

Investor Attrition and Fund Flows in Mutual Funds

Christoffersen, Susan E. K.; Xu, Haoyu

Document Version

Accepted author manuscript

Published in:

Journal of Financial and Quantitative Analysis

DOI:

[10.1017/S0022109017000291](https://doi.org/10.1017/S0022109017000291)

Publication date:

2017

License

Unspecified

Citation for published version (APA):

Christoffersen, S. E. K., & Xu, H. (2017). Investor Attrition and Fund Flows in Mutual Funds. *Journal of Financial and Quantitative Analysis*, 52(3), 867-893. <https://doi.org/10.1017/S0022109017000291>

[Link to publication in CBS Research Portal](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

If you believe that this document breaches copyright please contact us (research.lib@cbs.dk) providing details, and we will remove access to the work immediately and investigate your claim.

Download date: 18. Jun. 2025



Investor Attrition and Fund Flows in Mutual Funds

Susan E. K. Christoffersen and Haoyu Xu

Journal article (Accepted manuscript*)

Please cite this article as:

Christoffersen, S. E. K., & Xu, H. (2017). Investor Attrition and Fund Flows in Mutual Funds. *Journal of Financial and Quantitative Analysis*, 52(3), 867-893. DOI: 10.1017/S0022109017000291

This article has been published in a revised form in *Journal of Financial and Quantitative Analysis*
<http://dx.doi.org/10.1017/S0022109017000291>.

This version is free to view and download for private research and study only. Not for re-distribution, re-sale or use in derivative works. © Cambridge University Press

* This version of the article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the publisher's final version AKA Version of Record.

Uploaded to [CBS Research Portal](#): January 2019

Investor Attrition and Fund Flows in Mutual Funds

Susan E. K. Christoffersen
University of Toronto and Copenhagen Business School

Haoyu Xu*
University of Toronto

This Draft: April 14, 2016

Forthcoming: *Journal of Financial and Quantitative Analysis*

*Christoffersen (corresponding author), susan.christoffersen@rotman.utoronto.ca, Xu, haoyu.xu10@rotman.utoronto.ca, Rotman School of Management, 105 St. George Street, Toronto, ON, M5S3E6, Canada. The paper has benefited immensely from comments from Stephen Brown (the editor) and an anonymous referee. We are also grateful for comments by seminar participants at the University of Toronto, Copenhagen Business School, 2014 Midwest Finance Meetings, 2014 Financial Management Association European Meetings, 2014 Financial Management Association Meetings, and the 2014 Northern Finance Association meetings and from comments from Che-Kuan Chen, Hannah Lea Huhn, Puneet Jaiprakash, and Laleh Samarbakhsh. Christoffersen gratefully acknowledges research support from the Social Sciences and Humanities Research Council and the Global Risk Institute. All errors are our own.

Investor Attrition and Fund Flows in Mutual Funds

Abstract:

We explore the properties of equity mutual funds that experience a loss of assets after poor performance. We document that both inflows and outflows are less sensitive to performance because performance-sensitive investors leave or decide not to invest after bad performance. Consistent with the idea that attrition measures the sorting of performance-sensitive investors, we find that attrition has less of an impact on the fund's flow-performance sensitivity for institutional funds where there is less dispersion in investor performance-sensitivity. Also attrition has no effect on the flow-performance sensitivity when attrition arises after good performance or investors invest for non-performance reasons.

JEL Classification: G11, G23

Keywords: Flow-performance relation, mutual funds, attrition, convexity, investor heterogeneity

I. Introduction

Understanding the flow-performance relation for mutual funds has been extensively studied in the literature (see Christoffersen, Musto, and Wermers (2014)). The seminal papers by Sirri and Tufano (1998) and Ippolito (1992) identify a convex relation which pose potential risk-taking incentives of mutual fund managers (Chevalier and Ellison (1997) and Brown, Harlow, and Starks (1996)) or even incentives to manipulate performance through fees or other tactics (Christoffersen (2001) and Carhart, Kaniel, Musto, Reed (2002)). Bad performance does not appear to be punished with investors leaving the fund.

Lynch and Musto (2003) develop a rational model to explain this convexity where investors recognize that managers may change their strategies after bad returns, so investors rationally stay in a fund knowing that the manager will adjust their strategy. Christoffersen and Musto (2002) and later work by Elton, Gruber, and Busse (2004) show that bad performance looking forward is all but guaranteed for a subset of money and index funds because of excessively high fees. Even though manager strategies change through time, there remains a predictable component of bad performance which does not seem to go away, and investors still seem to want to stay. Huang, Wei, and Yan (2007) argue that participation and search costs may play an additional role in explaining this inertia since high costs can dissuade investors from leaving to other competing funds. Rather than explicit costs to investing, Christoffersen and Musto (2002) and Berk and Tonks (2007) argue that investor heterogeneity might explain the lack of sensitivity with some investors facing either behavioral or other constraints which prevent them from responding to poor performance.

Even with heterogeneity across investors, cross-sectional differences in fund performance sensitivities and flow convexity would only arise if some individuals were non-randomly assigned to different funds. To this end, Christoffersen and Musto (2002) draw from the

mortgage literature and first introduce a retention measure for money market mutual funds, which measures the current size of the fund relative to its maximum historical size. Funds, which have experienced a large outflow of investors after poor performance, are presumably left with a non-random sample of performance-insensitive investors since by leaving or staying, investors reveal their preference (or lack of) for high past performance (Samuelson (1938), (1948), and Varian (2006)).

Despite its intuitive appeal, little is known about the attrition measure and to what extent it can help identify funds with different performance sensitivities. The measure so far has been applied directly to money funds and used to explain changes in money fund pricing around mergers and the wide dispersion in money fund fees. Related work by Cooper, Halling, and Lemmon (2013) investigate how flow-performance sensitivities affect pricing in equity markets but do not directly look at attrition measures in equity markets. Gil-Bazo and Ruiz-Verdu (2008) investigate fee changes in equity funds and argue that increases in fees after attrition arise both from passive incentives to extract rents from inert investors as well as active incentives to target performance-insensitive investors. In their appendix, Gil-Bazo and Ruiz-Verdu (2008) show that the flow-performance relation of equity funds varies with a fund's attrition. We build on these results by controlling separately for attrition arising from poor returns and by exploring the effect of attrition on convexity, inflows, outflows, or investigate factors that may make attrition a better or worse predictor of the flow-performance sensitivity.

The purpose of this article is therefore to explore properties of attrition and under what conditions it serves as a good proxy for the revealed preferences of the investors in a fund. The specific questions for this study consider how attrition influences flow-performance sensitivities: (1) Is attrition a reliable measure to capture cross-sectional differences in performance sensitivities across funds? (2) Does attrition affect the price sensitivities of inflows as well as

outflows? (3) Does the measure of attrition capture the revealed preferences of investors and does this influence the effect of attrition on flows in the cross-section?

In short, the answer to all three questions is yes. To address all these questions, we first adopt a measure of investor attrition that we can easily apply to each fund at each point in time. As in Christoffersen and Musto (2002), attrition is defined as 1 minus the ratio of current asset size to the historical maximum size attained before the end of year t . One important contribution of this paper is to properly control for performance since poor performance will result in both a reduction in fund size as well as cause performance-sensitive investors to leave. It is the latter consequence that is of greater interest, as it is the self-selection of investors which helps to identify a fund with less performance sensitivity. In all our analysis, we therefore control explicitly for accumulated returns measured from the time when the fund reaches its historical maximum size to time t .

We are able to significantly predict the level of attrition in a fund (with R^2 values ranging around 30%) based on several determinants that align with our interpretation of investor seasoning. Lower accumulated returns, older funds, and smaller funds have higher attrition measures as one might expect. Load funds associate with low attrition, presumably because the broker intermediary is good at keeping investors in the fund. As found in Christoffersen and Musto (2002), higher expenses are charged to investors in high attrition funds which one might expect if a high attrition fund extracts rents from a pool of performance-insensitive investors. Interestingly, funds in large families typically have higher attrition levels than those in smaller families which could reflect that larger fund families are older with higher family levels of attrition over time.

In answering the first research question, we estimate flow-performance relations, but unlike prior estimations, we include attrition and an interaction term between attrition and performance

as explanatory variables. Consistent with our sorting interpretation, those funds, which have experienced a high amount of attrition after poor performance, exhibit a significantly flatter flow-performance relation. In contrast, investors leaving a fund after a period of good performance are not being sorted by their performance sensitivity. We hypothesize that the conditions which attract an investor to a fund are important when interpreting attrition. If performance is unimportant to investors when they enter and exit a fund, then attrition will not capture the selection effect of performance-sensitive investors leaving a fund. Therefore, while high attrition funds experience lower flow-performance sensitivities after a period of poor performance, attrition has no effect on the performance-sensitivity of the fund if returns have recently been good.

The relation between attrition and flows is separate from the well-known finding of Chevalier and Ellison (1997) that older funds exhibit much less sensitive flow-performance relations. Although old funds have more likely undergone greater amounts of attrition and are left with a core group of stable investors, both age and attrition independently dampen the flow-performance relation when both are included as independent variables. In fact, the importance of age in predicting flows diminishes once attrition is included as a dependent variable so the difference in flow sensitivities between old and young funds appears to reflect differences in attrition. Overall, the flow-performance relation is flatter in high attrition funds but only after a period of underperformance.

Our second research question considers whether attrition influences flow convexity and related to this, if attrition affects inflows as well as outflows. The idea that a fund with high attrition loses many performance-sensitive investors tells us something about the remaining investors in the fund and the predicted insensitivity of outflows to performance. Empirical estimates isolating outflows confirm that they have a very weak response to poor performance if

fund attrition is high. What about inflows? Attrition could equally have an effect on how investors enter the fund. Drawing from the vast literature on revealed preferences in economics, we know that a fund which has experienced a period of weak performance and significant attrition will not attract performance-sensitive investors who place importance on performance as an investment criterion. Instead, these funds will market to investors based on other fund characteristics aside from performance (i.e. services). In line with this, we find that high attrition funds attract inflows which are less sensitive to performance than their low attrition counterparts. Higher attrition therefore dampens both inflow and outflow performance sensitivities. As a result, a combined estimation of flow convexity for all attrition types may be more exaggerated if high attrition funds (with low performance sensitivity) more often associate with poor performance rankings and low attrition funds (with high performance sensitivity) more often are observed with good performance rankings.

Our final estimations hypothesize and test two conditions when attrition will have a more or less pronounced effect on flow-performance sensitivity. First, in a situation where investors are homogeneous and have exactly the same performance sensitivity, then attrition should have little effect on the flow-performance relation of the fund since there is no separation of investor types that occurs with attrition. In contrast, large investor heterogeneity implies an enhanced effect of attrition. To test this, we use institutional investment in the fund as a proxy for investor homogeneity with the assumption that institutional investors are more sophisticated and similar in their flow-performance sensitivity than retail investors. Consistent with this, we find attrition has little effect on the flow-performance relation in institutional share classes but it has a significant dampening effect on flow sensitivities amongst retail investors.

As mentioned above, attrition is only an effective proxy for lower performance-sensitivity when we know that investors are leaving for performance reasons. If attrition is arising for non-

performance reasons (i.e. poor service or fund family scandal), then it won't necessarily reveal the performance-sensitivity of the investor base. So what other factors might drive an investor's decision to invest aside from performance? We use two proxies to separate funds as ones where investors rely on non-performance versus performance rankings in making their investment decision. These proxies are: (1) index vs. active funds; and (2) size of the fund category.

