Mutual Funds Apart From the Crowd

Nadia Vozlyublennaia Youchang Wu*

June 23, 2017

Abstract

We construct measures of mutual fund uniqueness using cluster analysis of fund returns. We find that fund uniqueness persists over time, and is higher for more actively managed funds. More unique funds charge higher fees, but they do not deliver better net-of-fee performance. Fund uniqueness reduces the sensitivity of fund flows to past performance and increases performance persistence, especially when funds perform poorly. Non-unique funds exhibit neither convexity in the flow-performance relation nor performance persistence. Our results suggest that unique funds are better able to retain investors after poor performance, which may in turn increase the persistence of poor performance.

JEL codes: G23, G34

Key Words: mutual fund, fund flow, fund performance, cluster analysis

*Nadia Vozlyublennaia is at the U.S. Securities and Exchange Commission. Email: vozlyublennaian@sec.gov. Youchang Wu is at the Lundquist College of Business, University of Oregon. Email: ywu2@uoregon.edu; Phone: 541-3465126. We are grateful for comments and suggestions from Geoffrey Booth, John Chalmers, Hsiu-lang Chen, Diane Del Guercio, Roberto Gutierrez, Neal Stoughton, Zheng Sun, Lu Zheng, and seminar participants at the Securities and Exchange Commission, University of Oregon, University of Arizona, the 2016 China International Conference in Finance, and the 2017 Midwest Finance Association conference.

Mutual funds compete on both the asset and the liability sides of their balance sheets. On the asset side, they compete for profitable investment opportunities to generate high returns. On the liability side, they compete for fund flows to grow their assets under management. According to the 2016 Investment Company Fact Book published by the Investment Company Institute, there were 9,521 U.S.-registered open-end funds managing a total of $15.65 trillion assets at the end of the year 2015. Given the large number of funds in the industry, the competition on both fronts is intense.

The theory of industrial organization often suggests that the best way to win competition is to avoid competition, i.e., to establish a quasi-monopoly by creating products and services that are hard for competitors to mimic.[1] This implies that fund managers have an incentive to deviate from the crowd and employ innovative and unique investment strategies to escape competition. What are the characteristics of funds with more unique investment strategies? Does higher uniqueness allow funds to charge higher fees? How does fund uniqueness change the elasticity of investor demand for a mutual fund with respect to past performance? And how is it related to fund performance and the persistence of fund performance? These questions are fundamentally important for our understanding of the equilibrium and dynamics of the asset management industry.

To explore these questions, we develop a measure of mutual fund uniqueness using a two-stage cluster analysis of fund returns. We first use partitioning cluster analysis (PCA) to classify funds into different style groups, and then use hierarchical cluster analysis (HCA) to determine the number of steps it takes to separate one fund from other funds in the same style group. The more steps it takes, the lower is fund uniqueness. Based on this intuitive measure, we examine how fund uniqueness is related to fees and expenses, the response of investors to fund performance, as well as the persistence of fund performance. Using a sample of actively managed U.S. domestic equity funds, we find four main results:

First, fund uniqueness is persistent over time, and is generally higher for funds that are more actively managed. In addition, only 30% of the total variation in the return-based uniqueness

[1]See, for example, Aghion, Harris, Howitt, and Vickers (2001) and Aghion, Bloom, Blundell, Griffith, and Howitt (2005).

measures can be explained by the uniqueness of portfolio holdings.

Second, more unique funds have higher management fee and total expense ratios. In a univariate setup, as the uniqueness index increases from the 10th to the 90th percentile, the annual management fee and total expense ratios increase by 26% and 29% from their corresponding sample means, respectively. The positive relation between fees/expenses and fund uniqueness remains statistically and economically significant even after we control for a full set of fund characteristics to account for the potentially higher costs of more unique funds.

Third, fund uniqueness reduces the sensitivity of fund flows to fund performance, especially when performance is poor. Thus, it contributes to the convexity of the flow-performance relation. For a fund with alpha around zero, as the uniqueness index increases from the 10th to the 90th percentile, the sensitivity of fund flows to past performance drops by 43%. For the least unique funds, the response of fund flows to good and bad performance is largely symmetric, suggesting that the well-documented convexity in the flow-performance relation is driven by funds with a high degree of uniqueness.

Fourth, fund uniqueness is not directly related to fund performance, but increases the persistence of fund performance, especially the persistence of poor performance. Among the least unique funds, there is no persistence in performance. However, if the uniqueness index is one standard deviation above the mean, funds with a past four-factor alpha one standard deviation above the mean outperform those with a past alpha one standard deviation below the mean by 190 basis points over the next 12 months. These results suggest that mutual fund performance persistence found in the literature (for example, Brown and Goetzmann (1995) and Carhart (1997)) is mainly driven by high-uniqueness funds.

These results are largely the same when we measure fund uniqueness by an alternative index derived from a sequence of partitioning cluster analysis of fund returns. This alternative measure is based on the number of peers a fund has at various levels of classification granularity. Furthermore, we find similar, although slightly weaker, results when fund uniqueness is measured by a holding-based index.

Our study contributes to the literature on two fronts. First, we contribute to the understanding of the equilibrium in the mutual fund industry. Our findings of high management fees and lower sensitivity of fund flows to poor performance among more unique funds suggest that unique funds have stronger market power in setting fees, and that they are better able to retain investors after poor performance. The lack of a significant relation between fund uniqueness and net-of-fee performance further suggests that benefits from unique investment strategy, if any, are captured by fund managers instead of investors. These findings are potentially important for active fund managers, which face increasingly intensive competition from low-fee passive funds such as index funds and exchange-traded funds. They suggest that pursuing unique investment strategies may be one way for active funds to make their fees sustainable. In addition, these findings provide useful guidance for mutual fund investors on making investment decisions.

To the extent that performance evaluation is generally more difficult for funds with more unique return profiles, our finding of a positive relation between fund uniqueness and the convexity in the flow-performance relation provides support for the model of Huang, Wei, and Yan (2007), which predicts that higher participation costs lead to a higher convexity. Furthermore, our findings of lower flow sensitivity to performance and higher performance persistence among more unique funds are consistent with the Berk and Green (2004) equilibrium of the mutual fund industry, in which the performance-chasing flows act as an equilibrating force that eliminates performance persistence. Overall, our study demonstrates rich interaction between the asset and the liability sides of mutual fund balance sheets.

Second, we develop two return-based measures of mutual fund uniqueness, and show that fund uniqueness is a persistent fund characteristic closely related to mutual fund fees, fund flows, and fund performance. Our economically intuitive measures of uniqueness are well-founded on established statistical tools of cluster analysis. They can be applied to many other settings in which quantifying uniqueness is important, such as the study of product market competition. In the literature, PCA is used by Brown and Goetzmann (1997), Haslem and Scheraga (2001), Pattarina, Paterlinib, and Minerva (2004), and Sun, Wang, and Zheng (2012) to obtain return-based classification of mutual funds or hedge funds. HCA is used by Sun (2015) to analyze institutional holdings. None of these studies uses cluster analysis to characterize fund uniqueness.

Our paper is most closely related to the study of Sun, Wang, and Zheng (2012). They measure the distinctiveness of a fund's investment strategy by one minus the correlation of a fund's return with the average return of its style group, which they call Strategy Distinctiveness Index (SDI). While our measures of fund uniqueness are positively related to SDI, they differ from it by using the information in the entire hierarchy of the cluster structure on the return space. According to our measures, a fund's uniqueness is not only determined by its distance to the peer group mean, but also by its distances to other sub-groups within the peer group. More importantly, the focus of our paper is very different. Sun, Wang, and Zheng (2012) concentrate on the relation between hedge fund strategy distinctiveness and net-of-fee returns to investors, while we are interested in how benefits from mutual fund uniqueness, if any, are split between investors and fund managers, and how uniqueness alters the fund flow-performance sensitivity and performance persistence. Our findings are very different as well. They show that hedge funds with more distinctive strategies deliver better net-of-fee returns to investors. In contrast, we find that more unique mutual funds charge higher fees, but do not generate superior net performance. These different findings are likely due to the different market structures of the hedge fund and mutual fund industries. Compared to mutual fund investors, hedge fund investors are larger, more sophisticated, and face more contractual restrictions on share redemptions. They are more likely to be able to extract surplus generated by managerial skill. In contrast, the provision of capital in the mutual funds is much more competitive, which gives fund managers more market power.

Hoberg, Kumar, and Prabhala (2016) propose a measure of a fund's competition environment based on the number of funds with similar portfolio holdings (in terms of size, book-to-market, and momentum characteristics), and find that performance is persistent only when funds face less competition. Our study differs from theirs in two important aspects. First, they focus on the effect of competition on mutual fund return generation, while we consider the effects of fund uniqueness on both the asset (performance) and the liability (fund flow) sides of mutual fund balance sheets, as well as on mutual fund fees. Our results suggest that the higher performance persistence of unique funds is more likely due to the lack of investor response to poor performance, rather than due to the difficulty in mimicking the superior investment strategies of those funds. Second, they identify the close rivals of a mutual fund using quarterly holdings data, while we measure fund uniqueness based on cluster analysis of fund returns. Not only do our measures require much less data, they also have the advantage of being able to capture information not reflected in quarterly holdings, such as similarities between funds due to similar dynamic trading strategies.[2] While we find similar results when we replace the return-based uniqueness measure by a holding-based uniqueness measure, the results based on the holding-based measure are generally weaker.

Recent studies have found a number of fund characteristics that predict mutual fund performance. Kasperczyk, Sialm, and Zheng (2005) find that investment ability is more evident among managers who hold portfolios concentrated in a few industries. Kacperczyk and Seru (2007) show that skilled managers are less likely to rely on public information for their trades. Cremers and Petajisto (2009) and Petajisto (2013) find that funds deviating more from their benchmarks, i.e., funds with high Active Share, outperform closet indexers. Amihud and Goyenko (2013) find that funds with lower R-squared, obtained from a regression of fund returns on the benchmark factors, have better future performance. Our measures of fund uniqueness are positively correlated with Active Share and negatively correlated with R-squared, suggesting a positive connection between the deviation from the crowd and the deviation from passive benchmarks. Importantly, while funds with high Active Share, high SDI, and low R-squared are found to have superior future performance, our measures of fund uniqueness are not associated with future good or bad performance. Instead, they are associated with the persistence of fund performance, and the response of flows to performance. This suggests that uniqueness per se is not an indicator of good managerial skill. Furthermore, our results are obtained after we control for Active Share, R-squared, SDI, as well as other

[2]Kacperczyk, Sialm, and Zheng (2008) show that there is substantial cross-sectional variation in the gap between the fund return and the return on a hypothetical portfolio that invests in the previously disclosed fund holdings.

potential determinants of fund expenses, fund performance, and fund flows.

The rest of the paper is structured as follows. Section 1 outlines the hypotheses and describes our method of measuring uniqueness. Section 2 describes the data and summary statistics. We examine the determinants of fund uniqueness in Section 3, analyze how fund uniqueness is related to fees and expenses in Section 4, investigate the effect of fund uniqueness on the flow-performance relation in Section 5, and investigate the effect of fund uniqueness on performance persistence in Section 6. We describe several robustness tests in Section 7, and conclude in Section 8.

1 Hypotheses and Methodology

In this section, we first outline several hypotheses about mutual funds that deviate from the crowd, then describe our measures of fund uniqueness.

1.1 Hypotheses

Our first hypothesis is relatively straightforward. We expect more unique funds to charge higher fees, for at least three reasons. First, designing and implementing unique investment strategies is likely to be more costly than simply following the crowd. Second, unique funds, by definition, have no close substitutes. They may offer some unique return profiles preferred by certain investors, which gives them stronger market power in setting the fees. Third, some unique funds may be able to generate higher returns before fees, and this may allow unique funds in general to charge higher fees, as investors are unable to distinguish between the unique funds with and without skills. We therefore state our first hypothesis as follows:

H1. *Management fees and total expense ratios are higher for funds that are more unique.*

Fund uniqueness is likely to affect how investor demand for mutual fund shares responds to past performance. We expect fund flows to be less responsive to the past performance of unique funds. The reasons are as follows. First, unique funds are likely to be more difficult to evaluate than "mainstream" funds. Such funds do not have close peers that can be used as a benchmark to filter out noise in their performance. They may be exposed to risk factors that are not well captured by the standard performance evaluation models. Also, unique funds may focus on assets that are under-explored by other funds, and therefore, are more opaque and less understood. Consequently, investors are likely to be more cautious in interpreting and reacting to their past performance. Second, as we mention, unique funds, by definition, do not have close substitutes. They may offer some return profiles and risk exposures that are not available at other funds. Therefore, investors may continue to hold them even when they underperform due to the difficulty in finding a replacement. Third, unique funds may cater to a specific clientele of investors, which implies that they do not have a broad base of investors who can quickly react to past performance.

