Mutual Funds Apart From the Crowd

Nadia Vozlyublennaia Youchang Wu*

May 10, 2016

Abstract

We construct measures of mutual fund uniqueness using cluster analysis of fund returns. We find that more unique funds charge higher management fees and deliver better net-of-fee performance than do funds that are otherwise similar. Fund uniqueness significantly reduces the sensitivity of fund flows to fund performance, and increases performance persistence. In addition, both the dampening effects of fund uniqueness on the flow-performance sensitivity and the amplifying effect of fund uniqueness on performance persistence are stronger for underperforming than for well-performing funds. These results suggest that unique funds have better managerial skill, and are better able to retain investors after poor performance, potentially due to the lack of close substitutes. They also suggest that the slow adjustment of fund size toward the equilibrium level increases the persistence of poor performance of unique funds.

JEL codes: G23, G34

Key Words: mutual fund uniqueness, competition, fund flow, performance persistence

*Vozlyublennaia (vozlyublennaian@sec.gov) is at the U.S. Securities and Exchange Commission. Wu (ywu2@uoregon.edu) is at the Lundquist College of Business, University of Oregon. We thank seminar participants at the Securities and Exchange Commission and the University of Oregon for comments and suggestions.

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 asset management business. According to the 2016 Investment Company Fact Book published by the Investment Company Institute, there were 9,520 U.S.-registered open-end funds managing a total of $15.85 trillion assets at the end of the year 2014. 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. Which funds are more likely to deviate from the crowd? Are funds with more unique strategies rewarded by higher fees? Can they generate higher net-of-fee returns for investors? How does fund uniqueness change the elasticity of investor demand for a mutual fund with respect to past performance? And how does it affect 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 formulate several empirical hypotheses, and develop measures of fund uniqueness. We examine how these measures are related to mutual fund fees and performance, the response of investors to fund performance, as well as the persistence of fund performance. Our main measure of uniqueness is based on hierarchical cluster analysis of fund returns. Using a sample of actively managed U.S. domestic equity funds, we find three main results. First, more unique funds have higher management fees and higher total expense ratios, yet deliver better net-of-fee performance. All else equal, as the uniqueness index increases from the 25th to the 75th percentile, the annual management fee ratio increases by 9.2 basis points, and the total expense ratio increases by 10.7 basis points. Despite the higher fees and expenses, the net-of-fee Carhart (1997) four-factor alpha increases by about 41 to 59 basis points per

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

year. These results are novel, as the literature generally finds that high expenses are associated with poor performance (for example, Carhart (1997)). They suggest that unique funds have better managerial skill than do non-unique funds that are otherwise similar.

Second, 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. As the fund uniqueness increases from the 25th to the 75th percentile, the sensitivity of fund flows to past performance declines by about 22%. 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. The low flow-performance sensitivity of unique funds is likely due to the difficulty in evaluating such funds and their lack of a broad investor base. The additional dampening effect of fund uniqueness on the response of fund flows to poor performance is likely due to the lack of close substitutes after a unique fund performs poorly.

Third, fund uniqueness increases the persistence of fund performance, especially the persistence of poor performance. Among the least unique funds, there is a weak tendency of performance reversal instead of persistence. However, for a fund with a uniqueness index one standard deviation above the mean, a one standard deviation increase in past four-factor alpha predicts an increase of 94 basis points in the four-factor alpha over the next 12 months. These results suggest that mutual fund performance persistence found in the literature by, for example Brown and Goetzmann (1995) and Carhart (1997), is mainly driven by high uniqueness funds. The higher performance persistence of unique funds can arise either because of the difficulty faced by other funds in mimicking their investment strategies, or because of the stickiness of the money invested in those funds. However, the fact that fund uniqueness increases the persistence of poor performance more than the persistence of good performance suggests that the latter may be a more important reason.

These results remain significant even after controlling for the deviation a fund's portfolio from its benchmark portfolio, i.e., the active share. They are also robust to two alternative measures of fund uniqueness. One is based on a sequence of partitioning cluster analysis with increasing granularity, the other is the Strategy Distinctiveness Index (SDI) developed by Sun, Wang, and Zheng (2012) for analyzing hedge fund returns.

In summary, we show that managers of funds with more unique investment strategies are more skilled and are rewarded by higher fees. Part of the benefits from unique investment strategies is shared by investors, potentially as a compensation for the high search costs associated with such funds. In addition, fund uniqueness reduces the elasticity of investor demand with respect to past performance, and increases the convexity of the flow-performance relation. Since the equilibrating downward adjustment of fund size is slowed down, the poor performance becomes more persistent. These results demonstrate rich interaction between the competitions on the asset side (for return generation) and the liability side (for fund flows) of a mutual fund's balance sheet. The ability to employ innovative investment strategies not only allows a fund to generate higher asset returns and charges higher fees, but also makes it less substitutable, and thus improve its ability to retain fund investors after poor performance. This in turn affects the persistence of the fund's performance.

Our study contributes to the literature on two fronts. First, we contribute to the understanding of the mutual fund industry. To the extent that asset complexity and search frictions for finding an informed manager are higher for unique funds, our findings of high management fees and superior net-of-fee performance of unique funds provide support for the general equilibrium model of the asset and asset management markets recently developed by Gârleanu and Pedersen (2016). In addition, we show that the convexity in the flow-performance relation and performance persistence documented in the literature are largely due to funds with more unique investment strategies. By examining the patterns of asymmetry, we offer a coherent explanation for both the low flow-performance sensitivity and the high performance persistence of unique funds, in light of the Berk and Green (2004) argument of mutual fund industry equilibrium.

Second, we propose two return-based measures of mutual fund uniqueness, and show that they can be used to identify funds with managerial skill. One is based on Hierarchical Cluster Analysis (HCA), which considers the number of steps needed to separate one fund from others.

The other is based on Partitioning Cluster Analysis (PCA), which considers the number of funds in a fund's clusters at various levels of fund classification granularity.[2] In the literature, PCA has been 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. To the best of our knowledge, no previous study has used PCA or HCA to measure fund uniqueness.

Our paper is closely related to the study of Sun, Wang, and Zheng (2012), who show that hedge funds following distinctive investment strategies have better net-of-fee performance. 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 SDI. Our measures of fund uniqueness differ from the SDI by using the information in the whole hierarchical cluster structure. More importantly, the focus of our paper is very different from theirs. While Sun, Wang, and Zheng (2012) are mainly interested in the relation between strategy distinctiveness and the net-of-fee returns to hedge fund investors, we investigate how fund uniqueness is related to both fees and net-of-fee performance of mutual funds in a context of market equilibrium. Furthermore, we examine how uniqueness alters the flow-performance sensitivity and performance persistence of mutual funds.

Hoberg, Kumar, and Prabhala (2015) 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). Consistent with our finding of stronger performance persistence among unique funds, they show that fund performance is more persistent for funds with a smaller number of competing funds in their neighborhoods. Besides the difference between our return-based uniqueness measure and their holding-based measure of competition environment, our study differs from theirs in several other important aspects. First, they do not consider how a fund's competition environment affects the response of fund flows to performance, and how this may in turn affect performance persistence. They focus on the

[2] Both measures rank fund uniqueness within its style group. While a fund's style can be identified either by its declared investment objective or characteristics of its portfolio holdings, we choose to classify funds into different style groups based on PCA of realized returns, following Sun, Wang, and Zheng (2012).

competition pressure on the asset side of a fund's balance sheet. In contrast, we consider the effects of uniqueness on competitions on both the asset (return generation) and liability (fund flows) sides of a mutual fund, as well as the potential feedback effect of fund flows on asset performance. Second, they do not examine potential asymmetry in the effects of competition on performance persistence, which, as we show, contains important information about the source of the high performance persistence of unique funds. Third, they use the holding-based alpha developed by Daniel, Grinblatt, Titman, and Wermers (1997) to measure performance. This measure ignores transaction costs, fund expenses, and intra-quarter trades, and therefore, is silent about how the benefits from low competition, if any, are split between fund managers and investors.

Many other papers have studied potential indicators that are useful for identifying mutual fund managerial skill. Consistent with our findings that skilled managers are more likely to deviate from the crowd, 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, and exhibit strong performance persistence. Our paper contributes to this line of research by presenting fund uniqueness as a new indicator of manager skill. While our measures of fund uniqueness are positively correlated with active share, suggesting a positive connection between the deviation from the crowd of actively managed mutual funds and the deviation from passive benchmarks, our results remain largely unchanged after controlling for active share.

The rest of the paper is structured as follows. Section 1 outlines the hypotheses and describes our method of measuring uniqueness. Section2 describes the data and examines the determinants of fund uniqueness. We examine how fund uniqueness is related to fees and performance in Section 3, how fund uniqueness affects the flow-performance relation and performance persistence in Section 4, and investigate the asymmetries in the effects of fund uniqueness on fund flows and performance persistence in Section 5. We then present results from robustness tests in Section 6, and conclude in Section 7.

1 Hypotheses and Methodology

In this section, we first outline several hypotheses about mutual funds apart from the crowd, drawing on recent developments in the theoretical literature of delegated asset management, and then describe our measures of fund uniqueness.

1.1 Hypotheses

Gârleanu and Pedersen (2016) develop a general equilibrium model of the asset and asset management markets, in which both information about fund managers and about assets is costly. Their model links the efficiency of asset prices to the efficiency of the asset management market, and has several predictions that are relevant for our understanding of fees and performance of unique funds. First, as search costs for finding an informed manager and asset complexity increase, asset prices become less efficient, and asset management fees increase. Second, as a compensation for searching investors, the net-of-fee performance of informed managers increases as search costs increase. Third, skilled managers, i.e., managers with low information costs, are more likely to engage in collecting information about assets than are unskilled managers.

Stoughton, Wu, and Zechner (2011) present a partial equilibrium model of delegated asset management with search costs. In the model, investors can search for an informed manager directly, or invest through an intermediary, such as a financial adviser. In equilibrium, the net-of-fee performance of the actively managed fund is positively related to search costs in the absence of intermediaries, and is positively related to the costs of intermediation when intermediaries exist.

