{i,t-1}}$	0.124***	0.102***	0.118***	0.105***
	(13.59)	(9.26)	(10.85)	(8.96)
$Expenses_{i,t-1}$	-0.042***	-0.042***	-0.041***	-0.042***
	(-4.83)	(-4.90)	(-5.14)	(-5.20)
$Flow_{i,t-1}$	0.107***	0.106***	0.111***	0.111***
	(6.09)	(6.04)	(6.77)	(6.74)
$log(Age_{i,t-1})$	0.035***	0.033***	0.023***	0.022**
	(3.80)	(3.68)	(2.63)	(2.53)
$log(numStocks_{i,t})$	0.082***	0.101***	0.074***	0.092***
	(9.12)	(8.92)	(5.80)	(6.93)
$log(TNA_{i,t-1})$	-0.078***	-0.078***	-0.063***	-0.063***
	(-8.36)	(-8.37)	(-7.17)	(-7.16)
$Return_{i,t-1}$	0.286***	0.285***	0.286***	0.285***
	(12.13)	(12.12)	(12.90)	(12.88)
$Turnover_{i,t-1}$	-0.021*	-0.023*	-0.020*	-0.021*
	(-1.76)	(-1.92)	(-1.81)	(-1.89)
Fixed effects	Quart.	Quart.	Quart.	Quart.
Fixed effects			Style	Style
Adj-R^2	78.66%	78.67%	82.38%	82.38%
Nb Obs.	46,874	46,874	46,874	46,874