The difficulty in evaluating the performance of unique funds may not only reduce the overall sensitivity of fund flows to performance, but also increase the convexity of the flow-performance relation. That is, it may reduce the sensitivity of fund flows to poor performance more than it reduces the sensitivity to good performance. In an optimal learning model with participation costs, Huang, Wei, and Yan (2007) show that funds that are more difficult to analyze have more convex flows, because performance of such funds must be sufficiently high to induce investors to pay the costs to analyze them. In addition, while the lack of close substitutes tends to reduce the sensitivity of fund flows to poor performance, it does not necessarily weaken the response to good performance.

These considerations lead to our second hypothesis:

H2. *Fund uniqueness reduces the overall sensitivity of fund flows to performance, and increases the convexity of the flow-performance relation.*

In the canonical Berk and Green (2004) equilibrium of the mutual fund industry, the response of fund flows to past performance is an equilibrating force that eliminates performance persistence. Outperforming funds attract inflows, and inflows erode future performance due to diseconomies of scale in active management. Poor performance leads to outflows and a smaller fund size, which helps restore performance. Following this logic, if the response of fund flows to performance is weaker for unique funds, then the performance of unique funds should be more persistent. In addition, if fund uniqueness reduces the sensitivity to poor performance more than it reduces the sensitivity to good performance, as we hypothesize, we should expect the amplifying effect of fund uniqueness on performance persistence to be stronger for underperforming than for outperforming funds. These considerations lead to our third hypothesis:

H3. *Fund uniqueness increases the persistence of fund performance, especially the persistence of poor performance.*

Hypothesis **H3** links performance persistence directly to the sensitivity of fund flows to performance in light of the Berk and Green (2004) argument. However, the performance of unique funds may also be more persistent because investment strategies of these funds are difficult to mimic, which allows them to escape the competition and outperform consistently, as argued by Hoberg, Kumar, and Prabhala (2016). If this is indeed the case, we should expect fund uniqueness to be associated with stronger persistence of good performance, which is contrary to what we hypothesize.

1.2 Measuring Fund Uniqueness: Methodology

Our measures of funds' uniqueness are based on cluster analysis of fund returns. Cluster analysis is a machine learning technique of combining data into groups (clusters). It has been successfully applied in many fields such as medicine, biology, computer science, and social science.

There are two types of cluster analysis: partitioning (or nonhierarchical) and hierarchical. Partitioning cluster analysis (PCA) creates various partitions of mutually exclusive clusters with maximum similarity among members of the same cluster and maximum dissimilarity between clusters. The best partition is selected according to some criterion. The number of clusters has to be specified up front as an input parameter. Hierarchical cluster analysis (HCA) is designed to build a hierarchy of clusters by either progressively joining clusters (agglomerative or "bottom-up" strategy), which is the more popular approach, or recursively splitting up clusters (divisive or "top-down" strategy) (see Rokach and Maimon (2005)). The resulting data structure can be represented by a tree called Dendrogram, an example of which is given in Figure 1. Unlike PCA, no input parameters need to be specified up front in HCA.

Various measures of distance can be used, such as Euclidean distance, squared Euclidean distance, Manhattan distance, and Mahalanobis distance. Also, different linkage criteria, which determine the method of measuring the distance between two clusters, have been proposed, such as the average linkage, centroid linkage, complete linkage, and density linkage.[3] We adopt the widely-used centroid method, which defines the distance between two clusters as the squared Euclidean distance between the means (centroids) of the two clusters. The advantage of this method is a straightforward interpretation and robustness to outliers.

Cluster analysis is a natural approach for measuring mutual fund uniqueness. Consider a fund's return stream as an outcome of a particular strategy based on a given selection from all available investment vehicles and/or a method of changing this selection over time. Strategies of different funds may overlap, either due to similar security selection or due to similar methods of changing the selection over time, or both. The degree of overlap in strategies differs across funds. Funds with more overlap exhibit more similar streams of returns, while funds deviating more from the crowd have more unique return streams.

Consider the following example of "bottom-up" strategy, applied to A, B, C, D, E, and F funds in Figure 1. The dendrogram can be built by the following steps. First, pairwise distance measures are computed between every two funds, and the pair with the smallest distance are combined into a cluster. In this case, E and F are combined first to form the cluster EF. Next, an average for the newly formed cluster is computed, which is then considered as a separate element in place of E and F in the subsequent steps. Next, distances are computed across A, B, C, D, and the newly-formed unit EF, and a new cluster BC is formed based on the smallest distance. Next, pairwise distance is calculated across four remaining elements, A, BC, D, and EF, and the cluster DEF is formed. This procedure continues until all elements are merged

[3]See Kaufman and Rousseeuw (2005) for an overview of various basic methods.

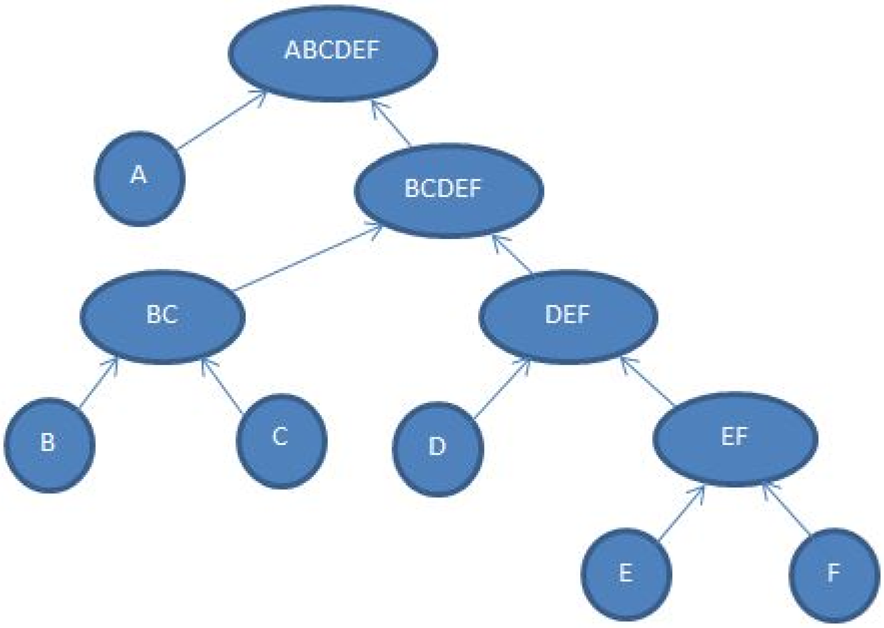

Figure 1: *An example of a Dendrogram in Hierarchical Cluster Analysis.*

into one group, ABCDEF.

By construction, the elements that are merged with other elements later (such as A) are members of a smaller number of clusters, and are more dissimilar (more distant) to other elements in the universe, i.e., more unique. Thus, the structure of the cluster membership, or the total number of clusters each element belongs to, can be viewed as an inverse measure of uniqueness, as it represents the total number of steps needed to separate it from other funds. According to this measure, A is the most unique fund, as it does not belong to any cluster other than the universe. B, C, D come second, as they all belong to three clusters. E and F are the least unique, as they belong to the largest number of clusters (four in this case).

In principle, one can apply HCA to any universe of funds and measure the uniqueness of each member of the universe. However, to simplify the comparison of performance across funds, we restrict our analysis to the commonly-used sample of actively managed U.S. domestic equity funds. Furthermore, we adopt a two-stage procedure, combining both PCA and HCA. In the first stage, we use the K-means PCA to split funds into K style group, so that fund returns are most similar within a group and most dissimilar across groups. This stage requires a pre-specified number of styles K. Following Sun, Wang, and Zheng (2012), we set K to be

10, but our results are not sensitive to this choice. After assigning funds to a given style, we proceed to measure fund uniqueness within the style using HCA. The advantage of this combined approach is that it takes into account the fact that some market sectors are inherently smaller than others. Using only HCA without the first-stage PCA will mechanically treat funds in the smaller market sector as more unique.[4]

We conduct our two-stage cluster analysis of monthly returns at the quarterly frequency, using a rolling window of 36 months. We require funds to have non-missing returns throughout a rolling window. This restriction is usually imposed in HCA, because missing observations make pairwise distances not fully comparable. We standardize returns in each month by the cross-sectional mean and standard deviation, which is customarily done in order to mitigate the influence of outliers. Each fund is then represented by a 1×36 vector of normalized returns. Following the centroid method, the distance between two funds, X_i and Y_i, is the squared Euclidean distance between the two vectors: $d(X_i, Y_i) = \sum_{j=1}^{36} (X_{ij} - Y_{ij})^2$. The distance between two clusters of funds, X and Y, is defined as the squared Euclidean distance between the two 1×36 mean vectors of these clusters. If a fund style resulting from the first stage PCA has 20 or less funds, we exclude it from further analysis.

As mentioned above, we use the total number of clusters that a fund belongs to uncovered by the HCA as an inverse measure of fund uniqueness. We normalize this number to be in the [0,1] interval, and use 1 minus the normalized number as our main measure of fund uniqueness. We call this the return-based uniqueness index.[5] Funds with a uniqueness index of 1 are the most unique, while funds with an index of 0 are the least unique.

To explore the relations between our return-based uniqueness measures and portfolio hold-

[4] One can also do the first-stage style classification based on the CRSP mutual fund style code. However, this approach has two drawbacks. First, the CRSP style code is not fully consistent over our sample period. It is based on the Wiesenberger Objective Code between 1990 and 1993, the Strategic Insight Objective Code between 1993 and 1998, and the Lipper Objective Code since 1998. Second, and more importantly, funds often deviate from their self-designated investment styles (see for example, Brown and Goetzmann (1997), Chan, Chen, and Lakonishok (2002). In particular, Sensoy (2009) finds that almost one-third of actively managed, diversified U.S. equity mutual funds specify a size and value/growth benchmark index in the fund prospectus that does not match the fund's actual style.

[5] Specifically, the index for fund i in a given period is calculated as $1 - \frac{N_i - N_{min}}{N_{max} - N_{min}}$, where N_i is the number of clusters (within its style group) that a fund belongs to, N_{max} and N_{min} are, respectively, the largest and smallest values of N_i across all funds in a style group for a given period.

ings, we also construct a holding-based measure of uniqueness. For this purpose, we examine various characteristics of stocks held by a fund, including size, book-to-market equity ratio, momentum, dividend yield, and the Amihud measure of illiquidity (Amihud (2002)). We first measure these characteritics at the stock level, then aggregate them to the fund level by taking a weighted average across stocks.[6] We then compute a uniqueness index based on the fund-level holding characteristics, following the same procedure that we use for computing our return-based uniqueness index. First, each characteristic is normalized by its cross-sectional mean and standard deviation. Second, funds are assigned into one of the ten style groups that we identify through the first-stage PCA of fund returns.[7] Third, an HCA of the holding characteristics is performed for each of the ten style groups. Fourth, a uniqueness index is computed for each fund based on the normalized number of clusters that the fund belongs to. We repeat this procedure each quarter to get a time series of the holding-based uniqueness index for each fund.

2 Data and Summary Statistics

2.1 Data

We use a sample of actively-managed domestic equity funds in the Center for Research in Security Prices (CRSP) Survivor-Bias-Free US Mutual Fund Database for our analysis. The database contains information about mutual fund historical returns, total net asset values (TNA), expense ratios, and other fund characteristics at the share class level. We use the MFLINKS database from the Wharton Research Data Services (WRDS) to link different share classes of the same fund, and use the CRSP fund objective code to identify domestic equity

[6] At the stock level, size is measured by the market capitalization (stock price times common shares outstanding). Book-to-market ratio is the ratio of the book value of common equity to market capitalization. Momentum is measured by the cumulative return over prior eleven months ending in month $t-2$. Dividend yield is the annual dividend per share divided by price per share. Amihud measure is the ratio of the absolute value of daily return to the dollar value of daily trading volume, scaled by 10^6 and averaged over a quarter. Each measure is winsorized at the 1st and the 99th percentile except the size. At the fund level, size is the natural logarithm of the weighted average of market capitalization across stocks. The other characteristics are weighted averages across stocks. Each stock is weighted by its weight in the fund portfolio.

[7] To maintained the consistency, a fund is assigned into the same style group for both the return- and holding-based analysis.

funds (the first two letters of *crsp_obj_cd* being "ED").

We drop ETFs, variable annuity funds, and index funds (identified either by the CRSP index fund flag, or the word "Index" in the fund name). Since our cluster analysis requires three years of non-missing return observations in each rolling window, and since the number of funds satisfying this requirement is small before 1991 (mainly due to missing TNA data needed for calculating returns at the fund level), we conduct our analysis on a sample that begins in January 1991 and ends in June 2014.

We combine the quarterly fund summary file with the monthly return and TNA files, and aggregate all share class level information to the fund level. A fund's TNA is the sum across all its share classes. Fund returns are the average returns across share classes weighted by the lagged TNA of each class. Other fund level variables, such as expense and management fee ratios, front-end and back-end loads, are averages across share classes weighted by the contemporaneous TNAs. To mitigate the incubation bias documented by Evans (2010), we include in our sample only funds whose TNA measured by the year 2009 dollar value has reached $10 million (the fund remains in the sample even if its TNA subsequently drops below this threshold to avoid selection bias). Furthermore, we drop funds with less than three years of data. Our final sample includes a total of 3,519 funds, with an average of 1,784 funds in a given quarter.[8]

To collect information about the holdings, we use the MFLINKS to link the CRSP database to the Thomson Reuters Mutual Fund Holdings database, which provides stock holdings at the quarterly frequency. We then use the CUSIP to link each stock in the holdings database to the CRSP/Compustat Merged Database, which allows us to compute various stock characteristics.