It is reasonable to expect that search costs of finding an informed manager and costs of intermediation increase with fund uniqueness. Evaluating a fund with unique strategies is more difficult than evaluating a "mainstream" fund. By definition, such funds do not have close peers that can be used as a benchmark to filter out noise in performance. They may be exposed to risk factors that are not captured by the standard asset pricing models. Also, financial advisers may be reluctant to channel investment into unique funds due to the difficulty in evaluating and monitoring such funds.

In addition, unique funds tend to focus on assets that are under-explored by other funds, and therefore, are less understood, more opaque, and more likely to be mispriced. Furthermore, developing and implementing innovative and unique investment strategies is clearly more challenging than herding with the crowd. Unskilled managers may therefore find it suboptimal to engage in such activities.

In light of the the models of Gârleanu and Pedersen (2016) and Stoughton, Wu, and Zechner (2011), the considerations above suggest that unique funds should have higher fees and deliver better net-of-fee performance than do non-unique funds. This leads to our first hypothesis:

H1. *Management fees and net-of-fee performance increase with fund uniqueness.*

Fund uniqueness may also affect investor demand. For several reasons, we expect fund flows to be less responsive to the past performance of unique funds. First, as we mention above, unique funds are more difficult to evaluate. Consequently, investors are likely to be more cautious in interpreting and responding to their past performance. Second, unique funds may cater to a specific clientele of investors, and therefore, they may not have a broad base of investors who can quickly react to past performance. Third, unique funds, by definition, do not have close substitutes, so even when they underperform, investors may still hold them because it is difficult to find a replacement. This leads to our second hypothesis:

H2. *The sensitivity of fund flows to performance decreases with fund uniqueness.*

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, which 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. This leads to our third hypothesis:

H3. *The persistence of fund performance increases with fund uniqueness.*

While the difficulty of performance evaluation and the lack of a broad investor base reduce the sensitivity of fund flows to both good performance and poor performance, the lack of close substitutes mainly works to reduce the response of fund flows to poor performance. Therefore, we may expect the dampening effect of fund uniqueness on the flow-performance sensitivity to be stronger for underperforming than for well-performing funds. According to the Berk and Green (2004) argument, the downward adjustment of fund size is a mechanism for an underperforming fund to restore performance. Since unique funds are better able to retain investors after poor performance, this equilibrating process is slowed down. As a result, their poor performance is more persistent. These considerations lead to our fourth hypothesis:

H4. *The dampening effect of fund uniqueness on the flow-performance sensitivity and the amplifying effect of fund uniqueness on performance persistence is stronger for underforming than for outperforming funds.*

Our hypothesis links performance persistence directly to the sensitivity of fund flows to performance, in light of the Berk and Green (2004) argument. However, the higher performance persistence of unique funds may also arise because their investment strategies are difficult to mimic, which allows them to escape the competition and outperform consistently, as argued by Hoberg, Kumar, and Prabhala (2015). If this is indeed the case, we should expect to see stronger persistence predominantly among well-performing funds.

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), and has been successfully applied in many fields such as medicine, biology, computer science, and social science. 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 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 differ across funds. Some overlap more, and therefore, exhibit a more similar stream of returns. Others overlap less and, therefore, have a more unique return streams, which are less substitutable.

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, Squared Euclidean, Manhattan, Mahalanobis, Maximum Distance. Also, different linkage criteria, which determine the method of comparing the distance between two clusters (or a cluster and an outside element), have been proposed, such as the average linkage, centroid linkage, complete linkage, density linkage.[3] We adopt the widely-used centroid method, developed by Sokal and Michener (1958). This method uses the squared Euclidean distance as the distance measure, and use the centroid linkage, which defines the distance between two clusters as the distance between the means (centroids) of the two clusters. The advantage of this method is a straightforward interpretation and robustness to outliers.

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

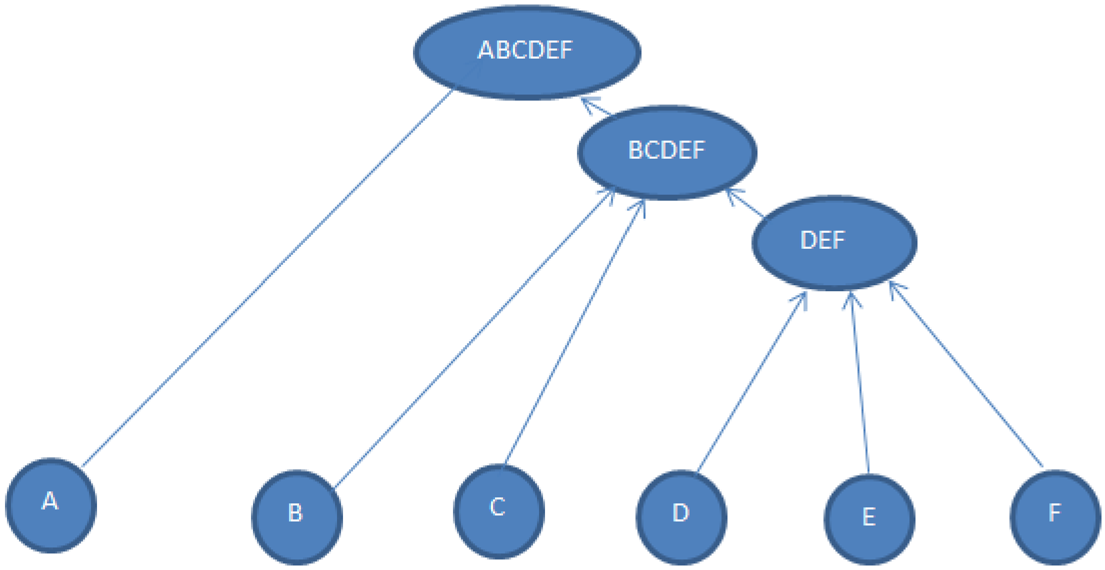

Figure 1: *An example of a Dendrogram in Hierarchical Cluster Analysis.*

Consider the following example of "bottom up" strategy, applied to A, B, C, D, E, F funds in Figure 1. Pairwise distance measures are computed in the first stage across all the elements, and the elements with the smallest distance are combined into a cluster. In this case, D, E, and F have about the same distance, which is smaller than the distance between elements of any other pair. Next, an average for the newly formed cluster is computed, which will be considered as a separate element in place of D, E, F in the subsequent steps. Next, distances are computed across A, B, C and DEF elements and a new cluster is formed (BCDEF) based on the smallest distance. At the last step, the last element A is joined with BCDEF, and the formation of the tree is complete. By construction, the elements that were joined later (such as A) are members of a smaller number of subclusters 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 subclusters each element belongs to, can be considered as an inverse measure of uniqueness, as it represents the number of steps needed in order to separate it from other funds. According to this measure, A is the most unique fund, as it does not belong to any subset of the universe. B and C are the second, as they belong to one subcluster. D, E, F are the least unique funds, as they belong to the largest number of subclusters (two 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 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 funds' uniqueness/substitutability within its style using HCA. The advantage 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.

We conduct our two-stage cluster analysis of monthly returns at the quarterly frequency, using a rolling window of 36 months. Funds are required to have non-missing returns in a given 36-month window to be included in the analysis. We standardize returns in each month by the cross-sectional mean and standard deviation, which is customarily done in order to preclude influence of outliers.

As mentioned above, the total number of subclusters that a fund belongs to according to the HCA can be interpreted as an inverse measure of the fund's 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 HCA-Based Uniqueness Index.[4] Funds with a uniqueness index of 1 are most unique and least substitutable, while funds with an index of 0 are least unique and most substitutable.

For robustness checks, we also compute two alternative measures of fund uniqueness. First, we replace the second-stage HCA by a sequence of PCA with increasing granularity. That is,

[4]Specifically, the HCA-Based Uniqueness Index is calculated as $1 - N_i/N_{max}$, where N_i is the number of subclusters other than the style group that a fund belongs to, N_{max} is the largest value of N_i across all funds in a style group for a given period.

for each of the ten style groups emerging 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 $\overline{N}_{K,i}$. We normalize this average to be in the [0, 1] interval, and use 1 minus the normalized value as our second measure of fund uniqueness, called the PCA-Based Uniqueness Index.[5] Like the HCA-Based Uniqueness Index, this index also uses information in the clustering structure of fund returns. Intuitively, the HCA-based index is inversely related to the *length* of a fund's subcluster chain, while PCA-based index is inversely related to the *width* of a fund's subcluster chain at given levels of granularity. However, the choice of the granularity levels to be considered is somewhat arbitrary.

Second, we use the SDI proposed by Sun, Wang, and Zheng (2012) as an additional measure of fund uniqueness. 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, and the style groups are identified by the first-stage PCA. Like the other two measures, we compute the SDI using a rolling window of 36 months.

2 Data, Summary Statistics, and Determinants of Fund Uniqueness

2.1 Data

We use a sample of actively-managed domestic equity funds in the CRSP survivor-bias-free database for our analysis. The database contains information about mutual fund historical

[5]Specifically, the PCA-Based Uniqueness Index is defined as $1-\frac{\overline{N}_{K,i}-\overline{N}_{K,min}}{\overline{N}_{K,max}-\overline{N}_{K,min}}$, where $\overline{N}_{K,max}$ and $\overline{N}_{K,min}$ represent maximum and minimum values of $\overline{N}_{K,i}$ across all funds in a fund style group for a given period, respectively.

return, total net asset value (TNA), expense ratio, and other fund characteristics at the share class level. We use the MFLINKS database from the WRDS to link different share classes of the same fund, and use the CRSP fund objective code to identify domestic equity funds (the first two letters of *crsp_obj_cd* being "ED"). We drop ETFs, variable annuity funds, 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, we conduct our analysis on a sample that begins in January 1991 and ends in June 2014. We drop funds with less than three years of data over this sample period.