Our hypothesis is that relative performance is a more important investment criterion in active funds and in categories with many competing funds. First, in comparing active and index funds, we argue that relative performance is critically important to investors when investing in an active fund, as opposed to an index fund, since investors are trying to determine the skill of the manager through observed performance. In a category with many competing funds, relative performance is an important determinant of flows because there are so many alternative options of where to invest: performance becomes the one dimension on which the fund can distinguish itself and compete. For both these proxies (active funds and large categories), attrition has a significantly greater effect on dampening the flow-performance relation consistent with the theory that attrition will only have an impact on the flow-performance sensitivity when performance is an important factor for investment.

Given the close association between attrition and flow-performance sensitivity, an attrition measure will be particularly helpful to researchers and practitioners needing to distinguish funds with different performance sensitivities and convexities. The only other way to estimate fund-level performance sensitivities is through time-series estimation of a fund's flow-performance relation, but this method is fraught with two main problems. First, it assumes that performance-sensitivity is constant through time for the mutual fund, which may not be the case. Second, it relies on a time series of data so eliminates young funds. Instead, the attrition measure is easy to measure fund-by-fund, changes over time, and does not rely on a large time-series of data. With

an attrition measure, one can easily include it as an additional fund characteristic to measure demand characteristics of the investor base.

The remainder of the paper is divided as follows. Section II develops the main hypotheses and Section III outlines the data and discusses the attrition measure and factors important to it. Section IV presents the main results relating attrition to flow performance. Finally, Section V concludes.

II. Hypothesis Development

We can learn a lot about the investors currently in a fund by conditioning on a fund's history. A fund which has lost investors reveals something about the investors who remain. This logic follows from the early work of Samuelson (1938), (1948) on revealed preferences where observing a consumer's purchasing decision and choice tells us something about their preferences. Take a simple example of an individual who has to choose where to shop. Two stores are the same on all dimensions except price and service where one store is a discount store with no service and the other a higher-priced store with better service. The individual selecting the discount store is revealing that price is more important than service and the opposite for someone going to the service-oriented store. If prices start to increase at the discount store, one might expect many of the shoppers to leave since this is an important quality of shopping for the discount shoppers. On the other hand, shoppers at the service-oriented store would be much less sensitive to increases in the price of goods because they have already revealed that service is a more important quality they seek when shopping.

In the literature on pricing mortgage-backed securities, Schwarz and Torous (1989) posit there is heterogeneity of investor sensitivity to interest-rate movements and that a period of low-interest rates will separate these two investor types as the interest-sensitive investors refinance their mortgage leaving a pool of interest-insensitive individuals. Christoffersen and Musto

(2002) carry this logic to money market mutual funds which have experienced a period of poor performance and then apply this to pricing strategies of a money fund.

This paper builds on the insights of Christoffersen and Musto (2002) but considers the effects of attrition in equity funds and then tests directly whether flow-sensitivity differs after periods of attrition and under what conditions attrition will affect flow-sensitivity. The premise of all tests is that there are two types of investors: those who respond vigorously to relative performance (performance-sensitive) and others who do not (performance-insensitive). We will also assume that fund investors reveal their type with bad performance realizations. Along all other dimensions, investors are assumed to be the same.

Suppose that the two investor types are represented in a fund and that there are also both types of investors who have not yet decided to invest in the fund. If the performance of the fund improves, one expects both investor types to stay and enter the fund. The performance-insensitive investors are indifferent to improvement in performance and the performance-sensitive investors are made better off so there is no separation in investor types when performance is good because everyone is content to stay. On the other hand, if performance worsens, we now observe two effects on the fund: performance-sensitive investors will leave the fund *and* will not enter. The overall effect is that the only investors left or the few entering the fund are the performance-insensitive investors who are indifferent about the changes in the relative performance of the fund. From this, our first hypothesis is:

Hypothesis 1. After a period of bad performance, performance-sensitive investors will leave as well as avoid investing the fund. Fund attrition therefore associates with a lower flow-performance sensitivity of outflows, inflows, and net flows.

So when will attrition not have an effect on the performance-sensitivity of outflows? This depends on two main criteria. First, there has to be heterogeneity in investor types otherwise there is no sorting that happens through attrition. Second, attrition will only affect the flow-performance sensitivity after poor performance. If attrition is caused by a non-performance characteristic or some other exogenous shock, then investors are leaving for reasons that are unrelated to performance so attrition is not sorting investors by performance-sensitivity. This insight provides two follow-up hypotheses:

Hypothesis 2. Attrition will have no effect on the flow-performance sensitivity of outflows, inflows, or net flows if all investors have homogenous performance sensitivities.

Hypothesis 3. Attrition will have no effect on the flow-performance sensitivity of outflows, inflows, or net flows if investors are leaving or avoiding a fund for non-performance reasons.

To illustrate the third hypothesis, consider the case where a fund has altered the services provided to investors but performance remains the same. In this case, investors are leaving the fund for reasons which are non-performance related. The sorting of investors in this case reveals something about how important service is to the investor clientele but it tells us nothing about the performance-sensitivity of who is leaving or avoiding the fund. Attrition will therefore not provide any information about the fund's performance sensitivity if performance is not the key criterion used by investors when deciding to invest or divest. In our empirical tests, we consider two cases where relative performance may be a less important criterion for investment and test

whether these coincide with cases when attrition has a minimal effect on flow-performance sensitivities.

Overall, we conclude that higher attrition should imply an overall weaker flow-performance relation than cases with low attrition. This relation is more apparent when investor heterogeneity is more disperse and when the investors choosing to leave and enter a fund base their decision on relative performance measures.

III. Data

The data we use in this study are from Center for Research in Security Prices (CRSP) Survivor-Bias-Free Mutual Fund Database. This data set provides daily returns of mutual funds, total assets under management, fee structure and other fund characteristics. Our sample is from 1998 to 2012 and is restricted to actively-managed U.S. domestic equity funds. To minimize potential incubation bias (Evans (2010)) and to isolate equity funds, we exclude funds which are younger than 3 years old, and any funds which are 3 years old with less than 3 million dollars under management. We also eliminate those funds with less than 80% average equity holding to focus on equity funds.

Many funds have multiple share classes where the same fund is offered with various fee structures. We aggregate share classes so each observation is at the fund-year level rather than by share class (see Wermers (2000)). FUND_SIZE is simply the summation of the total net assets of each share class and is expressed in millions of dollars while $\ln(\text{FUND_SIZE})$ is its natural logarithm. EXPENSES are calculated as the size-weighted average from each share class and are expressed in decimal form where 1.5% is represented as 0.015. RETURN is also the size-weighted average of net annual returns for each share class and is reported in decimals. The variable $\ln(\text{ACCUMULATED_RETURN})$ is the natural logarithm of one plus the compounded monthly net returns calculated from the end of the month when the fund reached

its historical maximum size to the end of year t . It is reported in decimals so a fund which is currently at its historical maximum size implies that the $\ln(\text{ACCUMULATED_RETURN})$ is equal to 0. In the flow regressions, lagged annual net returns are ranked between 0 and 1 and these ranked returns are denoted as RANKED_RETURN . In the flow regressions of convexity, variables are included to identify funds with low, medium, and high performance. As in Sirri and Tufano (1998), the piecewise variables, LOW , MED , and HIGH , are defined as $\min(\text{RANKED_RETURN}, 0.2)$, $\min(0.6, \text{RANKED_RETURN} - \text{LOW})$, and $\text{RANKED_RETURN} - \text{MED} - \text{LOW}$. We also use a simpler piecewise estimation of flows with only one kink at the median ranked return level so in this case LOW50 and HIGH50 are defined as $\min(\text{RANKED_RETURN}, 0.5)$ and $\text{RANKED_RETURN} - \text{LOW50}$.

There are situations where a new share class is introduced many years after the fund portfolio is created. For such cases, we use the birth date of the oldest share class as the starting date to determine the age of the fund and AGE is expressed in years. A fund is considered dead, if all of its share classes are liquidated or merged by other funds. Sometimes, a fund merges its own share classes into one, but we do not consider this a merger and continue to track the fund going forward.

FLOW is calculated as the change in total net assets between the last month in year $t+1$ to the last month in year t while removing the effects of the annual return on assets, as assumed in Sirri and Tufano (1998)

$$(1) \quad \text{FLOW}_{i,t+1} = \frac{\text{FUND_SIZE}_{i,t+1} - \text{FUND_SIZE}_{i,t} \times (1 + \text{RETURN}_{i,t+1})}{\text{FUND_SIZE}_{i,t}},$$

where $\text{RETURN}_{i,t+1}$ is the annual net return of the portfolio in year $t+1$. Flows are expressed in decimals as with expense ratios and returns.

To estimate attrition, we include AGE , FLOW , $\ln(\text{ACCUMULATED_RETURN})$, EXPENSES , and $\ln(\text{FUND_SIZE})$. In addition, we use four other control variables to capture

cross-sectional differences in attrition levels across funds. INSTITUTIONAL_SHARE is the proportion of institutional investors in a mutual fund. It is reported as a decimal where 5% is reported as 0.05. LOAD_DUMMY is a dummy variable taking the value 1 if a fund charges either a front- or back-load, and 0 otherwise. Both $\ln(\text{FAMILY_SHARECLASS})$ and $\ln(\text{FUNDS_IN_CATEGORY})$ are the natural logarithms of the number of share classes offered in a family and the number of funds in a category.

For all the estimations of future flows, we control for category flows, CATEGORY_FLOW, where a category average flow is calculated as an asset-weighted average of fund flows in the same category.¹ Also, lagged flows are included as a control for potential serial correlation in annual flows.

A. Measuring Attrition

One of the contributions of this paper is to analyze more closely the cross-sectional relevance of differences in attrition rates across funds. The concept of attrition comes from the early literature on revealed preferences (Samuelson (1948)) and its many applications (see references in Varian (2006)). In the finance literature the concept was introduced in the mortgage literature (i.e. Schwartz and Torous (1989)) where a recent decline in interest rates will cause some to refinance leaving a pool of mortgages that is less susceptible to refinancing risk going forward. Following the methodology of Christoffersen and Musto (2002), we apply this concept to mutual funds by measuring the current size of a fund to its maximum historical size. ATTRITION is therefore calculated as

$$(2) \quad \text{ATTRITION}_{i,t} = 1 - \frac{\text{FUND_SIZE}_{i,t}}{\text{HISTORICAL_MAX_FUND_SIZE}_{i,t}},$$

¹ Categories are first assigned based on classifications by Lipper. When a Lipper classification is not available, then category codes are used from the Strategic Insight and Weisenberg to fill in the missing category information.

where $FUND_SIZE_{i,t}$ is the total net assets (in millions) for each fund i in the end month of each year t , and $HISTORICAL_MAX_FUND_SIZE_{i,t}$ is the historical maximum fund size over all months prior to the end of year t . We calculate the attrition rate of each fund at the end of each year. Although the analysis focuses on data from 1998 to 2012, all the attrition and accumulated return measures are determined over the full history of the fund so the historical maximum size may have been realized prior to 1998 although the attrition measure itself is only calculated between 1998 and 2012. For much of our analysis relating flows to attrition, we use a dummy variable for attrition to indicate high and low levels of attrition. $ATTRITION_DUMMY$ takes the value 1 if $ATTRITION$ is greater than the median level across all funds and years, and 0 otherwise.