2.2 Summary statistics

Panel A of Table 1 presents the mean, standard deviation, and various percentiles of fund characteristics and performance at the fund-quarter level. *Ncluster* is the number of clusters

[8] Based on the last observation of the CRSP classification for each fund, our sample consists of 1,297 Growth Funds, 615 Growth and Income Funds, 125 Equity Income Funds, 540 Small Cap Funds, 346 Mid-Cap Funds, 39 Micro-Cap funds, 496 Sector Funds, and 61 Long/Short Funds.

that a fund belongs to in a given period (within its style group identified by the first-stage PCA) according to the HCA. It shows a large dispersion across funds, with a mean of 121 and a standard deviation of 84. *Uindex* is the return-based uniqueness index, our measure of fund uniqueness, which is equal to 1 minus the normalized *Ncluster*. Although normalized to be in the [0, 1] interval, it is not strictly evenly distributed over this interval. Both the mean (0.45) and the median (0.43) are below 0.5, suggesting a higher density at the lower end of the distribution. According to the *Uindex*, the three most unique funds at the end of our sample period are Fairholme Fund, Berkshire Focus Fund, and Wells Fargo Advantage Small Cap Value Fund. Interestingly, the investment adviser of the Fairholme Fund, Fairholme Capital Management, L.L.C., includes *"Ignore the crowd"* as a part of its company logo, suggesting that it embraces uniqueness as the defining feature of its investment strategies.[9]

Uindex(HLD) is the holding-based uniqueness index, obtained by conducting an HCA of the characteristics of stocks held by a fund, including size, book-to-market ratio, momentum, dividend yield, and the Amihud measure of illiquidity (Amihud (2002)), using quarterly holdings data. Compared to the return-based uniqueness measure, *Uindex(HLD)* has a lower mean and a smaller standard deviation. *SDI* is the Strategy Distinctiveness Index computed following the method of Sun, Wang, and Zheng (2012). This index is defined as one minus the correlation of a fund's return with the average return of all funds in its style group identified by the first-stage PCA. As with the *Uindex*, we compute the SDI also using a rolling window of 36 months. Its value ranges from 0 to 1.85. However, the mean and median are only 0.08 and 0.05, respectively, which are substantially lower than the mean (0.32) and median (0.29) values of SDI estimated by Sun, Wang, and Zheng (2012) for their hedge fund sample, suggesting that compared to hedge funds, mutual fund returns are more correlated within the style group. *R2* and *Alpha* are, respectively, the *R*-squared and alpha estimated from the Carhart (1997) four-factor model over a rolling window of 36 months. Our estimates suggest that the four common

[9]On the front page of Fairholme's website, it says, "How does Fairholme stand apart? We maintain that value inevitably attracts price. We swing big at our best ideas. Dips and downturns are our double espressos. If it's not ridiculously cheap it's not our cup of tea. And when the crowd stampedes left, we advance right – with courage of conviction. In short, we *Ignore the crowd*." Berkshire Focus Fund is a large-cap growth fund according to the Lipper fund classification, but its investment is highly concentrated in the electronic technology industry, with an average annual portfolio turnover rate of 732% from 2009 to 2013.

factors explain on average 87% of mutual fund return variation, and the average annualized fund alpha is -0.72%. These results are similar to those reported by Amihud and Goyenko (2013). *ActiveShare* is the Active Share measure developed by Cremers and Petajisto (2009), which measures the deviation of a fund's portfolio from its benchmark index portfolio.[10] The average Active Share for our sample of actively managed mutual funds is 0.80.

3 Characteristics of High-Uniqueness Funds

To determine what makes a fund stand out from the crowd, we first investigate the correlations of our measure of fund uniqueness and other fund characteristics. Panel A of Table 1 presents the correlation matrix. It shows that our return-based uniqueness index, *Uindex*, is positively correlated with the holding-based index, *Uindex(HLD)*, with a correlation coefficient of 0.48. This suggests that the uniqueness of a fund's return profile is partly generated by the uniqueness of its portfolio holdings. *Uindex* is also positively correlated with *SDI* and Active Share, and negatively correlated with R-squared (*R2*), suggesting that our measure of fund uniqueness captures funds deviating more from their benchmark portfolios, the mean of their style group, and common risk factors. Notably, the correlation between *SDI* and *R2* is -0.73, suggesting that funds deviating more from the mean of their style group also deviate more from common risk factors.

In addition, *Uindex* is positively correlated with the total expense ratio, volatility of fund returns, portfolio turnover, and portfolio concentration measured by the Herfindahl-Hirschman Index, and negatively correlated with fund size, fund age, the institutional share of fund assets, and the average market capitalization of stocks held in the portfolio. These results indicate

[10] Following Cremers and Petajisto (2009), we compute Active Share of each fund relative to 19 stock indexes using mutual fund quarterly holdings data from Thomson-Reuters, and use the minimum of the 19 resulting measures as a fund's Active Share. For 17 of the 19 indexes, we use the corresponding iShares ETF to represent the index portfolio. The remaining two indexes, Wilshire 5000 and Wilshire 4500, for which an iShare tracking portfolio is not available, are represented by the index funds Wilshire 5000 Index Portfolio and Fidelity Spartan Extended Market Index Fund, respectively. We compute Active Share starting in the last quarter of 2000, since holdings data for many of these benchmark index funds are not available before that time. For the earlier quarters, we use the data downloaded from Antti Petajisto's website, whom we thank for making the data available. We compare Active Share computed using our method and the item *activeshare_min* in Petajisto's dataset for the overlapping time periods (from the last quarter of 2000 to the 3rd quarter of 2009). The coefficient of correlation between them is 0.99.

that unique funds are more actively managed. They trade more, charge more, and deviate more from their benchmarks. They tend to hold smaller stocks and have a more concentrated portfolio. In addition, they are younger, smaller, and less likely to attract inflows through institutional share classes.[11]

We then investigate the determinants of fund uniqueness using regression analysis based on annual data. The results are reported in Table 3. In Column (1), we examine the persistence of *Uindex* by running a univariate regression. Since *Uindex* is measured over a rolling window of 36 months, we regress *Uindex* on its three-year, instead of one-year, lag to avoid an overlap of the measurement period. Also, since we are interested in the fraction of total variation in our uniqueness measure that can be explained by its own lag, we do not control for any fixed effects. Following Petersen (2009), we report t-statistics based on standard errors clustered by both fund and year. The result shows that uniqueness is a persistent fund characteristic. The coefficient on $Uindex_{t\text{-}3}$ is highly significant, both economically and statistically, with a point estimate of 0.575. The R-squared is 0.334, suggesting that one-third of the variation in a fund's uniqueness can be explained by its uniqueness three years ago.

Column (2) of Table 3 shows the extent to which the return-based uniqueness measure can be explained by the uniqueness of a fund's stock holdings. In addition to the concurrent *Uindex(HLD)*, we also add its two lags as the explanatory variables. As in model (1), we do not control for any fixed effects, but account for clustering of errors by both fund and year. The coefficients on all three terms of *Uindex(HLD)* are highly significant. However, together they explain only about 30% of the total variation in either *Uindex*: the R-squared is 0.303. This suggests that while the return-based uniqueness measure can be partly explained by the uniqueness of holdings, its variation is mainly driven by dynamic portfolio adjustments not fully revealed in the holdings data.

The last two columns of Table 3 show how our measures of fund uniqueness are related to other fund characteristics. All independent variables are measured with a one-year lag, and

[11]Many institutional share classes in the CRSP database are available to retail investors through employer-sponsored retirement plans. Therefore, they may not represent well the behavior of institutional investors.

the models control for both the fund style and the year fixed effects.[12] Consistent with the pairwise correlations reported in Panel B of Table 1, our return-based uniqueness measure is positively correlated with Active Share and SDI, and negatively correlated with the R-squared. In addition, it is positively correlated with the total expense ratio and portfolio concentration (*HHI*), and negatively correlated with fund size, load fees, and the share of institutional share classes in the total fund assets. The negative relation between load fees and fund uniqueness suggests that mutual fund brokers, who are usually compensated by load fees, may have less incentive to invest in unique funds.[13]

In summary, the results in this section show that fund uniqueness, measured by a fund's return profile, is persistent over time. Unique funds are more actively managed. They deviate more from their own benchmarks, the mean of their style group, and common risk factors, and they are less likely to receive money through institutional share classes. While the return-based fund uniqueness measure is partly driven by the uniqueness of holdings, the majority of its variation is not explained by holdings.

4 Uniqueness and Mutual Fund Fees

We now examine the relation between fund uniqueness and fees. We use two measures of asset management fees of mutual funds. The first is the annual management fee ratio, defined as a ratio of management fees to the average TNA over the year. These fees represent a payment to the fund manager (more precisely, to the fund management company) for the portfolio management services. The second is the total expense ratio, which includes management fees as well as fees charged by other service providers such as custodians or bookkeepers. According to **H1**, unique funds should have higher management fee and total expense ratios.

To evaluate the effects of fund uniqueness on management fee and total expense ratios, we regress these ratios, both expressed as a percentage, on the lagged uniqueness measure, *Uindex*, using three alternative model specifications. Since fee ratios are unlikely to change within a

[12]Fund styles are based on the CRSP_obj_code.

[13]See Stoughton, Wu, and Zechner (2011) for a theoretical analysis of the incentive of financial advisers.

given fiscal year, we conduct our tests at the annual frequency, controlling for both fund style and year fixed effects.

The results, reported in Table 4, reveal strong effects of fund uniqueness on management fee and total expense ratios. In the univariate regressions (Columns (1) and (4)), the point estimates of the coefficients on *Uindex* (0.238 and 0.486, respectively) suggest that, as the uniqueness index increases from the 10th percentile (0.09) to the 90th percentile (0.86), the annualized management fee ratio increases by 18 basis points, while the total expense ratio increases by 37 basis points. Given that the means of the management fee and total expense ratio are 70 and 129 basis points, respectively, these numbers imply an increase of 26% and 29%, respectively, from the corresponding sample means.

Part of these large increases in fees and expenses are due to higher costs. As we observe in Section 3, unique funds are more actively managed, and tend to be smaller. In Columns (2) and (4), we control for a vector of fund characteristics to account for the potentially higher costs of more unique funds, including return volatility, fund size and fund age, turnover rate, fraction of institutional share classes in the total TNA, the average size of stock held in the fund, and portfolio concentration. In Columns (3) and (6), we further control for Active Share, SDI, and R-squared. The magnitudes of the fund uniqueness effects on fees and expenses are reduced after we control for these fund characteristics. However, they remain statistically significant at the 1% level.

The effects of other fund characteristics on mutual fund fees are consistent with what one may expect. Funds with higher volatility, higher portfolio turnover rate, and higher portfolio concentration have higher expenses and management fees, as those funds are more actively managed, and therefore, have higher costs. For the same reason, funds deviating more from their own benchmarks or style means, as captured by high Active Share or SDI, also have higher fees and expense ratios. Larger funds and funds receiving a large fraction of investment through institutional share classes have a lower total expense ratio, perhaps because of economies of scale in mutual fund administrative services such as bookkeeping and investor communication. Interestingly, larger funds do not have a lower management fee ratio. In fact, the coefficient on

Log(TNA) is positive in Columns (2) and (3), suggesting a lack of economies of scale in active portfolio management.

To summarize, the results in this section support our hypothesis **H1**. Investors pay higher fees to invest in more unique funds, and the majority of the extra payment goes to fund managers. Part of the higher fees can be explained by higher costs due to more active management. However, fees are still significantly higher for more unique funds even after we control for a full set of observable fund characteristics. This suggests that more unique funds are likely to have stronger market power in setting fees.

5 Uniqueness and Fund Flow Sensitivity to Performance

We now investigate the effects of fund uniqueness on the flow-performance sensitivity, testing Hypothesis **H2**. According to this hypothesis, as fund uniqueness increases, fund flows become less sensitive to fund performance, especially when performance is poor, which leads to a convexity in the flow-performance relation. We conduct several tests of this hypothesis using quarterly data.

5.1 Does fund uniqueness reduce flow sensitivity?

We first test whether fund uniqueness reduces the sensitivity of fund flows to performance, using the following model specification,

$$Flow_{j,t+1} = \beta_0 + \beta_1 Alpha_{j,t} + \beta_2 Uindex_{j,t} + \beta_3 Alpha_{j,t} * Uindex_{j,t} + \sum_{n=1}^{N} \gamma_n Control_{n,j,t} + e_{j,t}, \quad (1)$$

where $Flow_{j,t+1}$ is the average monthly flows of fund i in quarter $j+1$, $Alpha_{j,t}$ and $Uindex_{j,t}$ are fund performance and uniqueness measured over a rolling window of 36 months till the end of quarter t.