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 total net asset value 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 total net asset value. 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). Our final sample includes a total of 3,538 funds, with an average of 1,785 funds in a given quarter.

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. *Nsub* is the number of subclusters that a fund belongs to in a given period (in addition to its style group) based on the HCA. It shows a large dispersion across funds, ranging from 0 to 411. In the last quarter of our sample period, we find 86 funds with a *Nsub* above 300. A majority (61) of them are large-cap funds, indicating that large-cap funds are generally less unique. 69 funds have a *Nsub* below 10. Among them 33 are sector funds, 14 are small-cap funds, and 3 are alternative long/short equity funds. This is not surprising because sector funds, small-cap funds, and alternative funds are likely to have more distinctive return profiles.

Uindex is the HCA-Based Uniqueness Index, our main measure of fund uniqueness, which is equal to 1 minus the normalized *Nsub*. *Uindex(PCA)* is the PCA-Based Uniqueness Index. Although both measures are normalized to be in the [0, 1] interval, neither of them is strictly evenly distributed over this interval because multiple funds can have a common index value. Therefore, their means and medians are slightly below 0.5. *SDI* is the Strategy Distinctiveness Index. Its value ranges from 0 to 1.85. However, the mean and median are only 0.08 and 0.05, respectively, suggesting a high correlation between the fund return and the group mean for a majority of funds. *AS* the active share measure developed by Cremers and Petajisto (2009), which measures the deviation of fund's portfolio from its benchmark index portfolio.[6]

Panel B of Table 1 presents the pairwise correlation coefficients. Not surprisingly, the three alternative measures of fund uniqueness are highly correlated. Especially, the correlation between *Uindex* and *Uindex(PCA)* is 0.76. The correlations of these two variables with the *SDI* are 0.48 and 0.39, respectively. Notably, these measures are also positively correlated with the active share, expense and management fee ratios, idiosyncratic volatility of fund returns, and portfolio turnover, and is negatively correlated with fund size, fund age, and the institutional share of fund assets. This indicates that more unique funds trade more, charge more, and deviate more from their benchmarks. They are younger, smaller, and less likely to attract inflows through institutional share classes. In particular, the correlation between *Uindex* and active share (AS) is 0.48, suggesting that the deviation from the benchmark portfolio can proxy for a fund's deviation from other active funds to some degree. Interestingly, while *Uindex* and

[6]Following Cremers and Petajisto (2009), we compute the 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. Active share data for the earlier quarters are downloaded from Antti Petajisto's website, whom we thank for making them available. We compare the 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 correlation is 0.99.

Uindex(PCA) are positively correlated with the total volatility of fund return, the correlation between *SDI* and the total volatility is -0.19. This negative correlation arises because *SDI* is inversely related to a fund's correlation with the style mean, and the latter is usually highly correlated with the market return.

2.3 Determinants of Fund Uniqueness

To further investigate the characteristics of unique funds, we regress our measures of fund uniqueness on lagged fund characteristics, including lagged fund uniqueness, at the annual frequency. We use two alternative methods to uncover the cross-sectional determinant of fund uniqueness. First, we run panel regressions with fixed year effects. Following Petersen (2009), we account for error clustering at both the fund and year levels. Second, we estimate the models using the Fama-Macbeth procedure (Fama and MacBeth (1973)). We use the Newey-West (Newey and West (1987)) correction to adjust for the effects of autocorrelation of the coefficient estimates (up to order three) on the t-statistics. The results, reported in Table 2, are largely consistent with the results on pairwise correlations reported in Panel B of Table 1. High uniqueness is consistently predicted by high expense ratio and high active share. While the coefficients are not always significant, high fund uniqueness is generally associated with high idiosyncratic volatility, low systematic volatility, low load fees, low institutional share of fund assets, and smaller fund size.

While one may expect institutional investors to invest more heavily in unique funds, as they might be better able to evaluate such funds, the negative coefficient on *InstRatio* in the first three columns of Table 2 shows the opposite. This result should be interpreted with caution. A large proportion of the institutional share classes in the CRSP database represents retail investment in mutual funds through the 401(k) retirement plan. Therefore, they do not capture the behavior of institutional investors. The negative relation between load fees and fund uniqueness suggests that fund brokers, which are usually compensated by load fees, tend to avoid unique funds, potentially because of their high idiosyncratic risk.

The negative relation between fund uniqueness and fund size, observed in several columns, suggests that unique funds grow slower, perhaps because they are difficult to evaluate for general investors. Also, large funds may find it difficult to deviate from the crowd as the price impact of their trades may be too big for "exotic" assets. Notably, while the univariate relation between fund uniqueness and fund age is negative, this relation becomes positive after controlling for other fund characteristics, suggesting that, all else equal, an older fund tends to be more unique than a younger fund.

The coefficients on lagged uniqueness measures are highly significant in all columns, suggesting that uniqueness measured over past three years are a good indicator of uniqueness in future three years. This demonstrates that uniqueness is a persistent fund characteristic.

3 Fees, Performance, and Fund Uniqueness

We now examine the fees and performance of unique funds, compared to funds that are otherwise similar. According to **H1**, unique funds should have higher fees and deliver better net-of-fee performance because they entail higher search costs, invest in more complex assets, and are more likely to be run by skilled managers.

3.1 Do unique funds charge more?

We use two measures of mutual fund fees. The first is the annual management fee ratio, defined as a ratio of the 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 services providers such as custodians or bookkeepers.

We regress the management fee and total expense ratios (expressed as a percentage of the TNA) reported at the year end on the lagged uniqueness measure, *Uindex*, as well as other fund characteristics potentially related to these ratios, including the logarithm of fund size and fund age, portfolio turnover rate, the volatility of monthly excess returns (measured over the prior 36 months), and the ratio of assets in a fund's institutional share classes to the fund's TNA (*InstRatio*). Since fee ratios are unlikely to change within a given fiscal year, we conduct our tests at the annual frequency, using both panel regression models (accounting for time fixed effects and two-way error clustering) and Fama-MacBeth regressions (with Newey-West correction for autocorrelation).

Table 3 reports the regression results. Both panel regressions and Fama-MacBeth regressions reveal a strong effect of fund uniqueness on fee ratios, significant at the 1% level. The magnitude of this effect is very similar across the four columns. Take the first two columns for example. The results suggest that, all else equal, as the uniqueness index increases from the 25th percentile (0.22) to the 75th percentile (0.68), the annualized management fee increases by 9.2 basis points, while the total expense ratio increases by 10.7 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 13% and 8%, respectively, from the corresponding sample means, which are quite substantial economically. Since management fees are a part of the total expenses, our point estimates also suggest that 86% (9.2 out of 10.7) of the increase in revenues associated with the increase in fund uniqueness goes to fund managers. Only 14% of the increase goes to other service providers such as custodian banks. This is not surprising, because fund managers are the main force in designing and implementing unique investment strategies.

The effects of other fund characteristics on mutual fund fees are consistent with what one may expect. Funds with higher volatility and portfolio turnover rate have higher expenses and management fees, as those funds are more actively managed. Larger funds, and funds receiving a large fraction of investment through institutional share classes have a lower total expense ratio, while older funds tend to have lower management fees.

3.2 Do unique funds perform better?

We now investigate whether unique funds deliver higher net-of-fees performance to fund investors. We use alpha estimated from the Carhart (1997) four-factor model as our main performance measure and regress it on the lagged uniqueness index and other fund characteristics. We consider holding periods of three and twelve months, and estimate our models at the quarterly frequency. For each holding period, we use the factor loadings estimated from the prior 36 months to compute monthly factor-adjusted fund returns. We then compute the average monthly alpha over the holding period from the compounded monthly factor-adjusted returns. As in the fee analysis, we run both panel regressions and Fama-MacBeth regressions. We estimate our models at the quarterly instead of the annual frequency to increase the statistical power.

Panel A of Table 4 reports results from the panel regressions, while Panel B reports results from the Fama-MacBeth regressions. In both panels, the first two columns show results for the four-factor alpha with the holding periods of three and twelve months, respectively. Both columns show a positive relation between fund uniqueness and future net-of-fee alphas. The estimated coefficient on *Uindex* varies from 0.078 to 0.111, indicating an increase in the range of 3.4 to 4.9 basis points per month (or 41 to 59 basis points per year) in the four-factor alpha as *Uindex* increases from the 25th to the 75th percentile. Note that we do not include expense ratio as a control variable in these models. Therefore, this coefficient estimate reflects the joint effect of managerial skill and management fees of unique funds. It suggests that more unique funds outperform less unique funds in net-of-fee alphas, despite the fact that they generally have higher expenses.

In Column (3) of both panels, we repeat the test in Column (2) by adding the total expense ratio as an additional control variable. This allows us to isolate the effect of fund uniqueness on net performance by removing the confounding effect of fees on net performance. Essentially, the coefficient on *Uindex* in Column (3) measures the marginal effect of fund uniqueness on fund performance holding constant the total expense ratio. Not surprisingly, this coefficient is more significant, both economically and statistically, than the same coefficient in Column (2).

Columns (4) to (6) repeat the tests in Column (3) using alphas estimated from the (Fama and French (1992)) three-factor model and the one-factor (CAPM) model, as well as the raw excess return, as the performance measures. With the exception of the last column in Panel

B, which shows an insignificant effect, these results generally confirm a significantly positive relation between fund uniqueness and fund performance.

Several models in Table 4 also show that funds with more volatile returns in the past tend to have a lower alpha, which is puzzling, and is potentially related to the idiosyncratic volatility puzzle in equity returns documented by Ang, Hoddrick, Xing, and Zhang (2006). The first two columns also show that funds with higher load fees tend to underperform. This effect becomes insignificant after we control for the total expense ratio, perhaps because of the strong positive correlation between total expense and load fee ratios (0.31 as reported in Panel B of Table 1).