Panel A of Table 1 presents the average, median, minimum, maximum, and standard deviation for the key variables in the analysis. We observe attrition ranges from 0 to 0.99 across the years. Average attrition is 0.33 or 33% of the historical maximum size. Attrition increases in years where there are much higher levels of negative returns such as in 2008. This highlights the importance of controlling for the accumulated returns since attrition levels are jointly affected by returns of the portfolio as well as individual investor decisions to enter and leave the fund. Average expenses and net annual flows of 1.12% and 8.4% match findings in other papers. Fund size is quite skewed since the average fund size is around \$962 million but can be as large as \$167 billion for the largest fund. To minimize the effect of skewness in the estimation, the natural logarithm of fund size is used throughout.

Panel B of Table 1 presents the correlations of raw attrition with control variables used in the analysis. As we expect, attrition is negatively correlated with size, the natural logarithm of accumulated return, and flow, and positively correlated with age and expenses. The highest level of correlation between attrition and any of the control variables is with fund age, 0.30,

since older funds are more likely to have experienced more attrition. Because the prior literature has found age to be an important determinant of fund flows, our analysis is careful to separate the effects of attrition from age. Even with a correlation of 0.30, it is clear that attrition is also capturing a different characteristic of the fund, separate from age.

Table 2 groups the data into five attrition quintiles and presents average levels of fund size, accumulated returns, expenses, age, and flows for each of the quintile groups to gain perspective on how the variables change economically. Older and smaller funds tend to be ones that have experienced more attrition and funds with more attrition have lower future expected inflows. We observe that high levels of attrition correlate strongly with funds that have experienced poor performance and higher fees. Both are consistent with the notion that attrition measures a separation of performance-sensitive investors. As pointed out in Christoffersen and Musto (2002), higher fees associate with higher attrition levels as performance-sensitive investors leave or avoid funds with high fees. In equilibrium, higher fees are therefore charged to a group of investors who are less responsive to the costs and fees they pay. To get an idea of the magnitude of these relations, Table 2 shows that fees in the lowest attrition quintile average 1.02% while fees for a fund in the highest attrition quintile are significantly higher at 1.33%. On average, accumulated returns are increasingly worse when there are higher levels of attrition even though from Table 1 we observe that accumulated returns can be both positive and negative for funds experiencing attrition. The next section identifies those variables which are important contributors to attrition and later what effects attrition has on flows.

B. Explaining Attrition

Our measure of raw attrition is influenced by both the performance of the fund and the decision of performance-sensitive investors to leave the fund. If we want to isolate the second effect of investors leaving the fund, we need to control for historical performance. Table 3

estimates those factors influencing attrition (estimated at the end of the year) where standard errors are clustered by fund and year fixed effects are included for two of the specifications following Petersen (2009). The first two columns of Table 3 are ordinary least squares regressions. Columns 3 and 4 are probit models of the ATTRITION_DUMMY which takes the value 1 if attrition is above its median level and 0 if below. Column 5 reports the marginal effects of the probit estimation in column 4.

In all estimations, $\ln(\text{ACCUMULATED_RETURN})$ enters significantly negative so worsening levels of accumulated returns increase the level of attrition in a fund. Economically, a 1 standard deviation decrease in $\ln(\text{ACCUMULATED_RETURN})$ (0.39) in column 1 will increase attrition by 4.96% so poor performance results in people leaving and not fully being replaced. Older and smaller funds correspond with higher attrition levels. Funds with loads have lower attrition levels, presumably because the brokers are good at keeping investors in a fund. As tested and posited in Christoffersen and Musto (2002), we find that expenses are higher in high attrition funds which is consistent with funds charging higher expenses to a subset of more inert investors. Overall, the baseline model of attrition in Table 3 provides estimates that are consistent with the notion of investor types being separated based on the historical realization of returns in the fund.

The high R^2 values in the baseline estimation demonstrate that we are capturing and explaining a large component of attrition. For our analysis estimating future flows, we include all the baseline controls for attrition in order to remove any confounding effects that these variables may have in explaining future flows. The inclusion of these controls isolate how residual and excess attrition affect future flows and flow-performance sensitivities and separates attrition from other potential effects on flows such as prolonged periods of poor performance. In

the next section, we explore whether funds experiencing a lot of unexplained attrition, exhibit weak flow-performance sensitivity and test the hypotheses outlined in Section II.

IV. Attrition and Flow-Performance Sensitivity

Having identified fund specific factors influencing attrition rates, we turn our attention to estimating different effects of attrition on fund flows by first considering its influence on the flow-performance relationship. For all the estimations, we report estimates of large panel regressions with robust standard-errors that cluster by fund and include year fixed effects following Petersen (2009) and Gormley and Matsa (2014). In unreported results, we also include fixed category effects in the regressions rather than including category averages and find similar results.

A. Flow-Performance Relation

The first part of our analysis determines if attrition provides a proxy for the performance-sensitivity of investors in a specific fund. Numerous studies have estimated the flow-performance relation of a large cross-section of funds, and in some cases, the funds are grouped into different subsamples, such as old and young funds (Chevalier and Ellison (1997)) to identify differences in sensitivities across funds. However, measures to identify performance-sensitivity of a specific fund is limited to time-series estimation which assumes that performance-sensitivity of a fund remains the same through time and is constrained by the time series data available for the fund. A fund-specific measure of flow-performance sensitivity could be immensely helpful to researchers and practitioners wanting an easy way to assess the stability of investors in a fund and the likeliness these investors will respond to future downturns in the fund's performance.

If high attrition identifies remaining individuals in the fund who are likely to stay in the future, then we expect higher attrition rates to correspond with lower flow-performance

relations. The regression setting follows Sirri and Tufano (1998) where we predict future fund flows with its lagged ranked return relative to other funds with the same investment objective.

Our baseline regression explaining future flows is

$$\begin{aligned}
 (3) \quad \text{FLOW}_{i,t+1} = & \alpha + \beta_1 \text{RANKED_RETURN}_{i,t} + \beta_2 \text{ATTRITION_DUMMY}_{i,t} + \\
 & \beta_3 (\text{ATTRITION_DUMMY}_{i,t} \times \text{RANKED_RETURN}_{i,t}) + \beta_4 \text{AGE}_{i,t} + \beta_5 (\text{AGE}_{i,t} \times \\
 & \text{RANKED_RETURN}_{i,t}) + \beta_6 \text{FLOW}_{i,t} + \beta_7 (\text{FLOW}_{i,t} \times \text{RANKED_RETURN}_{i,t}) + \\
 & \beta_8 \text{CATEGORY_FLOW}_{i,t+1} + \beta_9 \ln(\text{ACCUMULATED_RETURN}_{i,t}) + \\
 & \beta_{10} \ln(\text{FUND_SIZE}_{i,t}) + \beta_{11} \text{EXPENSES}_{i,t} + \beta_{12} \text{LOAD_DUMMY}_{i,t} + \\
 & \beta_{13} \text{INSTITUTIONAL_SHARE}_{i,t} + \beta_{14} \ln(\text{FAMILY_SHARECLASS}_{i,t}) + \\
 & \beta_{15} \ln(\text{FUNDS_IN_CATEGORY}_{i,t}) + e_{i,t}.
 \end{aligned}$$

Table 4 estimates this basic flow regression and interacts ATTRITION_DUMMY, AGE, and FLOW with RANKED_RETURN. Overall, the regressions in Table 4 provide strong evidence supporting H1: higher attrition associates with a weaker fund flow sensitivity. We observe this for net flows as well as for outflows and inflows. In column 1, we provide a baseline regression without any inclusion of attrition. When we add attrition to the baseline regression, we observe that on its own attrition enters significantly negative so higher attrition rates correspond with lower future flows to the fund. In addition, the coefficient on AGE drops from a significant value of -.0038 to an insignificant value of -.0009. Therefore, the documented effect of age on flows reflects the fact that in prior estimations, age has been proxying for the sensitivity of the investor base and the omitted effect of attrition.

To more precisely estimate whether higher attrition associates with lower performance sensitivity, the regression in column 3 includes an interaction term, ATTRITION_DUMMY \times RANKED_RETURN. The interaction term enters significantly negative and the dampening

effect of attrition on flows is very pronounced as a high attrition fund has a flow-performance sensitivity which is about 70% the sensitivity of a fund with low attrition.²

Lagged flows are included as an explanatory variable of next year's flow levels to control for potential serial correlation in flows that may influence both attrition and flows. In all specifications, $FLOW_t$ enters significantly positively with coefficients ranging from 0.15 to 0.18. While serial correlation is present in flows, the size of the coefficient suggests there is also a significant amount of surprise and variation in flows from one year to the next that cannot simply be predicted from past flows.

Because age has been used in prior literature to capture differences in flow-performance, we also interact $AGE \times RANKED_RETURN$ and include this variable in column 3. The interaction term between age and performance enters negatively and significantly but does not diminish the significance of the interaction term between attrition and performance. $ATTRITION_DUMMY \times RANKED_RETURN$ therefore captures an effect on the flow-performance sensitivity which is distinct from that of fund age.

Regression 3 also includes an insignificant interaction between $FLOW$ and $RANKED_RETURN$. The lack of significance of in this interaction term is important because even though attrition is determined to some extent by past flows, sorting on past flows is not sufficient to separate funds by flow-performance sensitivities: Attrition is a better measure of capturing differences in performance sensitivities.

Rather than using net flows, the last two columns of Table 4 use tobit estimations to analyze performance sensitivities separately for positive and negative flows. The left-hand side variable

² From Table 4, $0.494 - 0.1438 = 0.3502$ is the performance sensitivity of high attrition funds and 0.3502 divided by 0.494 is 70.9%. High attrition funds have 70.9% the sensitivity to performance as low attrition funds.

in column 4 is $\text{FLOWPOS}_{i,t+1}$ which is defined as $\max(\text{FLOW}_{i,t+1}, 0)$. In the last column, the dependent variable, $\text{FLOWNEG}_{i,t+1}$, is defined as $\min(\text{FLOW}_{i,t+1}, 0)$. Both columns (4) and (5) are tobit estimations with clustered standard errors by fund and yearly fixed effects.

These two tobit estimations reveal interesting differences between inflows and outflows. First in comparing the coefficients on RANKED_RETURN in columns 4 and 5, it is clear that inflows are more than four times more sensitive to past performance than are outflows. This highlights an inertia of individuals once they enter a fund since they are much less sensitive to performance when they leave. There is also more serial correlation in inflows than with outflows so more inflows coming in one year will predict more money arriving the following. This could be because of persistent advertising or even broker sales that attract money to the fund year to year. General category flows have a much stronger effect on inflows than on outflows, so it seems inflows into funds reflect general market trends into categories while outflows are less influenced by aggregate market movements.

To test H1, we interact $\text{ATTRITION_DUMMY} \times \text{RANKED_RETURN}$ to see if attrition affects the flow-performance sensitivity of inflows or outflows or both. The interaction term enters significantly and negatively for both inflows and outflows. Therefore, consistent with H1, attrition corresponds with a separation of investor types both in the people who leave a fund and the people who are attracted to it. In both inflows and outflows, the performance sensitivity is reduced by about a third when comparing a high attrition fund with a low attrition fund. It therefore seems the performance-insensitive investors remain and are the only types who are attracted to a fund after a long period of attrition.

B. Attrition and Performance

In Table 5, we test H1 more precisely by separating funds into those with positive or negative $\ln(\text{ACCUMULATED_RETURN})$. The basic intuition follows from the discussion of

the hypotheses. It is only possible to separate investors with different performance-sensitivities when performance is poor. With good performance, everyone is interested in the fund so no separation occurs, but with bad performance those individuals concerned with performance will either leave or not enter. This intuition is evident when comparing the first two columns in Table 5. Attrition after a period of good performance has no effect on the flow-performance relation, presumably because individuals sensitive to performance are not leaving and continue to enter the fund. However, attrition after a period of bad performance has a significant dampening effect on the flow-performance relation and cuts the sensitivity in half.