Following the standard practice in the literature, we calculate the fund flow in a given month as the TNA growth rate minus the realized fund return, and winsorize it at the 1st and

99th percentiles to mitigate the effects of outliers. Fund performance is measured by alphas estimated from two alternative models: the Carhart (1997) four-factor model, which accounts for the market, size, book-to-market, and momentum risk factors; and the one-factor model (CAPM), which accounts for the market risk.[14] We include the square term of alpha, *AlphaSQ*, as a regressor, to account for the nonlinearity in the flow-performance relation, and control for other potential determinants of fund flows, including fund style and quarter fixed effects. The t-statistics are based on standard errors clustered by both fund and quarter.

If fund flows are less sensitive to the performance of unique funds, as we hypothesize, the coefficient on the interaction term *Alpha*Uindex* should be significantly negative. The first column in Panel A of Table 5 shows the results from our baseline model specification, in which fund performance is measured by the four-factor alpha. The results provide strong support for the our hypothesis. The coefficient on *Alpha* is strongly positive, suggesting that fund flows respond strongly to performance for the least unique funds (i.e., *Uindex* close to zero). The coefficient on the interaction term *Alpha*Uindex*, however, is strongly negative, implying that the sensitivity weakens significantly as funds become more unique. The economic magnitude of the dampening effect of fund uniqueness is very large. The point estimates of the coefficients on *Alpha* (3.333) and *Alpha*Uindex* (-1.847) suggest that, for a fund with an alpha close to zero (so that the effect of the quadratic term *AlphaSQ* on the sensitivity is zero), as *Uindex* increases from the 10th to the 90th percentile (an increase of 0.77), the sensitivity of fund flows to past performance declines by about 43% (=1.847*0.77/3.333).

In Column (2) of the same panel, we extend our baseline model to also allow the sensitivity to past performance to vary with fund return volatility, fund age, and load fees. The results are similar to those in Column (1). Both the point estimate and the statistical significance of the coefficient on the interaction term *Alpha*Uindex* remain largely unchanged. The coefficient on *AlphaSQ* becomes significantly positive, consistent with the previous finding of a convex flow-performance relation in the literature (see, for example, Sirri and Tufano (1998)).

[14]The four-factor model has been commonly used to evaluate mutual fund performance (see, for example, Amihud and Goyenko (2013) and Kacperczyk, Sialm, and Zheng (2008)). However, Berk and van Binsbergen (2016) and Barber, Huang, and Odean (2016) find that investors respond most strongly to the CAPM-alpha.

Columns (4) and (5) report the results for the same tests when past performance is measured by the CAPM alpha. The coefficient on the interaction term *Alpha*Uindex* is negative in both columns, and it is statistically significant in the baseline model (Column (4)), providing further evidence for the dampening effect of fund uniqueness on the sensitivity of fund flows to performance.

Notably, the coefficients on both *Alpha*Vol* and *Alpha*Log(Age)* are significantly negative across the models, suggesting that investors are less responsive when returns are more volatile and when funds are older. These results are consistent with the optimal Bayes learning models of Berk and Green (2004), Dangl, Wu, and Zechner (2008), and Brown and Wu (2016). Investors learn less about managerial ability from past performance when returns are noisier. They also learn less when their uncertainty about managerial ability is lower, as in the case of older funds. As a result, they are less responsive to the performance of high-volatility funds and old funds.

5.2 Uniqueness and the convexity of fund flows

To examine whether the dampening effect of fund uniqueness on flow sensitivity to performance is stronger when funds perform poorly, we conduct two additional tests. First, we add another interaction term, *AlphaSQ*Uindex*, to the model in Columns (2) and (5) of Panel A. This allows us to assess directly the effect of fund uniqueness on the convexity of flow-performance relation. If fund uniqueness reduces the sensitivity to poor performance more than it reduces the sensitivity to good performance, the flow-performance relation should be more convex among more unique funds, and the coefficient on *AlphaSQ*Uindex* should be positive.

The results of this test, again based on two alternative alpha estimates, are reported in Columns (3) and (6) of Panel A in Table 5. The coefficient on *AlphaSQ*Uindex* is significantly positive in Column (6), where performance is measured by the CAPM alpha. This suggests that fund uniqueness significantly increases the convexity of fund flow response to performance. In fact, the coefficient on *AlphaSQ* becomes negative, significant at the 5% level. This implies that for the least unique funds, the flow-performance relation is actually concave. In Column

(3), although the coefficient on *AlphaSQ*Uindex* itself is insignificantly positive, adding this additional interaction term in the regression removes the statistical significance of the coefficient on *AlphaSQ*. This again implies that for the least unique funds, there is no convexity in the flow-performance relation. The convexity observed in Column (2) is, therefore, purely driven by funds that are sufficiently unique.

As an additional test for the effect of fund uniqueness on the convexity of fund flows, we adopt a piecewise linear structure to allow the dampening effect of fund uniqueness on the flow-performance relation to differ for funds with above- and below-median past performance. Specifically, we estimate the following model:

$$\begin{aligned} Flow_{j,t+1} &= \beta_0 + \beta_1 Alpha_{j,t} + \beta_2 Alpha_{j,t} * Uindex_{j,t} + \beta_3 Alpha_{j,t} * Above_{j,t} * Uindex_{j,t} \\ &\quad + \sum_{n=1}^{N} \gamma_n Control_{n,j,t} + e_{j,t}, \end{aligned} \tag{2}$$

where $Above_{j,t}$ is a dummy variable that equals one if $Alpha_{j,t}$ is above the median fund, and zero otherwise.

In this specification, the coefficient on *Alpha*, β_1, captures the response of fund flows to past performance for a fund with a uniqueness index close to zero (least unique funds), the coefficient on *Alpha*Uindex*, β_2, captures the effect of fund uniqueness on the sensitivity of flows to performance below the median, while the coefficient on the three-way interaction term *Alpha*Above*Uindex*, β_3, captures the differential effect of fund uniqueness on the flow-performance sensitivity between the above- and below-median funds. If the dampening effect is stronger for underperforming funds, β_2 should be significantly negative, while β_3 should be significantly positive. We control for each individual variable involved in the three-way interaction term, as well as all their two-way interactions. We include all other explanatory variables in Panel A of Table 5 as controls except *AlphaSQ*, which is dropped because the nonlinearity in the flow-performance relation is captured by a piecewise linear structure of the new model.

Panel B of Table 5 reports the results for this test under two model specifications and two alternative alpha estimates. The results confirm the dampening effects of fund uniqueness on the sensitivity of fund flow to performance, as indicated by the significantly negative coefficient on *Alpha*Uindex* in all columns of the table. They also show that this dampening effect is significantly weaker for funds with above-median performance, as the coefficient on the three-way interaction term *Alpha*Above*Uindex* is significantly positive in all four columns.

Interestingly, results in Panel B also suggest that there is no convexity in the flow-performance relation among the non-unique funds with a *Uindex* close to zero. For those funds, the coefficient on the interaction term *Alpha*Above* picks up the difference in the sensitivities of fund flows to above- and below-median performance. A positive coefficient indicates convexity and a negative coefficient indicates concavity. This coefficient is negative in all four columns in Panel B, and it is statistically significant in the last two columns, when performance is measured by the CAPM alpha. This implies that for the least unique funds, the flow-performance relation is either symmetric or concave, which suggests that the convexity in the flow-performance relation documented in the literature comes mainly from funds with a high degree of uniqueness.

To summarize, consistent with Hypothesis **H2**, we find that fund uniqueness significantly reduces the sensitivity of fund flows to performance. In addition, we find that it reduces the sensitivity to poor performance more than it reduces the sensitivity to good performance, generating a convexity in the flow-performance relation among more unique funds. To the extent that unique funds are more difficult to evaluate, this asymmetry in the dampening effect of fund uniqueness is consistent with the model of Huang, Wei, and Yan (2007), which predicts that higher participation costs lead to a higher convexity in the flow-performance relation. It is also consistent with the idea that unique funds may offer some unique risk profiles, which allows them to retain investors after poor performance due to the lack of close substitutes.

6 Uniqueness and Performance Persistence

Since performance-chasing fund flows are a mechanism to eliminate performance persistence in the mutual fund industry, the lower sensitivity of fund flows to performance and the higher convexity among more unique funds documented in Section 5 imply that performance, especially poor performance, should be more persistent among more unique funds, as we hypothesize in **H3**. We now test this hypothesis.

6.1 Is the performance of unique funds more persistent?

To investigate the effect of fund uniqueness on performance persistence, we estimate the following model using quarterly data:

$$Alpha_{j,t+1} = \beta_0 + \beta_1 Alpha_{j,t} + \beta_2 Uindex_{j,t} + \beta_3 Alpha_{j,t} * Uindex_{j,t} + \sum_{n=1}^{N} \gamma_n Control_{n,j,t} + e_{j,t}, \quad (3)$$

where the dependent variable $Alpha_{j,t+1}$ is the average fund performance over future 12 months, $Alpha_{j,t}$ and $Uindex_{j,t}$ are fund performance and uniqueness measured over a rolling window of 36 months till the end of quarter t. Our main interest is in the coefficient on the interaction term, β_2. If the performance of unique funds is more persistent, this coefficient should be significantly positive.

We use the alpha estimated from the Carhart (1997) four-factor model to measure past performance, and use a battery of measures for the performance in the holding period. In addition to the raw excess return, we use four models to estimate the risk-adjusted performance: the one-factor model (CAPM); the Carhart (1997) four-factor model; the Pastor and Stambaugh (2003) five-factor model, which accounts for liquidity risk in addition to the four risk factors in the Carhart (1997) model; and an alternative five-factor model, which adds an equity trend-following factor to the Carhart (1997) four-factor model. The last model is motivated by the idea that dynamic trading strategies employed by unique funds may generate option-like returns that may be better explained by risk factors that have embedded option features. Fung and Hsieh (2001) show that trend-following risk factors, which they estimate from returns of lookback straddles, explain returns of trend-following hedge funds better than standard asset indices. They construct trend-following factors for bond, equity, currency, and commodity markets. Since our sample focuses on domestic equity funds, we add the equity trend-following factor to the Carhart (1997) model to form an alternative five-factor model.[15]

To estimate alphas in the holding period, we first estimate factor loadings using a rolling window of 36 months, we then use these loadings to compute factor-adjusted fund returns in each of the next 12 months (we require fund returns to be available for at least 6 of the 12 months). Finally, we compute the compounded factor-adjusted return over the holding period and convert it into the average monthly alpha.

As in the fund flow analysis, we run panel regressions with both style and quarter fixed effects using quarterly data, accounting for standard errors clustered by both fund and quarter. Since fund performance is strongly correlated across funds in a given period, we also estimate the models using the Fama-Macbeth procedure (Fama and MacBeth (1973)). That is, we run cross-sectional regressions quarter by quarter and use the time series of the coefficient estimates to infer the mean and standard error of each coefficient. We control for style fixed effects, and estimate the t-statistics based on the Newey-West (Newey and West (1987)) corrected standard errors, which account for autocorrelation in the time series of coefficient estimates up to order three. The results from these two alternative approaches are reported in Panels A and B of Table 6, respectively.

We first test in Column (1) of both panels whether fund uniqueness itself is associated with future performance measured by the four-factor alpha. We control for the usual fund characteristics, but leave out the three variables that are highly correlated with fund uniqueness, i.e., ActiveShare, SDI and R2, to avoid the loss of statistical power due to multicollinearity (adding these variables into the model does not change the results). The coefficient on *Uindex* is indistinguishable from zero in both panels, suggesting that fund uniqueness per se does not predict performance. The results are similar if we use other performance measures. Therefore, investors in general do not benefit from investing in high-uniqueness funds.[16] This is

[15]We thank both Lubos Pastor and David Hsieh for making their factors data available on their websites.

[16]Note that we control for the expense ratio in the regressions, so the coefficient on *Uindex* capture the effect

in sharp contrast to the finding of Sun, Wang, and Zheng (2012) in the hedge fund industry, who show that hedge funds with more distinctive investment strategies deliver better net-of-fee performance. One explanation for this difference is that the provision of capital is much more competitive in the mutual fund industry than in the hedge fund industry. Therefore, mutual fund investors have little power to extract any rent generated by fund managers. The benefits from unique fund strategies, if any, are fully captured by fund managers, as argued by Berk and Green (2004).

The remaining columns of Table 6 show the effect of fund uniqueness on performance persistence. Column (2) reports results from a parsimonious model including only the lagged alpha, the lagged uniqueness index, and their interaction as regressors. They show a strong positive coefficient on the interaction term *Alpha*Uindex*, significant at the 1% and 5% levels, respectively, suggesting that performance is much more persistent among more unique funds. Interestingly, in both Panels A and B, the coefficient on *Alpha* is insignificantly negative. This indicates that for the least unique funds (with *Uindex* close to zero), there is no persistence in performance. If anything, there is a weak tendency of performance reversal. Therefore, persistence in mutual fund performance is purely driven by funds that are sufficiently unique.