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, suggesting that fund managers are able to capitalize on their ability to employ innovative investment strategies. In addition, unique funds deliver better net-of-fee performance, which serves as a compensation for the high search costs paid by their investors. Together, the higher fees and better net performance demonstrate that unique funds have better managerial skill than do funds that are otherwise similar.

4 Flow Sensitivity, Performance Persistence, and Fund Uniqueness

We now investigate the effects of fund uniqueness on the flow-performance sensitivity and performance persistence, and test two closely related hypotheses: **H2** and **H3**. According to these hypotheses, as fund uniqueness increases, fund flows should become less sensitive to fund performance, and fund performance should become more persistent.

4.1 Are flows less sensitive to the performance of unique funds?

To test the effect of fund uniqueness on the sensitivity of fund flows to performance, we regress the average fund flow in each quarter on the lagged fund performance, measured by the four-factor alpha, the lagged measure of fund uniqueness, and the interaction of these two variables. 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.

We include the square term of alpha, *AlphaSQ*, to account for the nonlinearity in the flow-performance relation, and control for other potential determinants of fund flows. We estimate our models at the quarterly frequency, where the fund flows in each quarter are measured by the average monthly flows within the quarter. As in the previous tables, we conduct our analysis using both the panel regression approach and the Fama-MacBeth approach.

Table 5 presents results for two alternative model specifications. Both sets of results are strongly supportive of the conjecture that fund uniqueness reduces the sensitivity of fund flows to past performance. In Columns (1) and (3), we only allow the sensitivity to past performance vary with fund uniqueness. In Columns (2) and (4), we also allow it to vary with fund return volatility and fund age. In all the four columns, the coefficient on the interaction term *Alpha*Uindex* is strongly negative, while the coefficient on *Alpha* itself is strongly positive. In Columns (1) and (3), the coefficient on *Alpha*Uindex* is about one half of the coefficient on *Alpha* in terms of the magnitude (1.552 vs. 3.059, 2.088 vs. 4.079). This suggests that as the fund uniqueness increases from the 25th to the 75th percentile, the sensitivity of fund flows to past performance declines by about 22%. This is an economically significant effect. Although the relative magnitude of the coefficient on *Alpha*Uindex* declines after we include two additional interaction terms in Columns (2) and (4), it is still highly significant, both economically and statistically.

The coefficients on both *Alpha*Vol* and *Alpha*Log(Age)* are significantly negative, 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), Huang, Wei, and Yan (2007), 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. The coefficient on *AlphaSQ* is significantly positive in three of the four columns, consistent with the previous finding of a convex flow-performance relation in the literature (see, for example, Sirri and Tufano (1998)).

4.2 Is the performance of unique funds more persistent?

To investigate the effect of fund uniqueness on performance persistence, we regress the average fund performance over a 12-month holding period on lagged performance and uniqueness index, as well as their interaction, all measured over the prior 36 months. The holding-period performance is measured either by the four-factor, three-factor, or one-factor alpha, or by the excess raw return. All alphas are calculated using the betas estimated over the prior 36 months. The past performance is always measured by the four-factor alpha. The regressions are run at the quarterly frequency. Our main interest is in the coefficient on the interact term *Alpha*Uindex*. If the performance of unique funds is more persistent, this coefficient should be positive.

The two panels of Table 6 present results from the panel and Fama-MacBeth regressions, respectively. Column (1) of both panels 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% level, suggesting that performance is much more persistent among unique funds.

In the remaining columns, 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. As the results in Columns (2) of both panels show, adding these controls has virtually no effect on the coefficient on *Alpha*Uindex*. Changing the performance measure for the holding period does not change the basic conclusion either. The coefficient on *Alpha*Uindex* is consistently positive in all columns. Interestingly, in all the five columns, the coefficient on *Alpha* is negative, although mostly statistically insignificant. This indicates that for the least unique funds (with *Uindex* close to zero), there is a weak tendency of performance reversal instead of persistence.

The strong positive coefficient on *Alpha*Uindex* suggests a large benefit for investors to invest in unique funds with good past performance. Take the results in Column (1) of Panel B as an example. The estimated coefficients suggest that for funds with a $Uindex$ one standard deviation above the mean (0.73), a one standard deviation increase in past four-factor alpha (0.47) predicts an increase of 94 basis points in the four-factor alpha over the next 12 months: (0.292*0.73-0.046)*47*12= 94. This is an economically significant effect.

In summary, the results in this section strongly support our conjectures that fund flows are less sensitive the performance and performance is more persistent for funds with more unique investment strategies. This is consistent with the Berk and Green (2004) equilibrium of the mutual fund industry, in which performance persistence is eliminated by performance-chasing fund flows due to diseconomies of scale.

5 Asymmetric Effects of Fund Uniqueness

We now examine potential asymmetries in the effects of fund uniqueness on the flow-performance sensitivity and performance persistence. According to the hypothesis **H4**, both the dampening effect of fund uniqueness on the flow-performance sensitivity and the amplifying effect of fund uniqueness on performance persistence should be stronger for underperforming than for outperforming funds. In contrast, if the higher performance persistence of unique funds is due to superior investment strategies not mimicked by competitors, we should expect the amplifying effect of fund uniqueness on performance persistence to be stronger among well-performing funds.

5.1 Asymmetric effects of fund uniqueness on fund flow sensitivity

To capture the potential asymmetry in the effect of fund uniqueness on the sensitivities of fund flows to good and poor performance, we create a three-way interaction term, *Alpha*Above*Uindex*, where *Alpha* is the four-factor alpha estimated over past 36 months, *Above* is a dummy variable that equals one if *Alpha* is above the cross-sectional median and zero otherwise, and *Uindex* is the uniqueness index estimated over past 36 months. We use this three-way interaction variable, together with the two-way interactions of the three variables, and the three variables themselves, to predict fund flows in each quarter. We include the same set of control variables as in Table 5 except the square term of alpha, *AlphaSQ*.[7] The results, estimated using both the fixed-effect model and the Fama-French approach, are reported in Panel A of Table 7.

In this extended model, the coefficient on the interaction term *Alpha*Above* picks up the difference in the sensitivities of fund flows to above- and below-median performance for funds with a uniqueness index equal to zero. A positive coefficient indicates convexity and a negative coefficient indicates concavity. Interestingly, this coefficient is insignificant in all columns in Panel A of Table 7, suggesting that for the least unique funds (*Uindex* close to zero), the flow-performance relation is largely symmetric. The coefficient on the three-way interaction term *Alpha*Above*Uindex*, is positive in all four columns, and is statistically significant in three of them. This suggests that fund uniqueness increases the convexity of the flow-performance relation. In fact, since the convexity is insignificant for the least unique funds, our results demonstrate that the convexity in the flow-performance relation well-documented in the literature comes mainly from funds with a high degree of uniqueness.

Another way to interpret this result is that, fund uniqueness reduces the sensitivity of fund flow to performance, as indicated by the significantly negative coefficient on *Alpha*Uindex* in all columns of the table. However, this dampening effect is weaker for funds with above-median performance. As a result, for funds that are sufficiently unique, fund flows respond to good performance more strongly than they respond to poor performance. This asymmetry in the dampening effect of fund uniqueness cannot be explained by the lack of a broad investor base, or the difficulty in evaluating the performance of unique funds, as these features of unique funds should affect the sensitivity of fund flows to both good and bad performance. It supports the conjecture that the lack of close substitutes allows unique funds to retain investors after poor performance.

Notably, the coefficient on *Alpha*Above*Uindex* is smaller than the coefficient on *Alpha*Uindex* in magnitude across all columns, which implies that the sensitivity of fund flow

[7] We drop *AlphaSQ* because the nonlinearity in the flow-performance relation is now captured by a piecewise linear structure.

to good performance is also reduced by fund uniqueness. Therefore, the sensitivity of fund flow to performance is generally weaker for unique funds, which is to be expected given the difficulty of evaluating such funds and their lack of a broad investor base.

5.2 Asymmetric effects of fund uniqueness on performance persistence

As in our tests for the asymmetry of the fund uniqueness effect on fund flow sensitivity, we use the three-way interaction variable, *Alpha*Above*Uindex*, to test for potential differential effects of fund uniqueness on the persistence of good and poor performance. We extend the full model in Table 6 by adding this interaction term, as well as *Above*, *Alpha*Above* and *Above*Uindex*, as predictive variables for future performance. Future performance is measured by either the average four- or three-factor alpha, or the average raw excess return, over twelves months. The results are reported in Panel B of Table 7.

As in Table 6, 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 four of them. This 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 is more associated with the persistence of poor performance rather than 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 most significantly the response of fund flows to poor performance. Since unique funds are better able to retain investors after poor performance due to the lack of substitutes, the equilibrating downward adjustment of fund size is slowed down. Therefore, unique funds are less disciplined by competition pressures, and their poor performance persist for a longer time. In summary, the results in this section provide support for our hypothesis **H4**, and the equilibrating mechanism of the mutual fund industry modeled by Berk and Green (2004).

6 Robustness Checks

We conduct several robustness checks of our results. First, we control for active share developed by Cremers and Petajisto (2009). Second, we use two alternative measures of fund uniqueness. Third, we control for fund fixed effects. Fourth, we exclude sector funds from our analysis.

6.1 Controlling for active share

Cremers and Petajisto (2009) and Petajisto (2013) show that funds with higher active share have better performance and exhibit stronger performance persistence. Due to the strong positive correlation between our measures of fund uniqueness and active share, we have not included active share in our baseline models in order to avoid problems arising from multicollinearity. However, it is an interesting question whether our results are still significant after controlling for active share.