In columns 3 to 6, we analyze how sensitivities differ for inflows and outflows. Inflows arising after good performance are 1.7 times³ more sensitive to performance than inflows coming in after poor performance, consistent with the idea that good performance attracts both performance sensitive and insensitive investors while bad performance does not. Also, attrition arising after good performance has no effect on either inflows or outflows but has a significant negative impact on the flow-performance sensitivity after a period of bad performance since the interaction term between ranked returns and attrition is only significantly negative after a period of negative accumulated returns. These results provide very convincing evidence that attrition is capturing the effect of separating investors by their performance-sensitivities and support the notion that the effect of attrition is more pronounced after a period of poor performance as highlighted in H1.

Note that the interaction terms $\text{RANKED_RETURN} \times \text{AGE}$ and $\text{RANKED_RETURN} \times \text{FLOW}$ are controlled for in all the estimations in Table 5 and either don't enter significantly or if they do, the results are not picking up the same relations as captured by attrition. For instance,

³ This compares the relative sensitivity between high and low accumulated return funds i.e. $1.3374/0.78$.

older funds tend to have weaker flow performance relations but only after good performance not bad performance and there is no consistent pattern with inflows and outflows.

C. Flow Convexity

Since attrition has such a significant impact on the flow performance relation, our next step is to see how attrition affects fund flow convexity. From the results in Table 4, attrition has a separate effect on inflows and outflows with a larger effect on the performance-sensitivity of inflows than outflows.

To test the effect on convexity more concretely, Table 6 estimates the flow convexity of funds with high and low attrition defined by whether the attrition dummy equals 1 or 0. The baseline model used to estimate flows is the same as used in Table 4 except `RANKED_RETURN` is transformed into piecewise variables to allow for different slopes over different ranges of lagged ranked returns. We estimate two different piecewise regressions to capture the convexity of flows. In the first two columns, `LOW50` and `HIGH50` allow the flow-performance relation to differ at the median rank return. Following the approach of Sirri and Tufano (1998), the estimation in the last two columns allows for kinks at the bottom and top quintiles of the return distribution over the performance regions `LOW`, `MED`, and `HIGH`.

As with other studies, we observe a convex flow-performance relationship in that the slope coefficient for high performance is significantly larger than for low performance for both high and low attrition funds. The new insight is that the combined estimation of fund flow convexity for all attrition types may in part be enhanced if high attrition funds (with low-performance sensitivity) associate more closely with poor performance and low attrition funds (with heightened performance sensitivity) associate more strongly with good performance. In columns 1 and 2, we observe that for below-median performers, low attrition funds seem to have a more responsive investor base with a significantly estimated flow-performance

coefficient of 0.3694 but amongst high attrition funds this coefficient is insignificant and economically half the size, 0.171. A similar pattern is observed amongst high performers, with the low attrition funds having a much more responsive investor base than high attrition funds. When we compare the slopes on LOW, MED, and HIGH, the high attrition funds have a more muted flow-performance relation at all levels of performance compared to the low attrition funds.

We are also careful to include AGE and $\text{AGE} \times \text{RANKED_RETURN}$ in the regressions. Research by Chevalier and Ellison (1997) documents that older funds have less convex flow-performance relations, but this is not driving our results. Attrition is capturing something different from simply the past historical returns of the fund and the age of the fund.

In support of our findings in Table 6, we also graphically depict the non-linear flow-performance relation in Figures 1 and 2. These figures shows the semi-parametric estimation of the flow-performance relation for high and low attrition funds using the kernel estimation methodology of Chevalier and Ellison (1997) to estimate the following semi-parametric regression:

$$\begin{aligned}
 (4) \quad \text{FLOW}_{i,t+1} = & f(\text{RANKED_RETURN}_{i,t}) + \alpha_1 \text{ATTRITION}_{i,t} + \\
 & \alpha_2 \ln(\text{ACCUMULATED_RETURN}_{i,t}) + \alpha_3 \text{FLOW}_{i,t} + \alpha_4 \text{AGE}_{i,t} + \\
 & \alpha_5 \ln(\text{FUND_SIZE}_{i,t}) + \alpha_6 \text{EXPENSES}_{i,t} + \alpha_7 \text{CATEGORY_FLOW}_{i,t+1} + \\
 & \alpha_8 \text{LOAD_DUMMY}_{i,t} + \alpha_9 \text{INSTITUTIONAL_SHARE}_{i,t} + \\
 & \alpha_{10} \ln(\text{FAMILY_SHARECLASS}_{i,t}) + \alpha_{11} \ln(\text{FUNDS_IN_CATEGORY}_{i,t}) + \epsilon_{i,t} .
 \end{aligned}$$

Equation (4) controls for the key explanatory variables explaining flows and also allows for a non-linear relation between RANKED_RETURN in year t and FLOW in year $t+1$. We adopt the two-stage method of Robinson (1988) to attain unbiased estimates of the control variable coefficients, α_1 to α_{11} , and then remove these effects from flows when estimating the semi-

parametric relation between ranked returns and flows that are plotted in Figures 1 and 2 for different attrition levels.

Figure 1 plots the resulting non-linear function relating flows and ranked returns for attrition levels which are above (high attrition) or below the median (low attrition). Figure 2 plots a similar flow function for attrition levels in the high- and low-attrition terciles. Consistent with our Table 6 estimates, the semi-parametric kernel estimations show that overall slope declines in the high attrition funds with a flattening in the flow-performance relation for both low and high performers.

D. Investor Heterogeneity and Attrition

From the hypothesis section, we claim that attrition proxies for the separation of investor performance-sensitivity types. There are two conditions that are necessary for this to be true: (1) investors are heterogeneous; and (2) investors have to be selecting funds based on relative performance and not some other criteria. To test the first condition, we separate funds into the institutional and retail share classes.

Unlike the earlier tables, Table 7 separates institutional and retail share classes. Flows, attrition, age, accumulated returns, fund size, expenses and the load dummy are created following the same methodologies as outlined in the data section except they are defined using institutional or retail share classes separately for each fund. In other words, asset-weighted institutional (retail) share classes are used to create these variables in the institutional (retail) regression. In the case where a fund only consists of either institutional or retail share classes, then the observation for the fund would be the same as in the prior estimations and the observation would show up as part of either the retail or institutional flow regressions. In the case that a fund has both retail and institutional share classes, then the fund variable would have both a retail and an institutional observation that would enter separately in the institutional and

retail flow regressions. Because of the change in variable construction all variable coefficients, including those of the control variables from Table 3, are shown. The portion of institutional share classes in a fund is not included as a control since retail and institutional share classes are estimated separately.

Table 7 therefore separates flows by whether the money is institutional or not. The underlying assumption is that institutional investors are more homogeneous and sophisticated in their performance-sensitivities than retail investors. As a result, an institutional share class that has experienced a lot of attrition will be left with a similar group of investors so the performance sensitivity of the remaining pool of investors should not differ. From H2, we therefore expect that our attrition proxy should have less of an impact on the flow-performance relation when the investor base is more homogenous, or has more institutional investors.

Consistent with this logic, Table 7 shows that attrition has no effect on the flow-performance relation if the investor base is institutional. While both institutional and retail funds have a significantly positive relation between lagged ranked returns and future flows, the coefficient on $\text{ATTRITION_DUMMY} \times \text{RANKED_RETURN}$ is insignificant for funds with institutional investors and is significantly negative for funds with retail investors. A similar pattern is observed when we separately estimate the flow-performance relation for inflows vs. outflows. We therefore conclude that attrition proxies for a separation of investor types amongst retail investors, but not for institutional investors which are more homogenous. The finding also supports the conclusions in Christoffersen and Musto (2002) showing retail funds increase fees after attrition to reflect the more inert investor base, while there are no similar fee changes observed for institutional funds.

The other notable aspect of the institutional and retail flow regressions is that institutional flows seem less affected by performance chasing than retail flows. We observe a much stronger

relation between past performance and flows when estimating net flows and inflows of retail rather than institutional funds. This is consistent with more sophisticated investors relying on other factors aside from past performance to determine their investment decisions.

E. Other Criteria for Investing

Investors choose funds for many reasons which are unrelated to performance. If attrition is an effective proxy for the separation of investor-sensitivity types, then we only expect attrition to affect investor-sensitivity when performance is an important criterion for investment. As in the example given in the hypothesis section, if investors are choosing funds because of some other investment factor (such as services) then this criterion will not separate investors by their performance sensitivity. In Table 5, we already observe that attrition only significantly affects the flow-performance relation when performance is poor since good performance does not separate investor types. Similarly, attrition should only matter if performance is an important investment criterion that investors care about. We use two different cuts of the data to test this idea as outlined in hypothesis H3.

Our first example is to compare index and active funds. In index investing, relative performance is a significantly less important investment factor than in active fund investments. Investors are not trying to extract information about managerial skill from realized performance in passive investments. In contrast, investment in an active fund relies heavily on using past realized returns to make some interpretation about managerial skill so performance is a much more important criterion for investment. Performance in index funds is almost entirely reliant on the market or index returns and so an investor's decision to invest in an index relies on an asset allocation decision rather than on how a fund performs against its peers.

Table 8 splits the sample into index and active funds. Consistent with the hypothesis that investors in active funds rely more heavily on relative returns in making their investment

decisions, we observe that `RANKED_RETURN` is significantly and positively related to future flows in active investing but not in index investing. In fact, flows to index funds appear to be driven more by category flows, while in actively managed funds category flows are much less important and relative ranked performance is the key investment criterion.

Attrition is a measure of how investors sort themselves based on realized relative performance. Given realized relative performance is a less important investment criterion for index funds, then from H3, we shouldn't expect attrition in index funds to measure the sorting of investors by their performance sensitivities. Table 8 confirms these two hypotheses and this intuition. In comparing Column 1 with Column 2 measuring the flow-performance relation of index and active funds respectively, we notice significant differences in how `ATTRITION` interacts with the `RANKED_RETURN`. Index funds experiencing a lot of attrition exhibit no change in their flow-performance relation while active funds that have had a lot of attrition exhibit a significantly more dampened flow-performance relation. The difference between the two fund types is consistent with the reliance of investors on performance as an investment criterion in active management, but not in passive management. Columns 3 to 6 separately estimate the differences between index and active funds for inflows and outflows and also finds that attrition in active funds negatively affects the sensitivity of both inflows and outflows.

Competition in the category is another dimension which affects how important performance is for investment decisions. When faced with a wide-range of competition, investors with limited cognitive abilities to choose funds along multiple dimensions for hundreds of different funds will compare funds across a single characteristic that makes it easier to rank funds. Either because of bounded rationality or prohibitive time costs, it is impossible for investors to look at all the details and nuances of a large selection of different funds. We argue that an easy decision metric is simply to compare across performance levels. As the number of funds in a category

shrinks then investors may look to other characteristics to supplement their decision based on performance. The premise is that when there is a wide selection of funds to choose from, investors will be unable to easily rank funds based on multiple criteria so will instead focus more on one characteristic, performance.

In a competitive category with a large number of funds competing for investors, we hypothesize that performance will be a more important decision factor for investment than in a category with few competing funds. Table 9 divides the sample into categories with more than and less than the median level of funds in a category. In looking at columns 1 and 2, we observe that in general there is more sensitivity (almost 2.1 times more) to RANKED_RETURN when there is more competition in the category. When we isolate the performance-sensitivity of inflows vs. outflows, we observe that inflows and outflows are almost 2 times more sensitive to performance in the competitive category than not. Taken together, the prior assumption that performance is a more important investment criterion in a competitive category is supported by the data.