In Columns (3) to (7) of both panels, we include the usual fund characteristics as controls. In particular, we allow the degree of performance persistence to vary with fund return volatility and fund age. A comparison between Columns (2) and (3) shows that adding these controls has virtually no effect on the coefficient on *Alpha*Uindex*. In addition, Columns (4) to (7) show that the coefficient on *Alpha*Uindex* is largely unaffected by the measures we use to evaluate the holding period performance. It is consistently positive in all columns, with relatively small variation in magnitude. In addition to fund uniqueness, there is also some evidence showing that performance is less persistent among funds with high return volatility (Panel A) and more persistent among older funds (Panel B).

The strong positive coefficient on *Alpha*Uindex* suggests a significant benefit of investing in unique funds with good past performance. Take the results in Column (2) of Panel B as

of fund uniqueness on performance after holding constant fund expenses. The results are similar if we do not control for the expense ratio.

an example. The estimated coefficients on *Alpha*Uindex* and *Alpha* are 0.234 and -0.002, respectively. This suggests that, if $Uindex$ is equal to 0.73, which is one standard deviation above the mean, funds with a past four-factor alpha one standard deviation (0.47) above the mean outperform those with a past alpha one standard deviation below the mean by 190 basis points over the next 12 months: (0.234*0.73-0.002)*47*2*12= 190. This is an economically significant effect.

While there is some evidence that future fund performance is positively related to Active Share and negatively related to R-squared, the coefficients on these variables are not robust.[17] Two other variables stand out in Table 6 as strong predictors of net performance, the total expense ratio and return volatility, both with a negative coefficient. The fact that total expense ratio is negatively related to the after-fee alpha is well established in the literature. The result that funds with more volatile returns in the past tend to have a lower alpha is consistent with the finding of Jordan and Riley (2015), and is potentially related to the idiosyncratic volatility puzzle in equity returns documented by Ang, Hoddrick, Xing, and Zhang (2006).

6.2 Fund uniqueness and asymmetry in performance persistence

To test for potential differential effects of fund uniqueness on the persistence of good and poor performance, we adopt a piecewise-linear model similar to Model (2):

$$\begin{aligned} Alpha_{j,t+1} &= \beta_0 + \beta_1 Alpha_{j,t} + \beta_2 Alpha_{j,t} * Uindex_{j,t} + \beta_3 Alpha_{j,t} * Above_{j,t} * Uindex_{j,t} \\ &\quad + \sum_{n=1}^{N} \gamma_n Control_{n,j,t} + e_{j,t}, \end{aligned} \tag{4}$$

where $Above_{j,t}$ is a dummy variable that equals one if $Alpha_{j,t}$ is above the median fund, and zero otherwise. The coefficient on the three-way interaction variable, *Alpha*Above*Uindex*, captures any potential asymmetry in the effect of fund uniqueness on performance persistence.

[17]The lack of predictability of fund performance by Active Share, R-squared, and SDI is largely due to the strong correlations between these predictors. When we replace these three variables by their first principal component, which is equal to $0.41*ActiveShare+0.62*SDI-0.66*R2$, we find a strong positive relation between this principal component and future fund performance in our panel regressions, confirming the findings of Cremers and Petajisto (2009) and Amihud and Goyenko (2013). The coefficients on *Uindex* and *Alpha*Uindex* remain largely unchanged under this alternative model specification.

We report in Panel C of Table 6 the results when future performance is measured by either the Carhart (1997) four-factor alpha, the Pastor and Stambaugh (2003) five-factor alpha, or the raw excess return. The results are similar when other performance measures are used. As in Panels A and B, the coefficient on the interaction term *Alpha*Uindex* is significantly positive, suggesting that fund uniqueness increases performance persistence. However, the coefficient on the three-way interaction *Alpha*Above*Uindex* is negative in all six columns, and is statistically significant in the Fama-MacBeth regressions (Columns (4) to (6)). The negative coefficient suggests that fund uniqueness increases the persistence of below-median performance more than it increases the persistence of above-median performance, an asymmetry echoing what is found in our fund flow regressions. Interestingly, none of the coefficients on *Alpha* or *Alpha*Above* is significantly different from zero, suggesting a lack of performance persistence for non-unique funds (with *Uindex* close to zero).

The result that fund uniqueness works primarily to increase the persistence of poor performance instead of good performance is inconsistent with the idea that unique funds employ superior investment strategies that are difficult to mimic. However, it is consistent with our finding that fund uniqueness reduces more significantly the response of fund flows to poor performance. Since unique funds are better able to retain investors after poor performance, the equilibrating downward adjustment of fund size is slowed down. Therefore, unique funds are less disciplined by competition pressures, thus their poor performance persists for a longer time.

In summary, the results in this section show that while fund uniqueness per se does not predict performance, it significantly increases the persistence of fund performance, especially the persistence of poor performance. These results provide support for the hypothesis **H3**. Combined with our finding of a lower sensitivity of fund flows to fund performance, especially to poor performance, these results also provide support for the equilibrating mechanism of the mutual fund industry modeled by Berk and Green (2004).

7 Robustness Checks

We conduct several robustness checks. First, we construct an alternative return-based uniqueness measure. Second, we repeat our tests using the holding-based uniqueness measure. Third, we exclude long-short funds and sectors funds. Fourth, we repeat our analysis using gross returns instead of net-of-fee returns.

7.1 PCA-based uniqueness measure

The uniqueness index described in Section 1.2 considers the number of clusters a fund belows to along its cluster chain; however, it does not account for the size of each cluster. For example, in Figure 1, the measure assigns the same level of uniqueness to funds B, C, and D. One may argue that D is less unique than B and C, because its parent cluster has three funds while the parent cluster of B and C has only two. As a robustness check, we construct an alternative measure of fund uniqueness, which we compute by replacing the second-stage HCA by a sequence of PCA with increasing granularity. That is, for each of the ten style groups resulting from the first-stage PCA, we further divide them into K groups using the K-means PCA, with K=5, 10, ..., 100. A higher K means a finer granularity of classification, as funds are split into more groups. For each K, we count the total number of funds in a fund's cluster, $N_{K,i}$. Since a large $N_{K,i}$ indicates a large number of funds with similar return profiles, $N_{K,i}$ is an inverse measure of the fund's uniqueness at the granularity level K. Obviously, as K increases, $N_{K,i}$ decreases. We average $N_{K,i}$ across all different Ks to get an average $\overline{N}_i$ for each fund i. We normalize $\overline{N}_i$ to be in the [0, 1] interval by converting it into $\frac{\overline{N}_i - \overline{N}_{min}}{\overline{N}_{max} - \overline{N}_{min}}$, where $\overline{N}_{max}$ and $\overline{N}_{min}$ are, respectively, the largest and smallest values of $\overline{N}_i$ across all funds in a fund style group for a given period. We use 1 minus this normalized value as our second return-based measure of fund uniqueness, and denote it by *Uindex(PCA)*. Like the HCA-based *Uindex*, this index also uses information in the overall cluster structure of fund returns. Intuitively, while *Uindex* considers the *length* of a fund's cluster chain, *Uindex(PCA)* considers the *width* of a fund's cluster chain.

Not surprisingly, the two return-based uniqueness indexes are highly correlated, with a correlation coefficient of 0.76, and their correlations with other fund characteristics are largely the same. We rerun our main regressions using the PCA-based *Uindex(PCA)* instead of the HCA-based *Uindex*, and summarize the main results in Panel A of Table 7. The first two columns of each panel in Table 7 replicate the main models in Table 4, the next two columns replicate the main models in Table 5, while the last two columns replicate the main models in Table 6. The effects of fund uniqueness on management fees, total expenses, the flow-performance sensitivity and performance persistence are similar to those reported earlier in the paper, as indicated by the coefficients on *Uindex(PCA)*, *Alpha*Uindex(PCA)*, and *Alpha*Above*Uindex(PCA)*. This suggests that the two return-based uniqueness measures capture similar information.

7.2 Additional tests

In addition, we repeat our tests using the holding-based uniqueness index, *Uindex(HLD)*, and report the results in Panel B. These results are largely similar to those obtained using the return-based uniqueness measures, but they are generally weaker. For example, the coefficient on *Uindex* in the total expense ratio regression in Column (2) and the coefficient on *Alpha*Above*Uindex* in the fund flow regression in Column (4) are only statistically significant at the 10% level, and the coefficient on *Alpha*Uindex* in the performance persistence regression in Column (5) is statistically insignificant. This suggests that, while the holding-based uniqueness measure is correlated with the return-based measures, it does not contain all information reflected in returns, perhaps because it does not adequately capture the patterns in dynamic trading strategies, especially the patterns in interim trading within quarters.

Sector funds and long-short funds tend to be more unique than other funds. As a result, one may wonder whether the effects we identify are mainly driven by these funds. To check this possibility, we redo our analysis excluding all sector and long-short funds from our sample. The main results are summarized in Panel C of Table 7. Excluding those funds reduces the sample size by about 9%, but does not significantly change any of our main results. High uniqueness is associated with higher management fee and total expense ratios, as indicated by the positive coefficient on *Uindex* in Columns (1) and (2). It is also associated with a lower flow-performance sensitivity and higher performance persistence (see the opposite signs of the coefficient on *Alpha*Uindex* in Columns (3) and (5)). Furthermore, both the dampening effect of fund uniqueness on the flow sensitivity and the amplifying effects of fund uniqueness on performance persistence are stronger for underperforming funds than for outperforming funds, as indicated by the opposite signs of the coefficient on *Alpha*Above*Uindex* in Columns (4) and (6).

Finally, we reconstruct our uniqueness measure using the HCA of gross instead of net fund returns, and rerun our tests using this alternative measure. The results are very similar to those reported in Tables 4 to 6. To save space, these results are not tabulated.

8 Conclusion

Based on a two-stage cluster analysis of historical returns, we construct a measure of mutual fund uniqueness. Intuitively, this measure is inversely related to the number of steps needed to separate a fund from other funds. We find that fund uniqueness persists over time, and is usually higher for funds that are more actively managed. More unique funds have higher management fee and total expense ratios, even after we control for a full set of fund characteristics to account for the potentially higher costs of those funds. Fund flows are less sensitive to the performance of unique funds than to the performance of non-unique funds. While fund uniqueness is not directly associated with good or poor performance, it is associated with stronger persistence of fund performance. In addition, both the dampening effect of fund uniqueness on the flow-performance sensitivity and the amplifying effect of fund uniqueness on performance persistence are stronger for underperforming funds. Our findings suggest that unique funds have stronger market power in setting fees, and that they are better able to retain investors after poor performance. They also suggest that both the convexity in the flow-performance relation and performance persistence documented in the literature are driven by funds that are sufficiently unique.

To the extent that performance evaluation is generally more difficult for funds with more unique return profiles, our finding of a positive relation between fund uniqueness and the convexity in the flow-performance relation is consistent with the theoretical prediction that higher participation costs lead to a higher convexity in fund flows. Furthermore, our findings of a lower flow sensitivity to performance and higher performance persistence among more unique funds also provide support for the Berk and Green (2004) equilibrium of the mutual fund industry, in which the performance-chasing flows act as an equilibrating force that eliminates performance persistence. They suggest that higher costs of performance evaluation and the lack of close substitutes reduce the sensitivity of fund flow to performance, especially when the performance is poor, and that this may in turn lead to higher persistence of poor performance. These results demonstrate rich interaction between the asset and the liability sides of mutual fund balance sheets.

Our findings have strong implications for active fund managers, whose fee schedules have been increasingly challenged by low-fee passive funds. Our finding of stronger performance persistence among more unique funds is useful for mutual fund investors. Furthermore, our statistics-based, economically intuitive measures of uniqueness can be applied to many other settings in which quantifying uniqueness is important.

References

Aghion, Philippe, Nick Bloom, Richard Blundell, Rachel Griffith, and Peter Howitt, 2005, Competition and innovation: an inverted-U relationship, *Quarterly Journal of Economics* 120, 701–728.

Aghion, Philippe, Christopher Harris, Peter Howitt, and John Vickers, 2001, Competition, imitation and growth with step-by-step innovation, *Review of Economic Studies* 68, 467–492.

Amihud, Yakov, 2002, Illiquidity and stock returns: Cross-section and time series effects, *Journal of Financial Markets* 5, 3156.

Amihud, Yakov, and Ruslan Goyenko, 2013, Mutual fund's R^2 as predictor of performance, *Review of Financial Studies* 26, 667–694.

Ang, Andrew, Robert J. Hoddrick, Yuhhang Xing, and Xiaoyan Zhang, 2006, The cross-section of volatility and expected returns, *Journal of Finance* 61, 259–299.

Barber, Brad M., Xing Huang, and Terrance Odean, 2016, Which Factors Matter to Investors? Evidence from Mutual Fund Flows, *The Review of Financial Studies* 29, 2600.

Berk, Jonathan, and Richard Green, 2004, Mutual fund flows and performance in rational markets, *Journal of Political Economy* 112, 1269–1295.

Berk, Jonathan B., and Jules H. van Binsbergen, 2016, Assessing asset pricing models using revealed preference, *Journal of Financial Economics* 119, 1–23.

Brown, David P., and Youchang Wu, 2016, Mutual fund flows and cross-fund learning within families, *Journal of Finance* 71, 383–424.

Brown, Stephen J., and William N. Goetzmann, 1995, Performance persistence, *The Journal of Finance* 50, 679–698.