For this purpose, we add active share to our baseline models in Tables 3 to 7. For the fund flow and performance persistence analysis, we also allow for an interaction effect between active share and alpha. The results are summarized in Panel A of Table 8. To save space, we only report results from the Fama-MacBeth regressions. As one can see, our results are largely unaffected by adding these controls. In fact, fund uniqueness appears to be more significant in predicting both management fees and net-of-fee performance, compared to active share. Also, it has a more significant effect on the flow-performance relation and performance persistence. This is remarkable, given that $Uindex$ is constructed using only return data. It suggests that the fund uniqueness effects we find are not simply due to the correlation between active share and fund uniqueness.[8]

6.2 Alternative uniqueness measures

We rerun our main regressions using two alternative fund uniqueness measures: the PCA-Based Uniqueness Index and the SDI of Sun, Wang, and Zheng (2012). The main results are summarized in Panels B and C of 8, respectively. These results are quite similar to those reported in Tables 3 to 7. High uniqueness is associated with higher management fees, higher net-of-fee performance, lower flow-performance sensitivity, and higher performance persistence. Furthermore, the effects of fund uniqueness on fund flow sensitivity and performance persistence appear to be stronger for underperforming funds than for outperforming funds.

6.3 Further tests

Since fund uniqueness is highly persistent, our tests focus mainly on variation across funds. Therefore, we control for time fixed effects but not fund fixed effects in panel regressions. As a robustness check, we rerun our tests in Tables 3 to 7, controlling for both time and fund fixed effects. The results remain largely unchanged, except that the positive relation between fund uniqueness and management fees becomes insignificant. This loss of significance is not very surprising, because both fees and fund uniqueness are highly persistent. Once we remove the fund fixed effects, the variation may be too small to reveal a significant effect.

Sector funds tend to be more unique than other funds. As a result, one may wonder whether the effects we identify are mainly driven by sector funds. To check this possibility, we redo our analysis excluding all sector funds. The sample size decreases by about 11%, but the results are largely unchanged. This shows that our findings are not simply driven by sector funds.

To save space, the results of these additional tests are tabulated.

[8]Without controlling for fund uniqueness, we find that higher active share significantly predicts higher alpha, and increases performance persistence, consistent with the findings of Cremers and Petajisto (2009) and Petajisto (2013). This suggests that the lack of statistical significance in the coefficients on AS and $Alpha*AS$ in Columns (2) and (4) are mainly due to multicollinearity.

7 Conclusion

Based on the cluster analysis of historical returns, we construct two measures of the uniqueness of a mutual fund's investment strategy. One considers the number of steps needed to separate a fund from other funds, the other considers the number of funds in a fund's clusters at various levels of fund classification granularity. We find funds that are more unique have higher management fee and total expense ratios. Despite the higher fees, unique funds deliver better future net-of-fee performance than do funds that are otherwise similar. More importantly, fund flows are less sensitive to the performance of unique funds than to the performance non-unique funds, and the performance of unique funds is more persistent. 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.

These results suggest that unique funds have better managerial skill than do funds that are otherwise similar, and that managers are able to capitalize on their ability to employ unique investment strategies. Part of the benefits from unique investment strategies is shared by investors, which potentially serves as a compensation for high search costs associated with such funds. The lack of close substitutes makes investors less likely to move money out of underperforming unique funds, generating a convexity in the flow-performance relation. Since the equilibrating downward adjustment of fund size is slowed down, the underperformance of unique funds is more persistent than that of non-unique funds.

Our study demonstrates rich interaction between competitions on the asset side (for return generation) and the liability side (for fund flows) of a mutual fund's balance sheet. Fund uniqueness created by innovative investment of fund assets can not only strengthen a fund's position in generating and sustaining good performance, but also improve a fund's ability in retaining fund investors after poor performance. In other words, unique funds are less disciplined by competition pressures imposed by fund flows. This in turn makes their poor performance more persistent.

Our measures of fund uniqueness are derived purely from realized fund returns. This makes it very straightforward to compute. However, it also leaves out other potentially useful sources of information. In principle, one can use cluster analysis to construct measures of fund uniqueness based on other fund characteristics, such as portfolio holdings. We leave this as an interesting topic for future research.

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. *Nsub* is the number of subclusters that a fund belongs to in a given period (in addition to its style group) based on HCA. *Uindex* is the HCA-Based Uniqueness Index, which is 1 minus the normalized *Nsub*. *Uindex(PCA)* is the PCA-Based Uniqueness Index. *SDI* is the strategy distinctiveness index developed by Sun, Wang, and Zheng (2012). *AS* the active share measure developed by Cremers and Petajisto (2009). *Flow* is the average monthly flow in a given quarter, calculated as the difference between a fund's TNA growth rate and realized returns. Alpha is the Carhart (1997) four-factor alpha calculated over a rolling window of 36 months. *Vol*, *IVol*, and *SVol* are total excess return volatility, idiosyncratic and systemic volatilities 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 millions of year 2009 dollars. *Age* is the number of years since fund inception.

A. Fund Characteristics

	Mean	SD	Min	P(25)	P(50)	P(75)	Max	N
#Fund per period	1785.96	544.83	662.00	1260.00	1965.50	2277.00	2447.00	94
Nsub	119.93	83.97	0.00	47.00	108.00	181.00	411.00	115674
Uindex	0.45	0.28	0.00	0.22	0.43	0.68	1.00	115674
Uindex(PCA)	0.43	0.26	0.00	0.21	0.40	0.64	1.00	118589
SDI	0.08	0.10	0.00	0.03	0.05	0.10	1.85	115674
AS	0.80	0.15	0.00	0.70	0.84	0.92	1.00	134690
Flow(% p.m.)	0.71	4.93	-18.51	-1.39	-0.22	1.56	34.81	167528
Alpha(% p.m.)	-0.06	0.47	-12.06	-0.27	-0.08	0.13	4.49	123233
Vol(% p.m.)	5.17	2.20	1.82	3.63	4.85	6.22	13.44	148050
IVol(% p.m.)	2.20	1.42	0.49	1.18	1.82	2.78	7.86	115674
SVol(% p.m.)	4.43	1.83	1.09	3.02	4.25	5.61	10.01	115674
Expense(% p.a.)	1.29	0.48	0.17	0.99	1.24	1.53	2.96	165182
MgmtFee(% p.a.)	0.70	0.38	-0.98	0.56	0.75	0.91	1.66	133727
Load(%)	1.97	2.14	0.00	0.00	1.00	4.00	10.41	167880
InstRatio	0.20	0.34	0.00	0.00	0.00	0.21	1.00	167880
Turnover(p.a.)	0.89	0.88	0.03	0.34	0.65	1.13	5.57	160375
Log(TNA)	5.30	1.82	-4.42	4.00	5.25	6.56	12.18	167366
Log(Age)	2.16	0.97	-2.48	1.59	2.22	2.78	4.50	167703

B. Correlations

	Uindex	Uindex(PCA)	SDI	AS	Flow	Alpha	Vol	IVol	SVol	Expense	MgtFee	Load	InstRatio	Turnover	Log(TNA)	Log(Age)
Uindex	1.00															
Uindex(PCA)	0.76	1.00														
SDI	0.48	0.39	1.00													
AS	0.48	0.43	0.22	1.00												
Flow	0.03	0.04	0.05	0.06	1.00											
Alpha	0.04	0.05	0.11	0.09	0.23	1.00										
Vol	0.12	0.21	-0.19	0.17	-0.03	0.03	1.00									
IVol	0.37	0.39	0.26	0.39	0.03	0.12	0.72	1.00								
SVol	-0.01	0.09	-0.40	0.03	-0.05	-0.05	0.93	0.43	1.00							
Expense	0.32	0.26	0.19	0.23	0.02	-0.11	0.19	0.31	0.07	1.00						
MgtFee	0.22	0.20	0.13	0.14	-0.14	0.02	0.11	0.21	0.05	0.36	1.00					
Load	0.02	0.01	0.03	0.02	-0.00	-0.02	-0.04	-0.01	-0.05	0.31	-0.08	1.00				
InstRatio	-0.19	-0.14	-0.15	-0.17	-0.04	0.00	0.03	-0.09	0.10	-0.28	-0.07	-0.28	1.00			
Turnover	0.15	0.16	0.09	0.07	0.02	-0.09	0.21	0.26	0.15	0.23	0.11	-0.03	-0.03	1.00		
Log(TNA)	-0.21	-0.12	-0.11	-0.18	-0.05	0.18	-0.09	-0.14	-0.03	-0.37	0.07	0.08	0.01	-0.17	1.00	
Log(Age)	-0.07	-0.06	-0.08	-0.10	-0.32	-0.04	-0.07	-0.16	-0.02	-0.15	0.10	0.10	-0.06	-0.10	0.42	1.00

Table 2: Determinants of fund uniqueness

This table shows the relation between fund uniqueness and lagged fund characteristics. Three alternative measures of fund uniqueness are calculated using monthly returns in a rolling-window of 36 months: the HCA-based *Uindex*, the PCA-based *Uindex(PCA)*, and the *SDI*. Independent variables are measured at the end of the year before the rolling window. See the caption of Table 1 for the details of variable definitions. The first three columns are estimated using panel regressions with year fixed effects and error clustering at both fund and year levels. The last three columns are estimated using the Fama-MacBeth regressions (with t-statistics Newey-West adjusted for autocorrelation up to order three). The models are estimated at the annual frequency. Significance at the 1, 5, and 10 percent levels are indicated by ***, **, and *, respectively.