What implications does this have for attrition? As predicted in H3, we expect that when performance is a more important investment criterion then attrition is more likely to capture the sorting of investors by performance-sensitivity. Table 9 shows that a fund that has experienced large amounts of attrition appears to more easily lose performance-sensitive investors when competition is high, presumably because investors are more focused on relative performance in choosing across competing funds. Similarly, there is much more sorting of investor types in a competitive landscape when investors are attracted to a fund. In comparing the performance-sensitivities of inflows after attrition, funds with more competition have a significantly lower flow-performance relation than funds with fewer competing funds in the category. Competition

therefore seems to enhance the sorting of investor types since investors who care about performance can more easily find an alternative investment where performance is better.

V. Conclusion

Mutual funds attract a heterogeneous group of investors with different risk-aversions, investment horizons, knowledge, investment objectives, and behavioral biases. The literature has in general treated these investors as homogeneous with some exceptions. For example, Johnson (2004) documents the costs of mixing short-term investors with long-term investors. One of the reasons for the literature not exploring the effects of investor heterogeneity on mutual funds is we know little about the composition of the investor base and how this differs in the cross-section of mutual funds.

Attrition is one measure which helps identify funds with a different set of investor demand characteristics. Christoffersen and Musto (2002) first applied this measure to explain changes in money market fees around fund mergers. This study expands the earlier research by estimating the effects of attrition in equity funds and exploring three additional questions. First, it finds that higher attrition corresponds to lower flow-performance sensitivity, both in inflows and outflows. Consistent with the sorting hypothesis of investor types, we observe that attrition only separates performance-sensitive investors from the pool of investors after consistent bad performance. The flow-performance relation dampens if investors leave after a long period of bad performance but not after good since in the latter case people are not being separated by their performance sensitivities.

Our second set of tests focus on the effect of attrition on convexity. Although convexity is present for both high and low attrition funds, the flow-performance sensitivity of both inflows and outflows is lower for high attrition funds. One insight of this paper is that the flow-performance sensitivity across all attrition types may exhibit greater convexity if high attrition

funds (with low performance-sensitivities) arise more often in poorly ranked funds and low attrition funds (with heightened performance-sensitivities) arise more often in highly ranked funds.

Our final tests focus on conditions when attrition will be particularly helpful in predicting a subsample of performance-insensitive investors. When we split the sample into institutional and retail share classes, we find that attrition only has a significant dampening effect on the flow-performance sensitivity amongst retail flows, not institutional flows, since retail investors are heterogeneous in their performance sensitivities while institutional investors are not.

We also argue that attrition will only change the flow-performance relation if investors are leaving or avoiding a fund for performance reasons. If investment decisions are based on other characteristics of the funds, then attrition will not reflect a separation of investor performance-sensitivity types. To test this, we compare index and active funds and find attrition only has an effect on the flow-performance relation for active funds where relative performance is a more important investment criterion used to distinguish the abilities of active managers. In addition, attrition in more competitive categories coincides more closely with investors leaving for performance reasons since investors rely more on relative performance to choose amongst a wide range of investment alternatives.

Overall, the paper provides compelling evidence that attrition captures this sorting of heterogeneous investors and the conditions when we expect attrition to be a particularly good measure of this sorting. On a practical note, this measure could be immensely helpful to researchers since attrition is such an easy measure to calculate and captures the cross-sectional differences across funds. Practitioners will find this measure useful in trying to determine the stability of their investor base going forward.

References

- Berk, J. B., and I. Tonks. “Return Persistence and Fund Flows in the Worst Performing Mutual Funds.” Working Paper, Stanford University (2007).
- Brown, K. C.; W. V. Harlow; and L. T. Starks. “Of Tournaments and Temptations: An Analysis of Managerial Incentives in the Mutual Fund Industry.” *Journal of Finance*, 51 (1996), 85–110.
- Carhart, M. M. “On Persistence in Mutual Fund Performance.” *Journal of Finance*, 52 (1997), 57–82.
- Carhart, M. M.; R. Kaniel; D. K. Musto; and A. V. Reed. “Leaning for the Tape: Evidence of Gaming Behavior in Equity Mutual Funds.” *Journal of Finance*, 57 (2002), 661–693.
- Chevalier, J., and G. Ellison. “Risk Taking by Mutual Funds as a Response to Incentives.” *Journal of Political Economy*, 105 (1997), 1167–1200.
- Christoffersen, S. E. K. “Why Do Money Fund Managers Voluntarily Waive Their Fees?” *Journal of Finance*, 56 (2001), 1117–1140.
- Christoffersen, S. E. K., and D. K. Musto. “Demand Curves and the Pricing of Money Management.” *Review of Financial Studies*, 15 (2002), 1499–1524.
- Christoffersen, S. E. K.; D. K. Musto; and R. Wermers. “Investor Flows to Asset Managers: Causes and Consequences.” *Annual Review of Financial Economics*, 6 (2014), 289–310.

Cooper, M.; M. Halling; and M. Lemmon. “Violations of the Law of One Fee in the Mutual Fund Industry.” Working Paper, University of Utah (2013).

Elton, E. J.; M. J. Gruber; and J. A. Busse. “Are Investors Rational? Choices among Index Funds.” *Journal of Finance*, 54 (2004), 261–288.

Evans, R. B. “Mutual Fund Incubation.” *The Journal of Finance*, 65 (2010), 1581–1611.

Gil-Bazo, J., and P. Ruiz-Verdú. “The Relation between Price and Performance in the Mutual Fund Industry.” *Journal of Finance*, 64 (2009), 2153–2183

Gormley, T. A., and D. A. Matsa. “Common Errors: How to (and Not to) Control for Unobserved Heterogeneity.” *Review of Financial Studies*, 27 (2014), 617–661.

Ippolito, R. A. “Consumer Reaction to Measures of Poor Quality: Evidence from the Mutual Fund Industry.” *Journal of Law & Economics*, 35 (1992), 45–70.

Lynch, A. W., and D. K. Musto. “How Investors Interpret Past Fund Returns.” *Journal of Finance*, 58 (2003), 2033–2058.

Namvar, E., and B. Phillips. “Commonalities in Investment Strategy and the Determinants of Performance in Mutual Fund Mergers.” *Journal of Banking & Finance*, 37 (2013), 625–635.

Petersen, M. A. “Estimating Standard Errors in Finance Panel Data Sets: Comparing Approaches.” *Review of Financial Studies*, 22 (2009), 435–480.

Samuelson, P. A. “A Note on the Pure Theory of Consumer’s Behavior.” *Economica*, 5, 17 (1938), 61–71.

Samuelson, P. A. “Consumption Theory in Terms of Revealed Preference.” *Economica*, 15, 60 (1948), 243–253.

Schwartz, E. S., and W. N. Torous. “Prepayment and the Valuation of Mortgage-Backed Securities.” *Journal of Finance*, 44 (1989), 375–392.

Sirri, E. R., and P. Tufano. “Costly Search and Mutual Fund Flows.” *Journal of Finance*, 53 (1998), 1589–1622.

Varian, H. R. “Revealed Preference” In *Samuelsonian Economics and the Twenty-First Century*, Michael Szenberg, Lall Ramrattan, and Aron A. Gottesman, eds. Oxford University Press, Chapter 5 (2006).

Wermers, R. “Mutual Fund Performance: An Empirical Decomposition into Stock-picking Talent, Style, Transactions, Costs, and Expenses.” *Journal of Finance*, 55 (2000), 1655–1695.

Figure 1. Non-Linear Flows in Above- and Below-Median Attrition Funds

Figure 1 plots the semi-parametric kernel estimation of fund flow functions and uses the methodology applied in Chevalier and Ellison (1997). Expected flow is expressed in decimals so 0.3 represents 30% annually and is the semi-parametric kernel estimation relating annual fund flow in year $t+1$ to RANKED_RETURN in year t after using the two-stage method of Robinson (1988) to control for the remaining independent explanatory variables of flows outlined in equation (4). Low attrition funds are defined as those with attrition levels below the median, while high attrition funds have attrition above the median level. $ATTRITION_{i,t}$ is the attrition rate of fund i at the end of year t (expressed in decimals) and is defined as 1 less the ratio of fund size in year t divided by the historical maximum fund size over all months prior to the last month in year t .

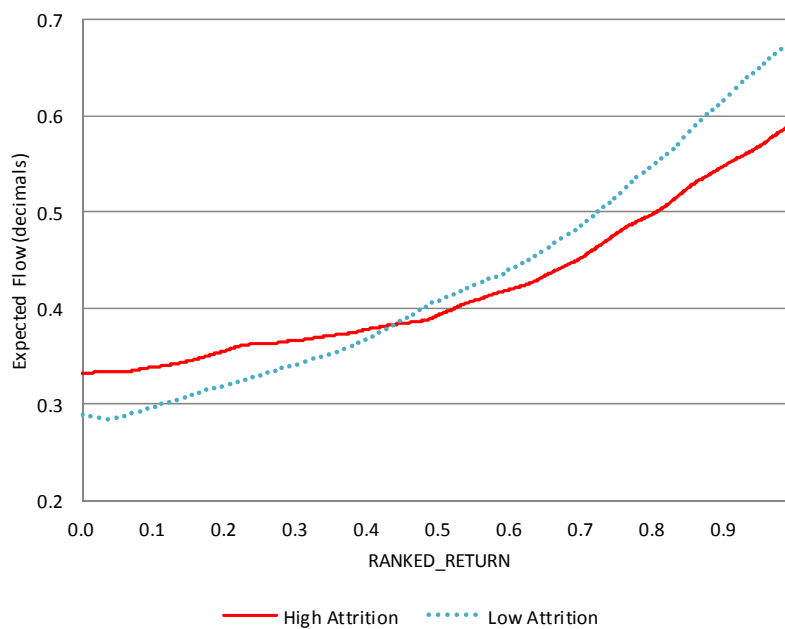


Figure 2. Non-Linear Fund Flows in High and Low Attrition Terciles

Figure 2 plots the semi-parametric kernel estimation of fund flow functions and uses the methodology applied in Chevalier and Ellison (1997). Expected flow is expressed in decimals so 0.3 represents 30% annually and is the semi-parametric kernel estimation relating annual fund flow in year $t+1$ to RANKED_RETURN in year t after using the two-stage method of Robinson (1988) to control for the remaining independent explanatory variables of flows outlined in equation (4). The low (high) attrition tercile is defined as those funds with attrition levels in the bottom (top) tercile of all funds. $ATTRITION_{i,t}$ is the attrition rate of fund i at the end of year t (expressed in decimals) and is defined as 1 less the ratio of fund size in year t divided by the historical maximum fund size over all months prior to the last month in year t .