Brown, Steven J., and William N. Goetzmann, 1997, Mutual fund styles, *Journal of Financial Economics* 43, 373–399.

Carhart, Mark, 1997, On persistence in mutual fund performance, *Journal of Finance* 52, 57–82.

Chan, Louis K. C., Hsiu-Lang Chen, and Josef Lakonishok, 2002, On mutual fund investment styles, *Review of Financial Studies* 15, 1407–1437.

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.

Dangl, Thomas, Youchang Wu, and Josef Zechner, 2008, Market discipline and internal governance in the mutual fund industry, *Review of Financial Studies* 21, 2307–2343.

Evans, Richard B., 2010, Mutual Fund Incubation, *Journal of Finance* 65, 1581–1611.

Fama, Eugene F., and James D. MacBeth, 1973, Risk and return: Some empirical tests, *Journal of Political Economy* 81, 607–636.

Fung, William, and David A. Hsieh, 2001, The risk in hedge fund strategies: Theory and evidence from trend followers, *Review of Financial Studies* 14, 313–341.

Haslem, John A., and Carl A. Scheraga, 2001, Morningstar's classification of large-cap mutual fundss, *Journal of Investing* 10, no. 1, 79–89.

Hoberg, Gerard, Nitin Kumar, and Nagpurnanand Prabhala, 2016, Mutual fund competition, managerial skill, and alpha persistence, University of Southern California working paper.

Huang, Jennifer, Kelsey Wei, and Hong Yan, 2007, Participation costs and the sensitivity of fund flows to past performance, *Journal of Finance* 62, 1273–1311.

Jordan, Bradford D., and Timothy B. Riley, 2015, Volatility and mutual fund manager skill, *Journal of Financial Economics* 118, 289–298.

Kacperczyk, Macin, and Amit Seru, 2007, Fund manager use of public information: New evidence on managerial skills, *Journal of Finance* 62, 485–528.

Kacperczyk, Marcin, Clemens Sialm, and Lu Zheng, 2008, Unobserved actions of mutual funds, *Review of Financial Studies* 21, 2379–2416.

Kasperczyk, Marcin, Clemens Sialm, and Lu Zheng, 2005, On the industry concentration of actively managed equity mutual funds, *Journal of Finance* 60, 1983–2011.

Kaufman, Leonard, and Peter J. Rousseeuw, 2005, *Finding groups in data: An introduction to cluster analysis*. (John Wiley & Sons New Jersey).

Newey, Whitney K., and Kenneth D West, 1987, A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance Matrix, *Econometrica* 55, 703–708.

Pastor, Lubos, and Robert F. Stambaugh, 2003, Liquidity risk and expected stock returns, *Journal of Political Economy* 111, 642–685.

Pattarina, Francesco, Sandra Paterlinib, and Tommaso Minerva, 2004, Clustering financial time series: An application to mutual funds style analysis, *Computational Statistics & Data Analysis* 47, 353–372.

Petajisto, Antti, 2013, Active share and mutual fund performance, *Financial Analysts Journal* 69, no. 4, 73–93.

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

Rokach, Lior, and Oded Maimon, 2005, *Clustering methods,* in Lior Rokach and and Oded Maimon (eds.)*, Data mining and knowledge discovery handbook,* 321-352. (Springer New York).

Sensoy, Berk A., 2009, Performance evaluation and self-designated benchmark indexes in the mutual fund industry, *Journal of Financial Economics* 92, 25 – 39.

Sirri, Erik R., and Peter Tufano, 1998, Costly search and mutual fund flows, *Journal of Finance* 53, 1589–1622.

Stoughton, Neal M., Youchang Wu, and Josef Zechner, 2011, Intermediated Investment Management, *Journal of Finance* 66, 947–980.

Sun, Zheng, 2015, Institutional clientele and comovement, UC Irvine working paper.

Sun, Zheng, Ashley Wang, and Lu Zheng, 2012, The road less traveled: Strategy distinctiveness and hedge fund performance, *Review of Financial Studies* 25, 96–143.

Table 1: Summary statistics of mutual funds

This table presents summary statistics of our sample, which covers the actively managed U.S. domestic equity mutual funds from January 1991 to June 2014. The statistics are measured at the fund-quarter level. *Ncluster* is the number of clusters that a fund belongs to within its style group, estimated from the hierarchical cluster analysis (HCA) of fund returns over a rolling window of 36 months. The return-based uniqueness index, *Uindex*, is equal to 1 minus the normalized *Ncluster*. *Uindex(HLD)* is the holding-based uniqueness index, obtained by cluster analysis of characteristics of stocks held by a fund, including size, book-to-market ratio, momentum, dividend yield, and the Amihud measure of illiquidity (Amihud (2002)). *SDI* is the Strategy Distinctiveness Index computed following Sun, Wang, and Zheng (2012). *ActiveShare* is the active share computed following Cremers and Petajisto (2009). *R2* and *Alpha* are, respectively, R-squared and alpha estimated from the Carhart (1997) four-factor model over a rolling window of 36 months. *Flow* is the average monthly flow in a given quarter, calculated as the difference between a fund's TNA growth rate and realized returns. is *Vol*, *IVol*, and *SVol* are total excess return volatility, idiosyncratic volatility and systematic volatility (estimated from the market factor model), respectively, all estimated over a rolling window of 36 months. *Expense* and *MgmtFee* are annual total expense ratio and management fee ratio, respectively. *Load* is the sum of maximum front-end and back-end loads. *InstRatio* is the ratio of assets invested through institutional share classes to the TNA. *Turnover* is the annual portfolio turnover rate. *TNA* is the total net asset value in million dollars. *Age* is the number of years since fund inception. *StockSize* is the natural logarithm of the weighted (by portfolio weights) average of market capitalization of stocks held by the funds. *HHI* is the Herfindahl-Hirschman Index of holding concentration, which equals the sum of the squared weights of all stocks in a fund's portfolio.

	Mean	SD	P(10)	P(25)	P(50)	P(75)	P(90)	N
#Fund per period	1784.05	546.46	900.00	1260.00	1963.50	2277.00	2304.00	94
Ncluster	120.95	83.94	20.00	48.00	109.00	182.00	240.00	115674
Uindex	0.45	0.28	0.09	0.22	0.43	0.68	0.86	115674
Uindex(HLD)	0.36	0.19	0.13	0.22	0.33	0.47	0.61	113722
ActiveShare	0.80	0.15	0.59	0.70	0.84	0.92	0.96	134614
R2	0.87	0.15	0.71	0.85	0.92	0.96	0.97	123232
SDI	0.08	0.10	0.02	0.03	0.05	0.10	0.18	115674
Flow(% p.m.)	0.78	4.84	-2.84	-1.37	-0.21	1.57	5.32	167160
Alpha(% p.m.)	-0.06	0.47	-0.51	-0.27	-0.08	0.13	0.41	123232
Vol(% p.m.)	5.17	2.20	2.67	3.63	4.85	6.22	7.84	148014
IVol(% p.m.)	2.20	1.42	0.85	1.18	1.82	2.78	4.01	115674
SVol(% p.m.)	4.43	1.83	2.23	3.02	4.25	5.61	6.83	115674
Expense(% p.a.)	1.29	0.48	0.77	0.99	1.24	1.53	1.92	165056
MgmtFee(% p.a.)	0.70	0.38	0.27	0.56	0.75	0.91	1.05	133702
Load(%)	1.97	2.14	0.00	0.00	1.00	4.00	5.09	167701
InstRatio	0.20	0.34	0.00	0.00	0.00	0.21	0.94	167701
Turnover(p.a.)	0.89	0.88	0.18	0.34	0.65	1.13	1.81	160254
Log(TNA)	5.31	1.82	2.97	4.00	5.25	6.57	7.67	167240
Log(Age)	2.16	0.97	0.92	1.59	2.22	2.78	3.34	167524
StockSize	9.56	1.65	7.18	8.23	9.84	11.04	11.43	154476
HHI	0.03	0.05	0.01	0.01	0.02	0.03	0.05	158089

Table 2: Correlation matrix

This table shows the correlation matrix of various fund characteristics, estimated using quarterly data. *Uindex* and *Uindex(HLD)* are the uniqueness indexes derived from cluster analysis of fund returns and fund holdings, respectively. See the caption of Table 1 for the details of variable definitions.

	Uindex	Uindex(HLD)	ActiveShare	R2	SDI	Flow	Alpha	Vol	Expense	MgtFee	InstRatio	Turnover	Log(TNA)	Log(Age)	StockSize	HHI
Uindex	1.00															
Uindex(HLD)	0.48	1.00														
ActiveShare	0.49	0.37	1.00													
R2	-0.50	-0.28	-0.37	1.00												
SDI	0.48	0.32	0.22	-0.73	1.00											
Flow	0.04	0.03	0.06	-0.06	0.05	1.00										
Alpha	0.05	0.02	0.09	-0.11	0.11	0.23	1.00									
Vol	0.13	0.05	0.17	-0.02	-0.19	-0.02	0.03	1.00								
Expense	0.32	0.19	0.23	-0.26	0.19	0.03	-0.11	0.19	1.00							
MgtFee	0.22	0.15	0.14	-0.14	0.13	-0.13	0.02	0.11	0.36	1.00						
InstRatio	-0.19	-0.13	-0.17	0.17	-0.15	-0.04	0.00	0.03	-0.28	-0.07	1.00					
Turnover	0.15	0.09	0.07	-0.12	0.09	0.02	-0.09	0.21	0.23	0.11	-0.03	1.00				
Log(TNA)	-0.21	-0.14	-0.18	0.14	-0.11	-0.06	0.18	-0.09	-0.37	0.07	0.01	-0.17	1.00			
Log(Age)	-0.07	-0.06	-0.10	0.05	-0.08	-0.33	-0.04	-0.07	-0.15	0.10	-0.06	-0.10	0.42	1.00		
StockSize	-0.19	-0.22	-0.62	0.18	-0.11	-0.07	-0.00	-0.17	-0.18	-0.20	0.05	-0.06	0.15	0.16	1.00	
HHI	0.12	0.20	0.23	-0.17	0.12	-0.00	-0.03	0.02	-0.01	-0.13	-0.04	-0.01	-0.10	-0.02	0.03	1.00

Table 3: Characteristics of high-uniqueness funds

This table shows the relation between fund uniqueness and fund characteristics. The dependent variable is the return-based uniqueness index (*Uindex*), derived from the two-stage cluster analysis of monthly returns in a rolling-window of 36 months. $Uindex(HLD)_{t}$, $Uindex(HLD)_{t\text{-}1}$, $Uindex(HLD)_{t\text{-}2}$ are uniqueness indexes based on quarterly holdings at the end of the current and two lagged years, respectively. $Uindex_{t\text{-}3}$ is estimated with a three-year lag to avoid an overlap of measurement period with the dependent variable. All other regressors are measured with a one-year lag. See the caption of Table 1 for the details of variable definitions. Results in the first two columns are estimated using pooled regressions without any fixed effects, while whose in the last two columns are estimated with both year and fund style fixed effects. The t-statistics (in parentheses) are based on standard errors clustered by both fund and year. All models are estimated using annual data. Significance at the 1%, 5%, and 10% levels are indicated by ***, **, and *, respectively.

	(1) Uindex	(2) Uindex	(3) Uindex	(4) Uindex
$\text{Uindex}_{\text{t-3}}$	0.575***			
	(29.75)			
$\text{Uindex(HLD)}_{\text{t}}$		0.481***		
		(24.86)		
$\text{Uindex(HLD)}_{\text{t-1}}$		0.245***		
		(11.60)		
$\text{Uindex(HLD)}_{\text{t-2}}$		0.261***		
		(11.00)		
SVol			-0.033***	0.014*
			(-4.07)	(1.92)
IVol			0.066***	0.015
			(4.19)	(1.36)
Log(Age)			0.011**	0.004
			(2.31)	(0.96)
Log(TNA)			-0.013***	-0.008***
			(-4.88)	(-3.78)
Alpha			0.024**	0.000
			(1.99)	(0.02)
Expense			0.081***	0.043***
			(7.12)	(4.87)
Load			-0.007***	-0.005***
			(-4.34)	(-3.53)
InstRatio			-0.093***	-0.051***
			(-9.98)	(-7.32)
Turnover			0.007	0.006
			(1.57)	(1.53)
StockSize			-0.023***	0.039***
			(-4.06)	(6.13)
HHI			2.187***	0.854***
			(3.45)	(3.04)
ActiveShare				0.886***
				(20.61)
R2				-0.422***
				(-3.54)
SDI				0.604***
				(5.04)
Constant	0.181***	0.098***	0.624***	-0.327**
	(16.51)	(9.45)	(12.14)	(-2.31)
Observations	20364	22146	24982	22938
R^2	0.334	0.303	0.326	0.474

Table 4: Fund uniqueness and fees

This table presents the regression results for management fee (Columns (1) to (3)) and total expense ratios (Columns (4) to (6)). *Uindex* is the return-based uniqueness index derived from cluster analysis. See the caption of Table 1 for the details of all other variable definitions. All models are estimated using annual data, with independent variables lagged by one year. The models are estimated using panel regressions with both year and fund style fixed effects. The t-statistics (in parentheses) are based on standard errors clustered by both fund and year. Significance at the 1%, 5%, and 10% levels are indicated by ***, **, and *, respectively.