	(1) Uindex	(2) Uindex(PCA)	(3) SDI	(4) Uindex	(5) Uindex(PCA)	(6) SDI
SVol	-0.015***	-0.015**	-0.008***	-0.013	-0.014	-0.007**
	(-3.01)	(-2.26)	(-2.85)	(-0.82)	(-0.79)	(-2.14)
IVol	0.010	0.029***	0.004*	0.007	0.033**	-0.000
	(1.56)	(4.80)	(1.69)	(0.52)	(2.17)	(-0.07)
Log(Age)	0.014***	0.005	0.004***	0.021***	0.009*	0.003**
	(3.59)	(1.17)	(2.70)	(4.02)	(1.92)	(2.21)
Log(TNA)	-0.007***	0.001	-0.001**	-0.004	0.012	-0.001**
	(-2.82)	(0.51)	(-2.27)	(-1.02)	(1.11)	(-2.39)
Alpha	0.010	-0.005	0.004	0.010	0.006	-0.003
	(1.04)	(-0.69)	(0.82)	(0.48)	(0.37)	(-1.34)
Expense	0.065***	0.062***	0.010***	0.062***	0.065***	0.008**
	(8.80)	(7.91)	(4.44)	(10.16)	(3.68)	(2.72)
Load	-0.006***	-0.006***	-0.001***	0.022	-0.010**	-0.002*
	(-4.70)	(-5.17)	(-3.31)	(0.80)	(-2.10)	(-2.02)
InstRatio	-0.044***	-0.040***	-0.007***	-0.011	-0.043	0.003
	(-4.59)	(-3.67)	(-4.46)	(-0.82)	(-1.18)	(0.54)
Turnover	-0.001	0.007**	0.003	-0.017	-0.011	0.001
	(-0.33)	(2.10)	(1.14)	(-1.14)	(-0.68)	(0.36)
AS	0.415***	0.321***	0.042***	0.451***	0.368***	0.045***
	(10.75)	(9.99)	(4.15)	(11.65)	(7.08)	(3.80)
$\text{Uindex}_{t\text{-}1}$	0.399***			0.359***		
	(19.78)			(16.09)		
$\text{Uindex(PCA)}_{t\text{-}1}$		0.303***			0.320***	
		(9.77)			(5.13)	
$\text{SDI}_{t\text{-}1}$			0.360***			0.439***
			(5.21)			(9.87)
Constant	-0.099***	-0.033	0.034***	-0.130**	-0.186	0.020
	(-3.00)	(-0.75)	(3.00)	(-2.62)	(-1.68)	(1.37)
Observations	17829	18459	17829	17829	18459	17829
R^2	0.409	0.330	0.419	0.340	0.297	0.245

Table 3: Fund uniqueness and fees

This table presents the regression results for fund expense and management fee ratios. See the caption of Table 1 for the details of variable definitions. All independent variables are lagged by one year. The first two columns are estimated using panel regressions with year fixed effects and error clustering at both fund and year levels. The last two columns are estimated using the Fama-MacBeth regressions (with t-statistics Newey-West adjusted for autocorrelation up to order three). The models are estimated at the annual frequency. Significance at the 1, 5, and 10 percent levels are indicated by ***, **, and *, respectively.

	Fixed effects		Fama-MacBeth	
	(1)	(2)	(3)	(4)
	MgtFee	Expense	MgtFee	Expense
Uindex	0.209***	0.244***	0.192***	0.231***
	(9.16)	(10.00)	(5.51)	(9.23)
Vol	0.024***	0.044***	0.033***	0.055***
	(5.68)	(9.35)	(5.21)	(9.28)
Log(Age)	-0.023**	-0.019	-0.023**	-0.016
	(-2.44)	(-1.56)	(-2.30)	(-1.10)
Log(TNA)	0.006	-0.080***	0.004	-0.084***
	(1.28)	(-17.88)	(0.58)	(-18.82)
Turnover	0.022***	0.044***	0.022***	0.047***
	(4.21)	(6.50)	(4.85)	(6.11)
InstRatio	-0.011	-0.332***	0.000	-0.237***
	(-0.82)	(-13.31)	(0.01)	(-3.69)
Constant	0.568***	1.319***	0.494***	1.428***
	(11.40)	(22.85)	(5.02)	(17.06)
Observations	25941	28922	25941	28922
R^2	0.072	0.334	0.050	0.284

Table 4: Fund uniqueness and performance

This table presents regression results on the net-of-fee performance. *Alpha4F3* is the average Carhart (1997) four-factor alpha over three months. *Alpha4F12*, *Alpha3F12*, *Alpha1F12*, and *ExRet* are the average Carhart (1997) four-factor alpha, the Fama and French (1992) three-factor alpha, the market-factor alpha, and the excess raw return over 12 months, respectively. See the caption of Table 1 for the definitions of other variables. All independent variables are lagged by one quarter. Results in Panel A are estimated with quarter fixed effects and error clustering at both fund and quarter levels. Results in Panel B are estimated using the Fama-MacBeth approach (with t-statistics Newey-West adjusted for autocorrelation up to order three). The models are estimated at the quarterly frequency. Significance at the 1, 5, and 10 percent levels are indicated by ***, **, and *, respectively.

A. Panel regressions with quarter fixed effects

	(1)	(2)	(3)	(4)	(5)	(6)
	Alpha4F3	Alpha4F12	Alpha4F12	Alpha3F12	Alpha1F12	ExRet12
Uindex	0.111**	0.095***	0.119***	0.091***	0.188***	0.117**
	(2.01)	(3.31)	(4.04)	(2.95)	(4.26)	(2.29)
Vol	-0.048	-0.055***	-0.051***	-0.034**	-0.074**	-0.042
	(-1.49)	(-4.51)	(-4.16)	(-2.14)	(-2.49)	(-1.15)
Expense			-0.087***	-0.077***	-0.079***	-0.065**
			(-4.05)	(-3.98)	(-3.05)	(-2.32)
Load	-0.008***	-0.008***	-0.003	-0.003	-0.004	-0.005
	(-2.83)	(-3.33)	(-0.82)	(-1.08)	(-1.14)	(-1.24)
Log(Age)	-0.007	-0.006	-0.010	-0.020*	-0.023*	-0.002
	(-0.63)	(-0.56)	(-0.97)	(-1.83)	(-1.72)	(-0.12)
Log(TNA)	0.006	0.006	-0.001	0.010*	-0.008	-0.015
	(0.83)	(1.14)	(-0.19)	(1.84)	(-1.07)	(-1.43)
Turnover	-0.020	-0.024	-0.020	0.011	0.000	-0.017
	(-0.96)	(-1.64)	(-1.37)	(0.65)	(0.02)	(-0.77)
InstRatio	0.005	0.004	-0.017	-0.022	0.008	0.006
	(0.25)	(0.25)	(-1.08)	(-1.50)	(0.39)	(0.28)
Constant	0.152	-0.039	0.078	-0.040	0.046	0.611***
	(1.16)	(-0.65)	(1.29)	(-0.59)	(0.39)	(3.88)
Observations	114783	107719	107719	107719	107719	107719
R^2	0.062	0.075	0.076	0.071	0.102	0.740

B. Fama-MacBeth regressions

	(1)	(2)	(3)	(4)	(5)	(6)
	Alpha4F3	Alpha4F12	Alpha4F12	Alpha3F12	Alpha1F12	ExRet12
Uindex	0.097*	0.078**	0.100***	0.062*	0.168***	0.079
	(1.76)	(2.37)	(3.02)	(1.76)	(2.85)	(1.09)
Vol	-0.039*	-0.052***	-0.046**	-0.021	-0.087	0.011
	(-1.69)	(-2.74)	(-2.44)	(-0.85)	(-1.49)	(0.22)
Expense			-0.106***	-0.098***	-0.085***	-0.093***
			(-3.94)	(-4.30)	(-2.85)	(-2.67)
Load	-0.006**	-0.007***	-0.001	-0.001	-0.001	-0.000
	(-2.00)	(-2.74)	(-0.15)	(-0.38)	(-0.24)	(-0.10)
Log(Age)	-0.010	-0.009	-0.015	-0.020	-0.028*	-0.017
	(-0.85)	(-0.95)	(-1.46)	(-1.54)	(-1.81)	(-1.10)
Log(TNA)	0.011	0.009	-0.000	0.010	-0.001	-0.004
	(1.60)	(1.65)	(-0.04)	(1.35)	(-0.06)	(-0.31)
Turnover	-0.009	-0.016	-0.011	0.021	0.020	0.018
	(-0.30)	(-0.63)	(-0.41)	(0.62)	(0.48)	(0.42)
InstRatio	0.001	0.005	-0.003	0.013	0.031	0.028
	(0.03)	(0.25)	(-0.17)	(0.62)	(1.54)	(1.23)
Constant	-0.013	0.036	0.188**	0.056	0.394	0.524*
	(-0.12)	(0.40)	(2.15)	(0.57)	(1.60)	(1.96)
Observations	114783	107719	107719	107719	107719	107719
R^2	0.002	0.010	0.012	0.010	0.004	0.002

Table 5: Fund uniqueness and fund flow sensitivity to performance

This table presents regression results on the flow-performance sensitivity. The dependent variable is the average monthly flow within a quarter. *AlphaSQ* is the square of the four-factor alpha. See the caption of Table 1 for the definitions of other variables. All independent variables are lagged by one quarter. The first two columns are estimated using panel regressions with quarter fixed effects and clustering of errors at both the fund and quarter levels. The last two columns are estimated using the Fama-MacBeth approach (with t-statistics Newey-West adjusted for autocorrelation up to order three). The models are estimated at the quarterly frequency. Significance at the 1, 5, and 10 percent levels are indicated by ***, **, and *, respectively.