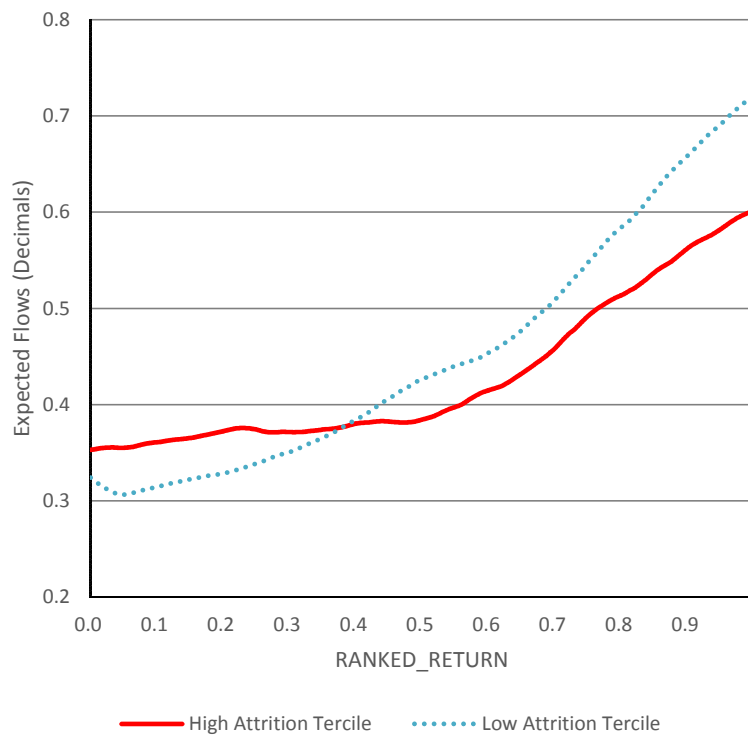


Table 1: Descriptive Statistics

Table 1 reports the summary statistics of fund variables. $ATTRITION_{i,t}$ is the attrition rate of fund i at the end of year t (expressed in decimals) and is defined as 1 less the ratio of fund size in year t divided by the historical maximum fund size over all months prior to the last month in year t . $FUND_SIZE$ is the total net assets of a fund aggregating across all fund shareclasses in the last month of year t (in millions of dollars). The variable $\ln(ACCUMULATED_RETURN)$ is the natural logarithm of one plus the compounded monthly net returns calculated from the end of the month when the fund reached its historical maximum size to the end of year t and is measured in decimals. $EXPENSES$ are calculated as the size-weighted average from each share class and are expressed in decimals so 5% is represented as 0.05. AGE is age of the fund in years from the date the first share class is reported in the data to the end of year t . $INSTITUTIONAL_SHARE$ is the proportion of dollars in institutional share classes at the end of year t as a portion of the total net assets of the fund and is reported in decimals. $FAMILY_SHARECLASS$ is the number of share classes in a fund family, aggregating across all categories of funds in the family at the end of year t . $FUNDS_IN_CATEGORY$ is the number of funds in the same investment category at the end of year t . $FLOW_{i,t+1}$ is expressed in decimals and defined as $FUND_SIZE_{i,t+1} - FUND_SIZE_{i,t} \times (1 + RETURN_{i,t+1})$ divided by $FUND_SIZE_{i,t}$ where $RETURN_{i,t+1}$ is the annual net fund return at the end of year $t+1$. $CATEGORY_FLOW$ is the size-weighted average flow of all funds in the same category (in decimals). In Panel A, we report the variable average, minimum, median, and maximum as well as standard deviation. Panel B reports the correlation between the variables. All variables are determined over year t except $FLOW$ and $CATEGORY_FLOW$ which are measured over year $t+1$.

Panel A: Summary Statistics

	Mean	St. Dev.	Min	Median	Max
ATTRITION	0.334	0.303	0	0.281	0.998
ln(ACCUMULATED_RETURN)	-0.087	0.392	-5.463	-0.014	1.967
EXPENSES	0.0112	0.0059	0.0000	0.0112	0.1242
FUND_SIZE	961.6	4057.0	0.1	209.4	167285.0
AGE	9	4	3	8	39
INSTITUTIONAL_SHARE	0.15	0.33	0	0	1
FAMILY_SHARECLASS	167	218	1	86	1250
FUNDS_IN_CATEGORY	109	81	1	91	306
FLOW	0.084	0.651	-1.475	-0.038	9.951
CATEGORY_FLOW	0.060	0.182	-0.772	0.032	5.306

Panel B: Correlation

	LN ACCUMULATED					
	ATTRITION	RETURN	EXPENSES	FUND_SIZE	FLOW	AGE
ATTRITION	1					
ln(ACCUMULATED_RETURN)	-0.169	1				
EXPENSES	0.199	-0.137	1			
FUND_SIZE	-0.124	0.001	-0.193	1		
FLOW	-0.119	0.066	-0.016	-0.014	1	
AGE	0.297	0.053	-0.052	0.194	-0.122	1

Table 2: Characteristics of Attrition Quintiles

Table 2 reports averages of six variables for five different attrition quintiles. $ATTRITION_{i,t}$ is the attrition rate of fund i at the end of year t (expressed in decimals) and is defined as 1 less the ratio of fund size in year t divided by the historical maximum fund size over all months prior to the last month in year t . $FUND_SIZE$ is the total net assets of a fund aggregating across all fund shareclasses in the last month of year t (in millions of dollars). The variable $\ln(ACCUMULATED_RETURN)$ is the natural logarithm of one plus the compounded monthly net returns calculated from the end of the month when the fund reached its historical maximum size to the end of year t and is measured in decimals. $EXPENSES$ are calculated in year t as the size-weighted average expense from each share class and are expressed in decimals so 5% is represented as 0.05. $FLOW_{i,t+1}$ is defined as $FUND_SIZE_{i,t+1} - FUND_SIZE_{i,t} \times (1 + RETURN_{i,t+1})$ divided by $FUND_SIZE_{i,t}$ where $RETURN_{i,t+1}$ is the annual net fund return at the end of year $t+1$. AGE is age of the fund in years from the date the first share class of the fund is reported in the data to the end of year t . All variables are determined over year t except $FLOW$ which is measured over year $t+1$.

Attrition Quintile	ATTRITION	FUND_SIZE	AGE	EXPENSES	LN ACCUMULATED RETURN	FLOW
1	0	1554	7.1	0.0102	0	0.321
2	0.073	1477	7.9	0.0103	-0.0145	0.063
3	0.277	955	8.6	0.0106	-0.0895	0.027
4	0.521	570	9.5	0.0118	-0.1638	-0.023
5	0.812	245	10.7	0.0133	-0.1638	0.024
Total	0.334	961	8.7	0.0112	-0.0866	0.084

Table 3: Explaining Attrition

Table 3 provides the estimated coefficients from regression and probit models explaining fund-level attrition and an attrition dummy. $ATTRITION_{i,t}$ is the attrition rate of fund i at the end of year t (expressed in decimals) and is defined as 1 less the ratio of fund size in year t divided by the historical maximum fund size over all months prior to the last month in year t .

$ATTRITION_DUMMY$ takes the value 1 if $ATTRITION$ is above its median level and 0 if below. $FUND_SIZE$ is the total net assets of a fund aggregating across all fund shareclasses in the last month of year t (in millions of dollars) and $\ln(FUND_SIZE)$ is its natural logarithm.

$FLOW_{i,t+1}$ is defined as $FUND_SIZE_{i,t+1} - FUND_SIZE_{i,t} \times (1 + RETURN_{i,t+1})$ divided by $FUND_SIZE_{i,t}$ where $RETURN_{i,t+1}$ is the annual net fund return at the end of year $t+1$. The variable $\ln(ACCUMULATED_RETURN)$ is the natural logarithm of one plus the compounded monthly net returns calculated from the end of the month when the fund reached its historical maximum size to the end of year t and is measured in decimals. $EXPENSES$ are calculated as the size-weighted average expenses from each share class at the end of year t and are expressed in decimals so 5% is represented as 0.05. AGE is age of the fund in years from the date the first share class of the fund is reported in the data to the end of year t . $LOAD_DUMMY$ takes the value 1 if a fund has either a front- or a back-load in each year t , and 0 otherwise.

$INSTITUTIONAL_SHARE$ is the proportion of dollars in institutional share classes as a portion of the total net assets of the fund and is reported in decimals at the end each year t . The natural logarithm of the number of share classes in the same family, $\ln(FAMILY_SHARECLASS)$, aggregates share classes across all categories of funds in the family at the end of each year t . The natural logarithm of the number of funds in the same investment category at the end of each year t is denoted as $\ln(FUNDS_IN_CATEGORY)$. Columns 1 and 2 provide regression estimates of

attrition while columns 3 and 4 provide probit estimates of the attrition dummy and column 5 lists the marginal effects of the probit estimation from column 4. The t -statistics are provided in brackets below each coefficient estimate and standard errors cluster by fund with year dummies included in columns 2 and 4. R^2 (and pseudo R^2 values for the probit estimates) are reported in decimals. Coefficients that are significant at the 1% and 5% levels are identified by ** and * respectively.

	ATTRITION		ATTRITION_DUMMY		Marginal Effects
ln(ACCUMULATED_RETURN)	-0.1272** (-6.08)	-0.0649** (-3.08)	-0.7807** (-8.22)	-0.3568** (-3.95)	-0.1046** (-4.00)
AGE	0.0286** (25.90)	0.0283** (19.88)	0.1282** (21.55)	0.1307** (17.25)	0.0383** (20.00)
EXPENSES	3.3829** (2.90)	4.4528** (3.99)	17.3222** (3.37)	26.4716** (4.95)	7.7622** (4.98)
ln(FUND_SIZE)	-0.0886** (-26.09)	-0.0838** (-25.83)	-0.3758** (-21.20)	-0.3626** (-19.93)	-0.1063** (-24.20)
LOAD_DUMMY	-0.0387** (-3.70)	-0.0374** (-3.66)	-0.1843** (-3.85)	-0.1817** (-3.66)	-0.0533** (-3.68)
INSTITUTIONAL_SHARE	-0.0033 (-0.20)	-0.0032 (-0.20)	-0.0074 (-0.10)	0.0117 (0.16)	0.0034 (0.16)
ln(FAMILY_SHARECLASS)	0.0368** (11.37)	0.0343** (10.56)	0.1642** (11.08)	0.1572** (10.14)	0.0461** (10.52)
ln(FUNDS_IN_CATEGORY)	0.0121* (2.19)	0.0086 (1.48)	0.0727** (2.94)	0.0542* (2.02)	0.0159* (2.02)
YEAR DUMMIES	No	Yes	No	Yes	Yes
CONSTANT	0.3334** (9.88)	0.2670** (7.66)	-0.1972 (-1.34)	-0.6354** (-3.84)	
No. of obs.	14799	14799	14799	14799	14799
R^2 (Pseudo R^2)	0.3449	0.3788	0.2038	0.2519	

Table 4: Attrition and Flows

Table 4 provides the estimated coefficients from the regression and tobit models explaining fund-level flows. FUND_SIZE is the total net assets of a fund, aggregating across all fund share classes in the last month of year t (in millions of dollars) and $\ln(\text{FUND_SIZE})$ is its natural logarithm. $\text{FLOW}_{i,t+1}$ is defined as $\text{FUND_SIZE}_{i,t+1} - \text{FUND_SIZE}_{i,t} \times (1 + \text{RETURN}_{i,t+1})$ divided by $\text{FUND_SIZE}_{i,t}$ where $\text{RETURN}_{i,t+1}$ is the annual net fund return at the end of year $t+1$. CATEGORY_FLOW is the size-weighted average flow of all funds in the same category (in decimals) in year $t+1$. RANKED_RETURN ranges from 0 to 1 and is the ranked annual net returns for a fund in year t where funds are ranked against other funds in their category. All remaining variables are defined in Table 3. The first three columns provide regression estimates of FLOW in year $t+1$ while the last two provide tobit estimates of FLOWPOS and FLOWNEG in year $t+1$ which are respectively defined as $\max(\text{FLOW}, 0)$ and $\min(\text{FLOW}, 0)$. The t -statistics are in brackets below each coefficient estimate and standard errors cluster by fund with year dummies included. R^2 (and pseudo R^2 values for the tobit models) are reported in decimals. Coefficients that are significant at the 1% and 5% levels are identified by ** and * respectively.