	(1) MgtFee	(2) MgtFee	(3) MgtFee	(4) Expense	(5) Expense	(6) Expense
Uindex	0.238***	0.116***	0.048***	0.486***	0.148***	0.082***
	(11.31)	(7.77)	(3.14)	(18.30)	(6.20)	(3.82)
Vol		0.007**	0.007**		0.031***	0.033***
		(2.26)	(2.06)		(7.22)	(6.83)
Log(Age)		-0.044***	-0.045***		-0.034***	-0.041***
		(-5.79)	(-5.72)		(-2.87)	(-3.44)
Log(TNA)		0.007	0.009**		-0.079***	-0.075***
		(1.61)	(2.09)		(-18.30)	(-17.95)
Turnover		0.010**	0.011**		0.031***	0.033***
		(2.16)	(2.33)		(5.04)	(5.10)
InstRatio		-0.041***	-0.031**		-0.343***	-0.339***
		(-3.08)	(-2.42)		(-15.29)	(-14.88)
StockSize		-0.036***	-0.018***		-0.020***	-0.001
		(-5.98)	(-2.82)		(-3.25)	(-0.09)
HHI		0.442***	0.242		0.432**	0.183
		(3.12)	(1.57)		(2.37)	(0.93)
ActiveShare			0.318***			0.248***
			(5.75)			(4.61)
R2			0.064			0.000
			(1.10)			(0.00)
SDI			0.181***			0.159**
			(3.01)			(2.00)
Constant	0.608***	1.192***	0.759***	1.041***	1.864***	1.498***
	(62.13)	(17.98)	(7.00)	(84.79)	(26.87)	(12.43)
Observations	26129	24530	23596	29150	27466	25287
R^2	0.051	0.108	0.123	0.132	0.348	0.358

Table 5: Fund uniqueness and fund flow sensitivity to performance

This table presents regression results on the effects of fund uniqueness on the sensitivity of fund flows to performance. The dependent variable, *Flow*, is the average monthly flow within a quarter. *Uindex* is the return-based uniqueness index derived from cluster analysis. *Alpha* is the Carhart (1997) four-factor alpha in the first half of each panel, and is the CAPM alpha in the second half of each panel. Both alphas are estimated over a rolling-window of 36 months. *Above* is a dummy variable equal 1 if *Alpha* is above the contemporaneous sample median, and 0 otherwise. *AlphaSQ* is the square of *Alpha*. See the caption of Table 1 for the definitions of other variables. All independent variables are lagged by one quarter, all models are estimated using quarterly data with both quarter and fund style fixed effects. The t-statistics (in parentheses) are based on standard errors clustered by both fund and quarter. The t-statistics for the usual controls are omitted to save space. Significance at the 1%, 5%, and 10% levels are indicated by ***, **, and *, respectively.

A. Uniqueness and the flow-performance sensitivity

	Four-factor alpha			One-factor alpha		
	(1)	(2)	(3)	(4)	(5)	(6)
	Flow	Flow	Flow	Flow	Flow	Flow
Alpha*Uindex	-1.847***	-1.597***	-1.585***	-0.530**	-0.217	-0.354
	(-6.67)	(-6.22)	(-6.37)	(-2.42)	(-0.99)	(-1.55)
AlphaSQ*Uindex			0.357			0.459***
			(1.34)			(2.96)
AlphaSQ	0.046	0.181**	-0.100	-0.045	0.047	-0.288**
	(0.54)	(2.29)	(-0.53)	(-0.63)	(0.65)	(-2.41)
Uindex	0.175	0.164	0.113	0.415***	0.441***	0.309**
	(1.45)	(1.41)	(0.97)	(3.44)	(3.59)	(2.36)
Alpha	3.333***	6.356***	6.311***	2.045***	4.370***	4.423***
	(13.73)	(15.57)	(15.91)	(10.91)	(10.76)	(11.46)
Alpha*Vol		-0.300***	-0.298***		-0.288***	-0.285***
		(-10.58)	(-10.76)		(-9.59)	(-9.54)
Alpha*Log(Age)		-0.447***	-0.437***		-0.246***	-0.236***
		(-4.94)	(-4.97)		(-3.07)	(-3.08)
Alpha*Load		-0.031	-0.033		0.011	0.008
		(-1.24)	(-1.31)		(0.48)	(0.37)
ActiveShare	0.477**	0.482**	0.523**	0.353	0.295	0.378*
R2	0.871	1.500**	1.445**	1.655***	2.393***	2.308***
SDI	0.469	0.471	0.402	0.423	0.351	0.223
Vol	-0.175***	-0.159***	-0.156***	-0.168***	-0.078**	-0.077**
Log(Age)	-0.352***	-0.368***	-0.367***	-0.357***	-0.334***	-0.335***
Log(TNA)	-0.014	-0.025*	-0.025*	-0.018	-0.036***	-0.036***
Expense	-0.126	-0.131	-0.129	-0.087	-0.100	-0.096
Load	0.009	0.010	0.010	0.003	0.008	0.008
Turnover	0.197***	0.211***	0.212***	0.167***	0.185***	0.184***
InstRatio	-0.287***	-0.281***	-0.280***	-0.294***	-0.281***	-0.281***
StockSize	-0.098**	-0.075	-0.072	-0.018	0.011	0.017
HHI	-0.858	-0.676	-0.728	0.065	0.436	0.346
Constant	0.136	-0.504	-0.487	-1.447*	-2.609***	-2.567***
Observations	97628	97628	97628	97628	97628	97628
R^2	0.098	0.108	0.108	0.095	0.107	0.107

B. Uniqueness and fund flow sensitivity: Piecewise linear models

	Four-factor alpha		One-factor alpha	
	(1)	(2)	(3)	(4)
	Flow	Flow	Flow	Flow
Alpha*Uindex	-2.752***	-2.408***	-1.740***	-1.396***
	(-6.19)	(-5.89)	(-5.04)	(-4.08)
Alpha*Above*Uindex	1.629**	1.460**	1.838***	1.607***
	(2.34)	(2.25)	(4.07)	(3.54)
Above*Uindex	0.777***	0.488**	1.300***	1.201***
	(4.32)	(2.51)	(10.95)	(9.80)
Above*Alpha	-0.377	-0.206	-0.783***	-0.573**
	(-0.88)	(-0.55)	(-2.78)	(-2.07)
Uindex	-0.506***	-0.327**	-0.645***	-0.524***
	(-3.13)	(-2.02)	(-4.44)	(-3.42)
Alpha	3.043***	5.944***	1.860***	4.093***
	(10.15)	(13.33)	(8.71)	(10.96)
Above	0.182**	0.130	0.496***	0.326***
	(2.09)	(1.55)	(9.64)	(6.08)
Alpha*Vol		-0.265***		-0.243***
		(-9.87)		(-8.41)
Alpha*Log(Age)		-0.452***		-0.288***
		(-5.11)		(-3.84)
Alpha*Load		-0.033		0.009
		(-1.34)		(0.41)
ActiveShare	0.540**	0.538**	0.377*	0.332
R2	0.909	1.447**	1.471**	2.115***
SDI	0.333	0.373	0.144	0.157
Vol	-0.178***	-0.160***	-0.165***	-0.086***
Log(Age)	-0.352***	-0.370***	-0.344***	-0.327***
Log(TNA)	-0.018	-0.026*	-0.033**	-0.046***
Expense	-0.129	-0.132	-0.094	-0.103
Load	0.011	0.011	0.008	0.011
Turnover	0.188***	0.204***	0.175***	0.190***
InstRatio	-0.277***	-0.274***	-0.277***	-0.271***
StockSize	-0.087*	-0.070	-0.021	0.002
HHI	-0.960	-0.764	-0.226	0.124
Constant	0.003	-0.530	-1.249*	-2.216***
Observations	97628	97628	97628	97628
R^2	0.102	0.110	0.107	0.115

Table 6: Fund uniqueness and performance persistence

This table presents results on the effects fund uniqueness on performance persistence. *Uindex* is the return-based uniqueness index derived from cluster analysis. *Alpha* is the Carhart (1997) four-factor estimated over past 36 months. *Above* is a dummy variable equal 1 if *Alpha* is above the contemporaneous sample median, and 0 otherwise. *Alpha4F*, *Alpha5F(I)*, *Alpha5F(II)*, and *Alpha1F* are the average alphas over future 12 months estimated from the Carhart (1997) four-factor model, the Pastor and Stambaugh (2003) five-factor model, an alternative five-factor model (Carhart (1997) factors plus an equity trend-following factor), and the CAPM, respectively. *ExRet* is the average excess raw return over future 12 months. See the caption of Table 1 for the definitions of other variables. All models are estimated using quarterly data, and all independent variables are lagged by one quarter. The fixed-effect models include both quarter and fund style fixed effects, and the t-statistics are based on standard errors clustered by both fund and quarter. The Fama-MacBeth regressions control for fund style fixed effects, and the t-statistics are Newey-West corrected for autocorrelation up to order three. The t-statistics for the usual controls are omitted to save space. Significance at the 1%, 5%, and 10% levels are indicated by ***, **, and *, respectively.

A. Fixed-effects models

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Alpha4F	Alpha4F	Alpha4F	Alpha5F(I)	Alpha5F(II)	Alpha1F	ExRet
Alpha*Uindex		0.313***	0.322***	0.331***	0.337***	0.242**	0.359***
		(3.24)	(3.32)	(3.42)	(3.36)	(2.24)	(3.24)
Uindex	0.042	0.015	-0.018	-0.011	-0.021	-0.010	-0.110*
	(1.16)	(0.48)	(-0.45)	(-0.27)	(-0.49)	(-0.19)	(-1.70)
Alpha		-0.101	-0.147	-0.102	-0.135	-0.286*	-0.238
		(-1.23)	(-1.07)	(-0.74)	(-0.92)	(-1.83)	(-1.47)
Alpha*Vol			-0.022**	-0.026***	-0.030***	0.006	-0.047***
			(-2.27)	(-2.63)	(-2.94)	(0.39)	(-2.87)
Alpha*Log(Age)			0.058	0.041	0.064	0.088**	0.111**
			(1.43)	(1.00)	(1.52)	(2.01)	(2.37)
ActiveShare			0.003	-0.015	0.007	0.282**	0.375***
R2			-0.444	-0.490*	-0.621**	-0.463	0.100
SDI			0.046	0.064	0.138	-0.261	0.538
Vol	-0.063***		-0.059***	-0.053***	-0.074***	-0.111***	-0.058*
Log(Age)	-0.010		0.005	0.007	0.011	0.014	0.018*
Log(TNA)	-0.003		-0.007*	-0.009**	-0.005	-0.013***	-0.012**
Expense	-0.123***		-0.097***	-0.078***	-0.075***	-0.136***	-0.124***
Load	-0.001		-0.002	-0.003	-0.003	0.000	-0.003
Turnover	-0.019		-0.025*	-0.026*	-0.009	0.008	-0.022
InstRatio	-0.022		-0.009	0.000	0.001	-0.012	-0.015
StockSize	-0.036**		-0.036**	-0.029*	-0.035**	-0.108***	-0.105***
HHI	-0.068		-0.189	-0.208	-0.269	-0.700*	-0.739
Constant	0.501***	-0.238***	0.835**	0.847**	1.052***	1.418***	2.229***
Observations	102440	108541	92045	92045	88256	92045	92045
R^2	0.082	0.069	0.101	0.087	0.113	0.152	0.793

B. Fama-MacBeth regressions

	(1) Alpha4F	(2) Alpha4F	(3) Alpha4F	(4) Alpha5F(I)	(5) Alpha5F(II)	(6) Alpha1F	(7) ExRet
Alpha*Uindex		0.234**	0.184***	0.182***	0.203***	0.127*	0.153**
		(2.60)	(2.83)	(3.21)	(2.87)	(1.95)	(2.35)
Uindex	0.021	0.029	0.019	0.032	0.026	0.014	0.016
	(0.49)	(0.74)	(0.34)	(0.60)	(0.45)	(0.39)	(0.36)
Alpha		-0.002	-0.061	-0.039	-0.082	-0.048	-0.053
		(-0.03)	(-0.56)	(-0.37)	(-0.65)	(-0.40)	(-0.45)
Alpha*Vol			-0.019	-0.018	-0.011	-0.032	-0.037**
			(-1.16)	(-1.22)	(-0.61)	(-1.58)	(-2.40)
Alpha*Log(Age)			0.081***	0.069***	0.080***	0.106***	0.108***
			(3.90)	(3.70)	(3.48)	(6.08)	(5.69)
ActiveShare			0.036	0.009	0.118	0.082	0.087
R2			0.187	0.033	0.350	0.196	0.414
SDI			0.358	0.235	0.787	0.148	-0.060
Vol	-0.066***		-0.061**	-0.064***	-0.041*	-0.087	0.010
Log(Age)	-0.010		0.005	0.007	0.015	0.008	0.010
Log(TNA)	-0.001		-0.008*	-0.010**	-0.008	-0.008*	-0.007
Expense	-0.135***		-0.104***	-0.090***	-0.092***	-0.100***	-0.101***
Load	0.001		0.001	-0.000	0.001	-0.001	-0.002
Turnover	-0.008		-0.004	-0.005	0.013	0.015	0.012
InstRatio	-0.001		0.005	0.010	0.018	-0.015	-0.015
StockSize	-0.026		-0.031	-0.029	-0.022	-0.049	-0.041
HHI	0.587		0.177	0.102	-0.540	0.175	0.205
Constant	0.588**	-0.170***	0.353	0.538	0.016	0.790	0.581
Observations	102440	108541	92045	92045	88256	92045	92045
R^2	0.014	0.005	0.022	0.014	0.023	0.025	0.034