	Fixed effects		Fama-MacBeth	
	(1)	(2)	(3)	(4)
	Flow	Flow	Flow	Flow
Alpha * Uindex	-1.552***	-1.442***	-2.088***	-1.862***
	(-5.27)	(-6.08)	(-7.01)	(-6.45)
Alpha * Vol		-0.302***		-0.296***
		(-11.64)		(-4.18)
Alpha * Log(Age)		-0.407***		-0.578***
		(-5.27)		(-5.58)
Uindex	0.072	-0.040	-0.156	-0.201
	(0.57)	(-0.37)	(-1.02)	(-1.44)
Alpha	3.059***	6.023***	4.079***	6.987***
	(12.61)	(19.00)	(15.06)	(14.43)
AlphaSQ	0.128	0.163**	0.661***	0.650***
	(1.45)	(2.54)	(3.96)	(3.78)
Vol	-0.130***	-0.118***	-0.028	-0.031
	(-3.94)	(-4.28)	(-0.46)	(-0.51)
Log(Age)	-0.372***	-0.394***	-0.366***	-0.400***
	(-11.06)	(-11.34)	(-9.85)	(-10.53)
Log(TNA)	0.004	-0.007	-0.000	-0.005
	(0.28)	(-0.51)	(-0.02)	(-0.41)
Expense	-0.084	-0.108	-0.050	-0.064
	(-1.14)	(-1.53)	(-0.58)	(-0.74)
Load	0.017	0.021**	0.025***	0.028***
	(1.59)	(2.10)	(2.69)	(2.96)
Turnover	0.146***	0.165***	0.159***	0.164***
	(3.54)	(4.00)	(3.30)	(3.34)
InstRatio	-0.223***	-0.226***	-0.066	-0.075
	(-3.07)	(-3.13)	(-0.51)	(-0.60)
Constant	2.083***	2.231***	1.250***	1.405***
	(10.80)	(12.01)	(3.63)	(4.12)
Observations	114037	114037	114037	114037
R^2	0.090	0.101	0.060	0.070

Table 6: Fund uniqueness and performance persistence

This table presents results on the effects fund uniqueness on performance persistence. *Alpha4F12*, *Alpha3F12*, *Alpha1F12*, and *ExRet* are the average Carhart (1997) four-factor alpha, the Fama and French (1992) three-factor alpha, the market-factor alpha, and the excess raw return over 12 months, respectively. See the caption of Table 1 for the definitions of other variables. All independent variables are lagged by one quarter. Results in Panel A are estimated with quarter fixed effects and error clustering at both fund and quarter levels. Results in Panel B are estimated using the Fama-MacBeth approach (with t-statistics Newey-West adjusted for autocorrelation up to order three). The models are estimated at the quarterly frequency. Significance at the 1, 5, and 10 percent levels are indicated by ***, **, and *, respectively.

A. Panel regressions with quarter fixed effects

	(1)	(2)	(3)	(4)	(5)
	Alpha4F12	Alpha4F12	Alpha3F12	Alpha1F12	ExRet12
Alpha * Uindex	0.308***	0.313***	0.180*	0.218**	0.381***
	(3.34)	(3.20)	(1.92)	(2.03)	(2.92)
Alpha * Vol		-0.029*	-0.011	-0.017	-0.068***
		(-1.89)	(-0.67)	(-0.91)	(-2.75)
Alpha * Log(Age)		0.049	0.081**	0.121***	0.167***
		(1.31)	(2.15)	(2.69)	(3.28)
Uindex	0.029	0.111***	0.076**	0.175***	0.124**
	(1.03)	(3.72)	(2.48)	(4.04)	(2.37)
Alpha	-0.085	-0.015	-0.096	-0.175	-0.215
	(-1.07)	(-0.10)	(-0.66)	(-1.06)	(-1.19)
Vol		-0.049***	-0.033**	-0.073**	-0.036
		(-4.06)	(-2.08)	(-2.47)	(-1.07)
Log(Age)		0.004	-0.002	-0.002	0.012
		(0.55)	(-0.25)	(-0.17)	(0.93)
Log(TNA)		-0.007	0.003	-0.015**	-0.015*
		(-1.63)	(0.63)	(-2.56)	(-1.93)
Expense		-0.079***	-0.065***	-0.066***	-0.061**
		(-3.99)	(-3.84)	(-2.76)	(-2.37)
Load		-0.003	-0.003	-0.004	-0.006
		(-0.93)	(-1.15)	(-1.23)	(-1.49)
Turnover		-0.014	0.019	0.009	-0.014
		(-0.90)	(1.09)	(0.39)	(-0.60)
InstRatio		-0.014	-0.017	0.014	0.007
		(-0.95)	(-1.23)	(0.69)	(0.32)
Constant	-0.217***	0.062	-0.059	0.016	0.556***
	(-16.83)	(1.00)	(-0.84)	(0.14)	(3.81)
Observations	108541	107719	107719	107719	107719
R^2	0.066	0.083	0.078	0.108	0.743

B. Fama-MacBeth regressions

	(1)	(2)	(3)	(4)	(5)
	Alpha4F12	Alpha4F12	Alpha3F12	Alpha1F12	ExRet12
Alpha * Uindex	0.292***	0.291***	0.174*	0.209**	0.216**
	(2.80)	(2.94)	(1.92)	(2.14)	(2.41)
Alpha * Vol		-0.019	0.006	-0.017	-0.011
		(-1.01)	(0.31)	(-0.48)	(-0.39)
Alpha * Log(Age)		0.070***	0.080***	0.127***	0.128***
		(2.82)	(3.80)	(5.23)	(5.05)
Uindex	0.037	0.103***	0.053	0.164***	0.089
	(0.97)	(3.35)	(1.56)	(2.66)	(1.16)
Alpha	-0.046	-0.081	-0.180	-0.243	-0.308*
	(-0.46)	(-0.60)	(-1.30)	(-1.51)	(-1.84)
Vol		-0.037*	-0.005	-0.063	0.036
		(-1.85)	(-0.20)	(-0.99)	(0.63)
Log(Age)		0.008	0.003	-0.003	0.005
		(1.03)	(0.37)	(-0.22)	(0.37)
Log(TNA)		-0.008*	0.002	-0.005	-0.006
		(-1.68)	(0.39)	(-0.90)	(-0.77)
Expense		-0.095***	-0.083***	-0.071***	-0.081**
		(-3.79)	(-4.05)	(-2.65)	(-2.53)
Load		-0.001	-0.003	-0.002	-0.002
		(-0.29)	(-0.71)	(-0.57)	(-0.44)
Turnover		-0.002	0.024	0.023	0.018
		(-0.10)	(0.82)	(0.65)	(0.50)
InstRatio		-0.001	0.015	0.029	0.028
		(-0.03)	(0.75)	(1.60)	(1.35)
Constant	-0.147***	0.120	-0.047	0.228	0.343
	(-4.72)	(1.23)	(-0.37)	(0.86)	(1.20)
Observations	108541	107719	107719	107719	107719
R^2	0.004	0.017	0.008	0.009	0.006

Table 7: Asymmetric effects of fund uniqueness

This table shows the asymmetric effects of fund uniqueness on the flow sensitivity (Panel A) and performance persistence (Panel B) of outperforming and underperforming funds. *Flow* is the average monthly flow within a quarter. *Alpha4F12*, *Alpha3F12*, and *ExRet* are the average Carhart (1997) four-factor alpha, the Fama and French (1992) three-factor alpha, and the excess raw return over 12 months, respectively. *Above* is a dummy variable equal 1 if the lagged four-factor alpha is above the contemporaneous sample median. See the caption of Table 1 for the definitions of other variables. All independent variables are lagged by one quarter. The models are estimated using panel regressions (with quarter fixed effects and clustering of errors at both the fund and quarter levels), and the Fama-MacBeth approach (with Newey-West adjustment for autocorrelation up to order three) at the quarterly frequency. Significance at the 1, 5, and 10 percent levels are indicated by ***, **, and *, respectively.

A. Asymmetric effects on performance sensitivity of fund flows

	Fixed effects		Fama-MacBeth	
	(1)	(2)	(3)	(4)
	Flow	Flow	Flow	Flow
Alpha * Uindex	-2.213***	-2.043***	-3.685***	-3.199***
	(-5.99)	(-6.17)	(-8.48)	(-7.90)
Above * Uindex	0.719***	0.414**	0.854***	0.603***
	(4.32)	(2.43)	(5.09)	(3.31)
Alpha * Above * Uindex	1.371**	1.188**	1.487*	1.161
	(2.15)	(2.14)	(1.76)	(1.50)
Alpha * Above	-0.121	-0.075	0.632	0.806
	(-0.29)	(-0.21)	(1.18)	(1.59)
Uindex	-0.561***	-0.458***	-0.896***	-0.762***
	(-3.50)	(-3.02)	(-4.23)	(-3.70)
Alpha	2.511***	5.498***	3.877***	6.609***
	(8.53)	(15.07)	(11.33)	(11.58)
Above	0.316***	0.194**	-0.008	-0.055
	(3.48)	(2.53)	(-0.10)	(-0.71)
Alpha * Vol		-0.266***		-0.249***
		(-9.93)		(-3.63)
Alpha * Log(Age)		-0.418***		-0.593***
		(-5.50)		(-5.67)
Vol	-0.128***	-0.118***	-0.032	-0.033
Log(Age)	-0.372***	-0.396***	-0.362***	-0.401***
Log(TNA)	-0.001	-0.008	-0.004	-0.008
Expense	-0.090	-0.110	-0.060	-0.073
Load	0.019*	0.022**	0.027***	0.028***
Turnover	0.142***	0.159***	0.151***	0.156***
InstRatio	-0.216***	-0.221***	-0.064	-0.074
Constant	2.014***	2.190***	1.296***	1.446***
Observations	114037	114037	114037	114037
R^2	0.094	0.103	0.066	0.073