	FLOW _{t+1}			FLOWPOS _{t+1}	FLOWNEG _{t+1}
RANKED_RETURN	0.3257** (14.26)	0.3089** (13.69)	0.4940** (8.73)	0.9108** (8.62)	0.2044** (9.91)
ATTRITION_DUMMY		-0.0915** (-6.88)	-0.0164 (-0.70)	-0.1161* (-2.16)	-0.0612** (-5.34)
ATTRITION_DUMMY x RANKED_RETURN			-0.1438** (-2.94)	-0.3709** (-4.02)	-0.0708** (-3.77)
AGE	-0.0038* (-2.24)	-0.0009 (-0.53)	0.0051 (1.84)	-0.0143* (-2.06)	0.0006 (0.50)
AGE x RANKED_RETURN			-0.0119* (-2.41)	0.0060 (0.59)	-0.0006 (-0.32)
FLOW _t	0.1809** (5.05)	0.1480** (4.32)	0.1741** (4.76)	0.4372** (6.59)	0.1454** (4.97)
FLOW _t x RANKED_RETURN	0.0152 (0.30)	0.0420 (0.86)	1.48E-05 (0.00)	-0.2248* (-2.56)	-0.1237** (-3.37)
CATEGORY_FLOW _{t+1}	0.4675** (7.33)	0.4651** (7.22)	0.4681** (7.22)	0.8640** (9.14)	0.2015** (8.12)
ln(ACCUMULATED_RETURN)	0.0681** (3.64)	0.0649** (3.44)	0.0711** (3.90)	0.2142** (4.52)	0.0302** (4.02)
ln(FUND_SIZE)	-0.0471** (-8.38)	-0.0551** (-9.36)	-0.0553** (-9.43)	-0.1007** (-8.40)	-0.0006 (-0.32)
EXPENSES	-2.9713 (-1.59)	-2.5210 (-1.41)	-2.5961 (-1.46)	-10.9766* (-2.23)	-2.6891** (-3.09)
LOAD_DUMMY	-0.0103 (-0.78)	-0.0149 (-1.14)	-0.0146 (-1.13)	-0.0608* (-2.03)	-0.0067 (-1.09)
INSTITUTIONAL_SHARE	-0.0597** (-3.47)	-0.0595** (-3.49)	-0.0589** (-3.46)	-0.0511 (-1.43)	-0.0259** (-2.89)
ln(FAMILY_SHARECLASS)	0.0123** (3.17)	0.0160** (4.07)	0.0160** (4.08)	0.0261** (3.31)	-0.0036* (-2.08)
ln(FUNDS_IN_CATEGORY)	-0.0029 (-0.43)	-0.0015 (-0.22)	-0.0017 (-0.26)	-0.0209 (-1.45)	-0.0075* (-2.42)
YEAR DUMMIES	Yes	Yes	Yes	Yes	Yes
CONSTANT	0.1773** (3.15)	0.2151** (3.85)	0.1184* (1.97)	-0.0562 (-0.47)	-0.0994** (-4.08)
No. of obs.	12969	12969	12969	12969	12969
R ² (Pseudo R ²)	0.1078	0.1110	0.1125	0.0920	0.3844

Table 5: Attrition, Flows, and Past Performance

Table 5 provides the estimated coefficients from regression and tobit models explaining fund flows conditional on $\ln(\text{ACCUMULATED_RETURN})$ being positive or negative. The variable $\ln(\text{ACCUMULATED_RETURN})$ is the natural logarithm of one plus the compounded monthly net returns calculated from the end of the month when the fund reached its historical maximum size to the end of year t and is measured in decimals. FUND_SIZE is the total net assets of a fund aggregating across all fund shareclasses in the last month of year t (in millions of dollars) and $\ln(\text{FUND_SIZE})$ is its natural logarithm. $\text{FLOW}_{i,t+1}$ is defined as $\text{FUND_SIZE}_{i,t+1} - \text{FUND_SIZE}_{i,t} \times (1 + \text{RETURN}_{i,t+1})$ divided by $\text{FUND_SIZE}_{i,t}$ where $\text{RETURN}_{i,t+1}$ is the annual net fund return at the end of year $t+1$. CATEGORY_FLOW is the size-weighted average flow of all funds in the same category (in decimals) in year $t+1$. RANKED_RETURN ranges from 0 to 1 and is the ranked annual net return for a fund in year t where funds are ranked against other funds in their category. All remaining variables are defined in Table 3 and the set of control variables include EXPENSES , LOAD_DUMMY , $\text{INSTITUTIONAL_SHARE}$, $\ln(\text{FAMILY_SHARECLASS})$, and $\ln(\text{FUNDS_IN_CATEGORY})$. The first two columns provide regression estimates of FLOW in year $t+1$ while the last four provide tobit estimates of FLOWPOS and FLOWNEG in year $t+1$ which are respectively defined as $\max(\text{FLOW}, 0)$ and $\min(\text{FLOW}, 0)$. The t -statistics are provided in brackets below each coefficient estimate and standard errors cluster by fund with year dummies included. R^2 (and pseudo R^2 values for the tobit estimations) are reported in decimals. Coefficients that are significant at the 1% and 5% levels are identified by ** and * respectively.

	FLOW _{t+1}		FLOWPOS _{t+1}		FLOWNEG _{t+1}	
			ln(ACCUMULATED_RETURN)			
	> 0	≤ 0	> 0	≤ 0	> 0	≤ 0
RANKED_RETURN	0.5761** (5.40)	0.4683** (7.04)	1.3374** (5.35)	0.7799** (6.93)	0.1489** (4.03)	0.2254** (9.14)
ATTRITION_DUMMY	-0.0501 (-1.04)	-0.0127 (-0.46)	-0.0990 (-0.77)	-0.1350* (-2.13)	-0.0443* (-2.43)	-0.0764** (-5.20)
ATTRITION_DUMMY x RANKED_RETURN	0.0797 (0.82)	-0.2343** (-4.27)	-0.0403 (-0.19)	-0.5102** (-5.02)	-0.0090 (-0.29)	-0.0896** (-4.01)
AGE	0.0109* (2.01)	0.0019 (0.61)	-0.0009 (-0.06)	-0.0226** (-3.05)	0.0012 (0.57)	-0.0003 (-0.19)
AGE x RANKED_RETURN	-0.0286** (-2.94)	-0.0065 (-1.14)	-0.0343 (-1.53)	0.0179 (1.61)	-0.0017 (-0.53)	-0.0003 (-0.15)
FLOW _t	0.1435 (1.78)	0.1683** (4.19)	0.4388* (2.37)	0.3838** (5.67)	0.1425* (2.29)	0.1324** (4.27)
FLOW _t x RANKED_RETURN	0.0198 (0.16)	-0.0025 (-0.04)	-0.2034 (-0.87)	-0.1968* (-2.18)	-0.1020 (-1.38)	-0.1273** (-3.30)
CATEGORY_FLOW _{t+1}	0.6801** (5.21)	0.3850** (6.23)	1.4833** (8.95)	0.6718** (6.85)	0.2412** (5.64)	0.1776** (6.28)
ln(ACCUMULATED_RETURN)	0.0874* (2.36)	0.0488 (1.54)	0.4119** (4.54)	0.1110 (1.27)	0.0610** (3.45)	0.0240 (1.95)
ln(FUND_SIZE)	-0.0744** (-6.28)	-0.0491** (-7.39)	-0.1745** (-6.85)	-0.0835** (-6.69)	-0.0079* (-2.39)	0.0011 (0.52)
YEAR DUMMIES	Yes	Yes	Yes	Yes	Yes	Yes
CONTROL VARIABLES	Yes	Yes	Yes	Yes	Yes	Yes
CONSTANT	0.0635 (0.53)	0.1375* (1.98)	-0.4344 (-1.61)	0.0652 (0.51)	-0.1207** (-2.86)	-0.0906** (-3.20)
No. of obs.	3501	9468	3501	9468	3501	9468
R ² (Pseudo R ²)	0.1057	0.1271	0.0875	0.1047	0.3376	0.4283

Table 6: Attrition and Convexity

Table 6 provides the estimated coefficients from regression models explaining piecewise estimations of fund flows. FUND_SIZE is the total net assets of a fund aggregating across all fund shareclasses in the last month of each year (in millions of dollars) and $\ln(\text{FUND_SIZE})$ is its natural logarithm. $\text{FLOW}_{i,t+1}$ is defined as $\text{FUND_SIZE}_{i,t+1} - \text{FUND_SIZE}_{i,t} \times (1 + \text{RETURN}_{i,t+1})$ divided by $\text{FUND_SIZE}_{i,t}$ where $\text{RETURN}_{i,t+1}$ is the annual net fund return at the end of year $t+1$. CATEGORY_FLOW is the size-weighted average flow of all funds in the same category (in decimals) in year $t+1$. RANKED_RETURN ranges from 0 to 1 and is the ranked annual net return for a fund at the end of year t where funds are ranked against other funds in their category. LOW, MED, and HIGH are defined as $\min(\text{RANKED_RETURN}, 0.2)$, $\min(0.6, \text{RANKED_RETURN} - \text{LOW})$, and $\text{RANKED_RETURN} - \text{MED} - \text{LOW}$. We also use a simpler piecewise estimation of flows with only 1 kink at the median ranked return level so in this case LOW50 and HIGH50 are defined as $\min(\text{RANKED_RETURN}, 0.5)$ and $\text{RANKED_RETURN} - \text{LOW50}$. All remaining variables are defined in Table 3 and the set of control variables include EXPENSES, LOAD_DUMMY, INSTITUTIONAL_SHARE, $\ln(\text{FAMILY_SHARECLASS})$, and $\ln(\text{FUNDS_IN_CATEGORY})$. High (low) attrition is defined as those funds with attrition greater than (less than) the median levels of attrition. The t -statistics are provided in brackets below each coefficient estimate and standard errors cluster by fund with year dummies included. R^2 values are reported in decimals. Coefficients that are significant at the 1% and 5% levels are identified by ** and * respectively.

	FLOW _{t+1}			
	High Attrition	Low Attrition	High Attrition	Low Attrition
LOW50	0.1710 (1.74)	0.3694** (5.15)		
HIGH50	0.4914** (4.43)	0.7265** (6.84)		
LOW			0.1675 (0.89)	0.3850* (2.07)
MED			0.2754** (2.90)	0.4917** (6.86)
HIGH			1.0048** (2.78)	1.3446** (3.93)
AGE	0.0031 (0.82)	0.0106** (2.90)	0.0031 (0.83)	0.0109** (2.96)
AGE x RANKED_RETURN	-0.0096 (-1.37)	-0.0208** (-3.26)	-0.0099 (-1.42)	-0.0214** (-3.32)
FLOW _t	0.2197** (2.87)	0.1675** (4.05)	0.2225** (2.89)	0.1727** (4.11)
FLOW _t x RANKED_RETURN	-0.0321 (-0.26)	-0.0056 (-0.09)	-0.0370 (-0.30)	-0.0167 (-0.27)
CATEGORY_FLOW _{t+1}	0.5296** (5.09)	0.4242** (6.06)	0.5268** (5.08)	0.4233** (6.02)
ln(ACCUMULATED_RETURN)	0.0708** (3.41)	0.0349 (0.81)	0.0698** (3.38)	0.0309 (0.72)
ln(FUND_SIZE)	-0.0657** (-7.64)	-0.0431** (-6.14)	-0.0657** (-7.64)	-0.0436** (-6.19)
YEAR DUMMIES	Yes	Yes	Yes	Yes
CONTROL VARIABLES	Yes	Yes	Yes	Yes
CONSTANT	0.2587* (2.46)	0.0533 (0.80)	0.2556* (2.37)	0.0448 (0.62)
No. of obs.	6556	6413	6556	6413
R ²	0.0798	0.1337	0.0800	0.1346

Table 7: Attrition and Institutional and Retail Investors

Table 7 provides the estimated coefficients from fund flow regression models of institutional and retail shareclasses. Unlike the earlier estimations, this table separates institutional from retail share classes. Flows, attrition, age, accumulated returns, fund size, expenses and the load dummy are created following the same methodologies as outlined in Tables 3 and 4 except they are defined using institutional or retail shareclasses separately for each fund: Only asset-weighted institutional (retail) shareclasses are used to create these variables in the institutional (retail) regression. The first two columns provide regression estimates of FLOW in year $t+1$ while the last four provide tobit estimates of FLOWPOS and FLOWNEG in year $t+1$. Columns 1, 3, and 5 estimate net flows, inflows, and outflows for the subset of institutional shareclasses of each fund using their institutional level variables where appropriate and columns 2, 4, and 6 report the same regressions for the subset of retail shareclasses. The t -statistics are provided in brackets below each coefficient estimate and standard errors cluster by fund with year dummies included. R^2 values (and pseudo R^2 values for the tobit estimations) are reported in decimals. Coefficients that are significant at the 1% and 5% levels are identified by ** and * respectively.