C. Asymmetric effects of uniqueness on performance persistence

	Fixed-effects models			Fama-MacBeth regressions		
	(1)	(2)	(3)	(4)	(5)	(6)
	Alpha4F	Alpha5F(I)	ExRet	Alpha4F	Alpha5F(I)	ExRet
Alpha*Uindex	0.447**	0.454**	0.612***	0.453***	0.478***	0.439***
	(2.34)	(2.29)	(2.96)	(3.61)	(3.98)	(3.52)
Alpha*Above*Uindex	-0.105	-0.094	-0.185	-0.474***	-0.497***	-0.531***
	(-0.47)	(-0.41)	(-0.74)	(-3.55)	(-3.52)	(-3.50)
Above*Uindex	-0.102	-0.114	-0.205***	-0.069	-0.093*	-0.042
	(-1.40)	(-1.53)	(-2.69)	(-1.33)	(-1.80)	(-0.79)
Above*Alpha	-0.186	-0.146	-0.221	0.079	0.114	0.096
	(-1.24)	(-0.98)	(-1.29)	(0.94)	(1.33)	(0.90)
Above	0.012	0.016	0.060*	0.023	0.027	0.020
	(0.43)	(0.55)	(1.83)	(0.86)	(1.00)	(0.68)
Uindex	0.074	0.082	0.057	0.130**	0.157***	0.117**
	(0.95)	(1.02)	(0.58)	(2.15)	(2.66)	(2.05)
Alpha	-0.038	-0.013	-0.152	-0.063	-0.061	-0.060
	(-0.25)	(-0.08)	(-0.80)	(-0.47)	(-0.46)	(-0.45)
Alpha*Vol	-0.022**	-0.027**	-0.047***	-0.024	-0.024	-0.043***
	(-2.14)	(-2.56)	(-2.82)	(-1.34)	(-1.51)	(-2.91)
Alpha*Log(Age)	0.046	0.032	0.095**	0.068***	0.057***	0.093***
	(1.28)	(0.87)	(2.23)	(3.52)	(3.20)	(5.18)
ActiveShare	0.004	-0.015	0.372***	0.019	-0.007	0.071
R2	-0.505*	-0.538*	0.017	0.145	-0.004	0.376
SDI	0.032	0.053	0.518	0.387	0.268	-0.025
Vol	-0.053***	-0.049***	-0.051*	-0.056**	-0.059**	0.016
Log(Age)	0.005	0.006	0.017	0.003	0.005	0.009
Log(TNA)	-0.006	-0.009**	-0.011**	-0.008*	-0.010**	-0.006
Expense	-0.094***	-0.075***	-0.120***	-0.104***	-0.089***	-0.100***
Load	-0.002	-0.004	-0.003	0.001	-0.000	-0.002
Turnover	-0.022	-0.023	-0.018	-0.003	-0.003	0.013
InstRatio	-0.010	-0.001	-0.016	0.004	0.009	-0.012
StockSize	-0.037**	-0.031*	-0.107***	-0.032	-0.029	-0.041
HHI	-0.176	-0.194	-0.714	0.244	0.171	0.301
Constant	0.879***	0.877**	2.267***	0.353	0.526	0.576
Observations	92045	92045	92045	92045	92045	92045
R^2	0.103	0.089	0.793	0.024	0.015	0.039

Table 7: Robustness Checks

This table presents results of three sets of robustness checks on the effects of fund uniqueness on management fee and total expense ratios (Columns (1) and (2)), flow-performance sensitivity (Columns (3) and (4)), and performance persistence (Columns (5) and (6)). Panel A presents results when fund uniqueness is measured by the return-based uniqueness index derived from partitioning cluster analysis (*Uindex(PCA)*), Panel B presents results when fund uniqueness is measured by the holding-based uniqueness index (*Uindex(HLD)*), and Panel C presents results when sector funds and long-short funds are excluded. *Alpha* is the Carhart (1997) four-factor alpha estimated over past 36 months. *Above* is a dummy variable that equals 1 if *Alpha* is above the contemporaneous sample median, and 0 otherwise. *Alpha4F* is the average Carhart (1997) alpha over future 12 months. See the caption of Table 1 for the definitions of other variables. The management fee and total expense models are estimated using annual data. All other models are estimated using quarterly data. The first five models are estimated using panel regressions with both time and fund style fixed effects, and with standard errors clustered by both fund and time. The last model is estimated using the Fama-MacBeth approach with t-statistics Newey-West adjusted for autocorrelation up to order three. The t-statistics for the usual controls are omitted to save space. Significance at the 1%, 5%, and 10% levels are indicated by ***, **, and *, respectively.

A. PCA-based uniqueness measure

	(1) MgtFee	(2) Expense	(3) Flow	(4) Flow	(5) Alpha4F	(6) Alpha4F
Uindex(PCA)	0.052***	0.067***	0.334***	-0.195	-0.057	0.055
	(3.67)	(3.22)	(2.59)	(-1.22)	(-1.33)	(1.17)
Alpha*Uindex(PCA)			-1.298***	-2.137***	0.325***	0.490***
			(-5.04)	(-5.15)	(3.04)	(3.46)
Alpha*Above*Uindex(PCA)				1.341**		-0.448***
				(2.08)		(-2.76)
Above*Uindex(PCA)				0.549***		-0.100*
				(2.74)		(-1.85)
Above*Alpha				-0.098		0.078
				(-0.27)		(0.79)
Alpha			5.958***	5.549***	-0.097	-0.043
			(15.19)	(12.85)	(-0.74)	(-0.32)
AlphaSQ			0.167**			
			(2.11)			
Above				0.150*		0.033
				(1.74)		(1.47)
Alpha*Vol			-0.286***	-0.254***	-0.026**	-0.026
			(-9.67)	(-9.04)	(-2.51)	(-1.36)
Alpha*Log(Age)			-0.425***	-0.439***	0.053	0.061***
			(-4.72)	(-5.02)	(1.34)	(3.14)
Alpha*Load			-0.029	-0.031		
			(-1.13)	(-1.24)		
ActiveShare	0.333***	0.284***	0.562**	0.561**	-0.002	0.017
R2	0.043	-0.031	1.452**	1.377**	-0.451*	0.092
SDI	0.168***	0.164**	0.173	0.126	0.101	0.483
Vol	0.007**	0.033***	-0.166***	-0.167***	-0.057***	-0.051**
Log(Age)	-0.045***	-0.040***	-0.363***	-0.367***	0.004	0.004
Log(TNA)	0.008*	-0.076***	-0.029**	-0.030**	-0.006	-0.008*
Expense			-0.131	-0.134*	-0.097***	-0.102***
Load			0.009	0.010	-0.002	0.001
Turnover	0.010**	0.032***	0.210***	0.203***	-0.025*	-0.002
InstRatio	-0.031**	-0.340***	-0.281***	-0.273***	-0.010	0.001
StockSize	-0.016**	0.002	-0.061	-0.062	-0.038***	-0.032
HHI	0.254*	0.207	-0.592	-0.697	-0.197	0.354
Constant	0.756***	1.484***	-0.636	-0.569	0.866***	0.404
Observations	23596	25287	97628	97628	92045	92045
R^2	0.123	0.358	0.107	0.109	0.101	0.025

B. Holding-based uniqueness measure

	(1)	(2)	(3)	(4)	(5)	(6)
	MgtFee	Expense	Flow	Flow	Alpha4F	Alpha4F
Uindex(HLD)	0.060***	0.039*	0.117	-0.198	-0.075*	0.113
	(3.43)	(1.83)	(1.15)	(-1.23)	(-1.83)	(1.59)
Alpha*Uindex(HLD)			-1.036***	-1.579***	0.144	0.440**
			(-3.71)	(-3.99)	(1.57)	(2.41)
Alpha*Above*Uindex(HLD)				1.360*		-0.415**
				(1.86)		(-2.21)
Above*Uindex(HLD)				0.164		-0.122
				(0.73)		(-1.59)
Above*Alpha				0.117		-0.044
				(0.40)		(-0.63)
Alpha			5.829***	5.114***	-0.006	0.012
			(13.95)	(12.19)	(-0.05)	(0.09)
AlphaSQ			0.165**			
			(1.97)			
Above				0.378***		0.020
				(4.59)		(0.65)
Alpha*Vol			-0.317***	-0.276***	-0.017*	-0.016
			(-11.19)	(-10.46)	(-1.80)	(-0.95)
Alpha*Log(Age)			-0.420***	-0.442***	0.045	0.070***
			(-4.57)	(-4.79)	(1.23)	(3.65)
Alpha*Load			-0.029	-0.029		
			(-1.14)	(-1.16)		
ActiveShare	0.346***	0.314***	0.707***	0.673***	-0.028	0.007
R2	0.045	-0.024	1.320**	1.328**	-0.412	0.097
SDI	0.198***	0.229***	0.418	0.407	0.103	0.381
Vol	0.008**	0.036***	-0.151***	-0.153***	-0.060***	-0.055**
Log(Age)	-0.045***	-0.041***	-0.360***	-0.366***	0.003	0.003
Log(TNA)	0.009**	-0.076***	-0.028**	-0.029**	-0.007	-0.009*
Expense			-0.121	-0.124	-0.097***	-0.107***
Load			0.007	0.008	-0.002	0.001
Turnover	0.010**	0.032***	0.212***	0.206***	-0.024	-0.001
InstRatio	-0.032**	-0.343***	-0.296***	-0.285***	-0.009	0.001
StockSize	-0.015**	0.004	-0.059	-0.059	-0.041***	-0.034
HHI	0.157	0.094	-0.897	-0.916	0.028	0.428
Constant	0.737***	1.443***	0.406	0.279	0.888**	0.443
Observations	23593	25284	97589	97589	92022	92022
R^2	0.123	0.357	0.106	0.108	0.099	0.022

C. Sample excluding sector and long/short funds

	(1) MgtFee	(2) Expense	(3) Flow	(4) Flow	(5) Alpha4F	(6) Alpha4F
Uindex	0.039**	0.087***	-0.020	-0.533***	0.047	0.135***
	(2.35)	(4.19)	(-0.17)	(-3.41)	(1.21)	(3.07)
Alpha*Uindex			-0.845***	-1.763***	0.201***	0.321***
			(-3.22)	(-5.23)	(2.90)	(3.56)
Alpha*Above*Uindex				1.339**		-0.344**
				(2.00)		(-2.37)
Above*Uindex				0.556***		-0.088**
				(3.08)		(-2.39)
Above*Alpha				0.326		0.121
				(0.87)		(1.20)
Above				0.004		0.052**
				(0.05)		(2.28)
Alpha			6.824***	6.373***	-0.150	-0.179
			(15.71)	(13.98)	(-1.33)	(-1.51)
AlphaSQ			0.287**			
			(2.24)			
Alpha*Vol			-0.398***	-0.369***	-0.002	0.012
			(-10.44)	(-9.41)	(-0.17)	(0.74)
Alpha*Log(Age)			-0.400***	-0.405***	0.044	0.039**
			(-3.96)	(-4.07)	(1.58)	(2.28)
Alpha*Load			-0.033	-0.032		
			(-1.19)	(-1.18)		
ActiveShare	0.298***	0.236***	0.264	0.300	0.016	-0.008
R2	-0.134	-0.239*	0.633	0.844	-0.201	-0.111
SDI	0.158	0.036	1.351**	1.422**	-0.058	0.216
Vol	0.005	0.034***	-0.125***	-0.132***	-0.093***	-0.076***
Log(Age)	-0.038***	-0.032***	-0.367***	-0.368***	0.004	0.002
Log(TNA)	0.009**	-0.069***	-0.020	-0.022	-0.007**	-0.009*
Expense			-0.030	-0.033	-0.098***	-0.101***
Load			0.023**	0.023**	-0.002	0.000
Turnover	0.006	0.042***	0.069*	0.062	-0.007	0.006
InstRatio	-0.023*	-0.327***	-0.163**	-0.159**	-0.027**	-0.011
StockSize	-0.018***	-0.002	-0.032	-0.028	-0.030**	-0.024
HHI	0.228	0.203	-1.260	-1.356	-0.434	-0.054
Constant	0.945***	1.669***	0.924	0.716	0.721***	0.654
Observations	20580	22253	85938	85938	81142	81142
R^2	0.134	0.353	0.140	0.142	0.131	0.024