B. Asymmetric effects on performance persistence

	Fixed effects			Fama-MacBeth		
	(1)	(2)	(3)	(4)	(5)	(6)
	Alpha4F12	Alpha3F12	ExRet12	Alpha4F12	Alpha3F12	ExRet12
Alpha * Uindex	0.498***	0.393**	0.745***	0.480***	0.344**	0.500***
	(3.01)	(2.29)	(3.70)	(3.16)	(2.41)	(3.28)
Above * Uindex	-0.082	-0.043	-0.252***	-0.061	0.006	-0.071
	(-1.26)	(-0.64)	(-3.18)	(-1.14)	(0.11)	(-1.05)
Alpha * Above * Uindex	-0.291	-0.387*	-0.389	-0.390**	-0.434**	-0.503**
	(-1.43)	(-1.94)	(-1.58)	(-2.53)	(-2.47)	(-2.54)
Alpha * Above	0.039	0.023	-0.021	0.046	0.050	0.112
	(0.25)	(0.15)	(-0.11)	(0.30)	(0.35)	(0.56)
Uindex	0.214***	0.183***	0.351***	0.200***	0.128***	0.207*
	(3.67)	(3.00)	(3.84)	(4.74)	(3.01)	(1.98)
Alpha	0.011	-0.052	-0.211	-0.032	-0.133	-0.325
	(0.08)	(-0.37)	(-0.99)	(-0.22)	(-0.82)	(-1.64)
Above	0.010	-0.002	0.091**	0.019	-0.001	0.030
	(0.34)	(-0.04)	(2.18)	(0.86)	(-0.06)	(1.29)
Alpha * Vol	-0.031**	-0.011	-0.070***	-0.028	0.003	-0.019
	(-2.04)	(-0.76)	(-2.85)	(-1.34)	(0.13)	(-0.61)
Alpha * Log(Age)	0.037	0.063*	0.145***	0.063***	0.067***	0.119***
	(1.10)	(1.91)	(3.08)	(2.97)	(3.83)	(4.96)
Vol	-0.046***	-0.028	-0.030	-0.034*	-0.001	0.040
Log(Age)	0.003	-0.004	0.010	0.007	0.003	0.005
Log(TNA)	-0.007	0.003	-0.015*	-0.008*	0.002	-0.007
Expense	-0.078***	-0.063***	-0.058**	-0.098***	-0.083***	-0.084***
Load	-0.003	-0.004	-0.006	-0.001	-0.002	-0.002
Turnover	-0.012	0.022	-0.010	-0.001	0.025	0.019
InstRatio	-0.014	-0.018	0.006	0.000	0.015	0.027
Constant	0.030	-0.095	0.466***	0.091	-0.080	0.293
Observations	107719	107719	107719	107719	107719	107719
R^2	0.084	0.080	0.744	0.019	0.010	0.012

Table 8: Robustness Checks

This table presents results of several robustness checks on the effects of fund uniqueness on management fees, net-of-fee performance, flow-performance sensitivity, and performance persistence. In Panel A, we control for the effects of active share (*AS*). In Panel B, fund uniqueness is measured by the PCA-based uniqueness index. In Panel C, fund uniqueness is measured by the *SDI*. *Alpha4F12* is the average Carhart (1997) four-factor alpha over 12 months. *Above* is a dummy variable equal 1 if the lagged four-factor alpha is above the contemporaneous sample median. See the caption of Table 1 for the definitions of other variables. All independent variables are lagged by one quarter. The models are estimated at the quarterly frequency using the Fama-MacBeth approach (with t-statistics Newey-West adjusted for autocorrelation up to order three). Significance at the 1, 5, and 10 percent levels are indicated by ***, **, and *, respectively.

A. Controlling for active share

	(1) MgtFee	(2) Alpha4F12	(3) Flow	(4) Alpha4F12	(5) Flow	(6) Alpha4F12
Uindex	0.149***	0.098**	-0.325***	0.091**	-0.903***	0.205***
	(3.69)	(2.61)	(-3.25)	(2.12)	(-5.60)	(4.44)
Alpha * Uindex			-1.764***	0.229*	-3.140***	0.391**
			(-5.17)	(1.86)	(-8.62)	(2.48)
Above * Uindex					0.668***	-0.092
					(3.41)	(-1.59)
Alpha * Above * Uindex					0.952	-0.327**
					(1.18)	(-2.09)
Alpha * Above					1.089**	-0.051
					(2.14)	(-0.38)
Alpha			7.514***	-0.285	6.998***	-0.365
			(7.78)	(-1.25)	(7.11)	(-1.47)
Above					-0.075	0.035
					(-0.89)	(1.09)
AS	0.179**	0.194	-0.370	0.205	-0.371	0.225*
	(2.20)	(1.65)	(-0.81)	(1.60)	(-0.82)	(1.80)
Alpha * AS			-0.541	0.350	-0.361	0.556**
			(-0.57)	(1.11)	(-0.43)	(2.14)
Alpha * Vol			-0.228***	-0.037**	-0.198**	-0.048**
			(-2.92)	(-2.16)	(-2.60)	(-2.46)
Alpha * Log(Age)			-0.641***	0.081***	-0.678***	0.073***
			(-5.58)	(4.20)	(-6.11)	(4.24)
Vol	0.030***	-0.066***	0.003	-0.057**	0.001	-0.052**
Log(Age)	-0.029**	-0.015	-0.371***	0.006	-0.372***	0.005
Log(TNA)	0.004	-0.000	-0.012	-0.007	-0.014	-0.008
Expense		-0.088***	-0.007	-0.078***	-0.007	-0.080***
Load		-0.000	0.027**	-0.001	0.026**	-0.000
Turnover	0.018***	-0.008	0.152***	0.002	0.145***	0.004
InstRatio	0.007	-0.004	-0.096	0.001	-0.107	0.001
AlphaSQ			0.849***			
Constant	0.406***	0.108	1.558***	0.041	1.589***	-0.015
Observations	24169	94032	99907	94032	99907	94032
R^2	0.055	0.018	0.064	0.024	0.069	0.027

B. PCA-based uniqueness measure

	(1)	(2)	(3)	(4)	(5)	(6)
	MgtFee	Alpha4F12	Flow	Alpha4F12	Flow	Alpha4F12
Uindex(PCA)	0.242***	0.065*	-0.119	0.066**	-0.593**	0.192***
	(4.96)	(1.86)	(-0.65)	(2.14)	(-2.44)	(3.73)
Alpha * Uindex(PCA)			-1.760***	0.323***	-2.877***	0.551***
			(-5.45)	(2.80)	(-6.09)	(3.26)
Above * Uindex(PCA)					0.569***	-0.124*
					(3.38)	(-1.97)
Alpha * Above * Uindex(PCA)					0.771	-0.417**
					(1.11)	(-2.02)
Alpha * Above					1.048**	0.084
					(2.11)	(0.50)
Alpha			6.692***	-0.054	6.090***	-0.031
			(13.59)	(-0.41)	(10.50)	(-0.21)
Above					-0.010	0.043*
					(-0.14)	(1.87)
Alpha * Vol			-0.266***	-0.022	-0.208***	-0.030
			(-3.91)	(-1.22)	(-3.22)	(-1.47)
Alpha * Log(Age)			-0.575***	0.066***	-0.590***	0.056***
			(-5.65)	(2.69)	(-5.77)	(2.65)
Vol	0.036***	-0.045**	-0.038	-0.036*	-0.036	-0.033
Log(Age)	-0.014	-0.013	-0.407***	0.009	-0.410***	0.008
Log(TNA)	-0.006	-0.002	-0.004	-0.010*	-0.004	-0.010**
Expense		-0.101***	-0.072	-0.090***	-0.081	-0.093***
Load		-0.001	0.025***	-0.001	0.026***	-0.001
Turnover	-0.024	-0.010	0.168***	-0.002	0.160***	-0.001
InstRatio	-0.003	-0.009	-0.096	-0.005	-0.093	-0.005
AlphaSQ			0.637***			
Constant	0.558***	0.208**	1.429***	0.140	1.387***	0.099
Observations	26439	107719	115692	107719	115692	107719
R^2	0.036	0.011	0.069	0.016	0.072	0.018

C. Fund uniqueness measured by SDI

	(1) MgtFee	(2) Alpha4F12	(3) Flow	(4) Alpha4F12	(5) Flow	(6) Alpha4F12
SDI	0.669***	0.421***	-1.188***	0.343***	-2.332***	0.825***
	(3.19)	(3.29)	(-2.91)	(3.10)	(-3.07)	(3.45)
Alpha * SDI			-7.386***	0.994**	-11.182***	1.760***
			(-7.92)	(2.52)	(-6.25)	(3.46)
Above * SDI					0.321	-0.352
					(0.31)	(-1.17)
Alpha * Above * SDI					5.509***	-1.483**
					(2.90)	(-2.57)
Alpha * Above					1.157***	-0.036
					(4.27)	(-0.27)
Alpha			7.623***	-0.135	6.839***	-0.039
			(14.78)	(-0.92)	(12.34)	(-0.30)
Above					0.201**	0.015
					(2.39)	(0.84)
Alpha * Vol			-0.508***	0.012	-0.497***	0.003
			(-6.64)	(0.77)	(-6.31)	(0.20)
Alpha * Log(Age)			-0.554***	0.064**	-0.563***	0.054***
			(-5.90)	(2.52)	(-5.91)	(2.77)
Vol	0.046***	-0.038**	-0.074	-0.029	-0.081	-0.025
Log(Age)	-0.024**	-0.014	-0.401***	0.007	-0.403***	0.005
Log(TNA)	0.001	-0.002	-0.010	-0.010*	-0.012	-0.009*
Expense		-0.111***	-0.054	-0.097***	-0.067	-0.098***
Load		-0.000	0.026***	-0.001	0.028***	-0.001
Turnover	0.016***	-0.013	0.184***	-0.004	0.176***	-0.004
InstRatio	-0.014	-0.009	-0.072	-0.008	-0.072	-0.007
AlphaSQ			0.654***			
Constant	0.496***	0.180**	1.574***	0.114	1.433***	0.090
Observations	25941	107719	114037	107719	114037	107719
R^2	0.040	0.012	0.075	0.012	0.076	0.013

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.

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.

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

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.

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.

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

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

Fama, Eugene F., and Kenneth R. French, 1992, The Cross Section of Expected Stock Returns, *Journal of Finance* 47, 427–466.

Fama, Eugene F., and James D. MacBeth, 1973, Risk and Return: Some Empirical Tests, *Journal of Political Economy* 81, 607–636.

Gârleanu, Nicolae, and Lasse Heje Pedersen, 2016, Efficiently Inefficient Markets for Assets and Asset Management, UC Berkeley working paper.

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, 2015, 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.

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

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.

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).

Sirri, Erik R., and Peter Tufano, 1998, Costly Search and Mutual Fund Flows, *Journal of Finance* 53, 1589–1622.

Sokal, Robert R., and Charles D. Michener, 1958, A Statistical Method for Evaluating Systematic Relationships, *University of Kansas Science Bulletin* 38, 1409–1438.

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

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.