	FLOW _{t+1}		FLOWPOS _{t+1}		FLOWNEG _{t+1}	
	Institutional	Retail	Institutional	Retail	Institutional	Retail
RANKED_RETURN	0.3715** (3.34)	0.5512** (9.02)	0.6092** (3.55)	1.0479** (9.25)	0.2333** (3.80)	0.2374** (11.21)
ATTRITION_DUMMY	-0.0814 (-1.35)	-0.0184 (-0.66)	-0.2268* (-2.05)	-0.1570* (-2.49)	-0.1271** (-3.97)	-0.0602** (-5.17)
ATTRITION_DUMMY x RANKED_RETURN	0.0832 (0.74)	-0.1505** (-2.71)	0.0854 (0.46)	-0.3538** (-3.34)	0.0007 (0.01)	-0.0666** (-3.60)
AGE	-0.0146* (-2.53)	0.0078* (2.49)	-0.0441** (-4.43)	-0.0107 (-1.46)	-0.0044 (-1.56)	0.0032** (2.78)
AGE x RANKED_RETURN	-0.0028 (-0.29)	-0.0170** (-3.29)	0.0124 (0.93)	-0.0042 (-0.41)	-0.0013 (-0.29)	-0.0046* (-2.50)
FLOW _t	0.2022** (3.64)	0.1681** (3.67)	0.3001** (4.43)	0.4167** (5.98)	0.1129** (2.89)	0.1500** (5.52)
FLOW _t x RANKED_RETURN	-0.0809 (-0.90)	0.0054 (0.09)	-0.1517 (-1.40)	-0.2046* (-2.27)	-0.0737 (-1.33)	-0.1294** (-3.82)
CATEGORY_FLOW _{t+1}	0.0547 (0.44)	0.5062** (7.97)	0.2639 (1.34)	0.9771** (9.55)	0.0908 (1.37)	0.2182** (8.74)
ln(ACCUMULATED_RETURN)	0.0553 (1.49)	0.0789** (4.19)	0.1660 (1.93)	0.2437** (4.93)	0.0296 (1.18)	0.0295** (3.44)
ln(FUND_SIZE)	-0.0394** (-3.78)	-0.0504** (-8.89)	-0.0568** (-3.39)	-0.0882** (-8.15)	0.0096* (2.29)	-0.0002 (-0.13)
YEAR DUMMIES	Yes	Yes	Yes	Yes	Yes	Yes
CONTROL VARIABLES	Yes	Yes	Yes	Yes	Yes	Yes
CONSTANT	0.3546* (2.56)	0.0595 (1.01)	0.1607 (0.71)	-0.2294* (-2.02)	0.0385 (0.53)	-0.1425** (-6.13)
No. of obs.	3104	11867	3104	11867	3104	11867
R ² (Pseudo R ²)	0.1104	0.1176	0.0765	0.0957	0.1693	0.4209

Table 8: Attrition and Index vs. Active Funds

Table 8 provides the estimated coefficients from regression models explaining fund flows conditional on whether the fund is classified as an index or active fund. INDEX identifies if a fund follows an index strategy while ACTIVE determines if the fund deviates from a passive index strategy and are based on Lipper category classifications. All remaining variables are defined in Tables 3 and 4. The first two columns provide regression estimates of FLOW in year $t+1$ while the last four provide tobit estimates of FLOWPOS and FLOWNEG in year $t+1$. The t -statistics are provided in brackets below each coefficient estimate and standard errors cluster by fund with year dummies included. Control variables include EXPENSES, LOAD_DUMMY, INSTITUTIONAL_SHARE, $\ln(\text{FAMILY_SHARECLASS})$, and $\ln(\text{FUNDS_IN_CATEGORY})$ as described in Table 3. R^2 (and pseudo R^2 values for the tobit estimates) are reported in decimals. Coefficients that are significant at the 1% and 5% levels are identified by ** and * respectively.

	FLOW _{<i>t</i>+1}		FLOWPOS _{<i>t</i>+1}		FLOWNEG _{<i>t</i>+1}	
	Index	Active	Index	Active	Index	Active
RANKED_RETURN	0.0470 (0.35)	0.6082** (9.62)	0.1329 (0.69)	1.1759** (9.16)	-0.0173 (-0.35)	0.2482** (10.90)
ATTRITION_DUMMY	0.0315 (0.49)	-0.0127 (-0.50)	0.0059 (0.06)	-0.1189 (-1.82)	-0.0614* (-2.52)	-0.0521** (-3.87)
ATTRITION_DUMMY x RANKED_RETURN	-0.1021 (-0.82)	-0.1705** (-3.21)	-0.2378 (-1.34)	-0.4425** (-4.17)	-0.0437 (-1.06)	-0.0850** (-3.94)
AGE	-0.0073 (-0.84)	0.0069* (2.39)	-0.0234 (-1.54)	-0.0159* (-1.98)	-0.0042 (-1.42)	0.0004 (0.36)
AGE x RANKED_RETURN	0.0017 (0.12)	-0.0155** (-2.94)	0.0118 (0.56)	0.0049 (0.41)	0.0071 (1.61)	-0.0021 (-1.08)
FLOW _{<i>t</i>}	0.0929 (1.00)	0.2014** (4.89)	0.2072 (1.87)	0.5520** (6.20)	0.0338 (0.74)	0.1766** (4.74)
FLOW _{<i>t</i>} x RANKED_RETURN	-0.0321 (-0.23)	-0.0082 (-0.14)	-0.1346 (-0.82)	-0.3151** (-2.76)	-0.0351 (-0.56)	-0.1450** (-3.05)
CATEGORY_FLOW _{<i>t</i>+1}	0.5419** (4.76)	0.3479** (4.19)	0.8036** (5.97)	0.7414** (5.75)	0.2608** (4.56)	0.1122** (4.32)
ln(ACCUMULATED_RETURN)	0.0779 (1.40)	0.0663** (3.53)	0.0988 (1.12)	0.2500** (4.49)	0.0242 (1.03)	0.0347** (4.41)
ln(FUND_SIZE)	-0.0435** (-4.03)	-0.0570** (-8.78)	-0.0534** (-3.18)	-0.1198** (-8.42)	0.0050 (1.03)	-0.0029 (-1.50)
YEAR DUMMIES	Yes	Yes	Yes	Yes	Yes	Yes
CONTROL VARIABLES	Yes	Yes	Yes	Yes	Yes	Yes
CONSTANT	0.2221 (1.27)	0.0309 (0.44)	-0.0508 (-0.21)	-0.2597 (-1.57)	0.0903 (1.29)	-0.1946** (-6.33)
No. of obs.	2622	10347	2622	10347	2622	10347
<i>R</i> ² (Pseudo <i>R</i> ²)	0.1041	0.1251	0.0713	0.1028	0.2127	0.4529

Table 9: Attrition and Category Competition

Table 9 provides the estimated coefficients from regression models explaining fund flows conditional on the number of funds in a category. The sample is split into those funds in categories with more (less) funds than the median number of funds in a category. Columns 1, 3, and 5 report estimations for funds with more than the median number of funds in the category while columns 2, 4, and 6 report estimations for funds with less than the median number of funds. All variables are defined in Tables 3 and 4. The first two columns provide regression estimates of FLOW in year $t+1$ while the last four provide tobit estimates of FLOWPOS and FLOWNEG in year $t+1$. The t -statistics are provided in brackets below each coefficient estimate and standard errors cluster by fund with year dummies included. Control variables include EXPENSES, LOAD_DUMMY, INSTITUTIONAL_SHARE, $\ln(\text{FAMILY_SHARECLASS})$, and $\ln(\text{FUNDS_IN_CATEGORY})$ as described in Table 3. The t -statistics are provided in brackets below each coefficient estimate and standard errors cluster by fund with year dummies included. R^2 (and pseudo R^2 values for the tobit estimates) are reported in decimals. Coefficients that are significant at the 1% and 5% levels are identified by ** and * respectively.

	FLOW _{t+1}		FLOWPOS _{t+1}		FLOWNEG _{t+1}	
			Number of Funds in Category			
	< Median	≥ Median	< Median	≥ Median	< Median	≥ Median
RANKED_RETURN	0.3199** (3.93)	0.6725** (8.45)	0.5403** (3.95)	1.3664** (8.17)	0.1429** (4.56)	0.2631** (9.85)
ATTRITION_DUMMY	-0.0340 (-0.97)	0.0125 (0.39)	-0.1287 (-1.78)	-0.0596 (-0.74)	-0.0698** (-3.85)	-0.0494** (-3.39)
ATTRITION_DUMMY x RANKED_RETURN	-0.0637 (-0.90)	-0.2297** (-3.37)	-0.2134 (-1.72)	-0.5782** (-4.17)	-0.0462 (-1.58)	-0.0957** (-4.00)
AGE	0.0039 (0.93)	0.0060 (1.59)	-0.0185* (-2.04)	-0.0103 (-0.97)	-0.0011 (-0.70)	0.0008 (0.53)
AGE x RANKED_RETURN	-0.0039 (-0.53)	-0.0208** (-3.08)	0.0204 (1.50)	-0.0142 (-0.92)	0.0028 (1.02)	-0.0041 (-1.71)
FLOW _t	0.1900** (2.97)	0.1603** (4.00)	0.4048** (4.28)	0.4928** (5.13)	0.1447** (3.08)	0.1441** (3.85)
FLOW _t x RANKED_RETURN	-0.0096 (-0.11)	0.0046 (0.07)	-0.2002 (-1.65)	-0.2782* (-2.16)	-0.1290* (-2.25)	-0.1174* (-2.45)
CATEGORY_FLOW _{t+1}	0.4571** (6.05)	0.4549** (3.33)	0.7845** (7.45)	0.8953** (4.83)	0.2239** (7.09)	0.0965* (2.57)
ln(ACCUMULATED_RETURN)	0.0893** (2.96)	0.0671** (2.96)	0.2067** (3.43)	0.2567** (3.63)	0.0322** (2.65)	0.0338** (3.55)
ln(FUND_SIZE)	-0.0645** (-6.95)	-0.0498** (-6.51)	-0.1056** (-6.49)	-0.1066** (-6.19)	0.0022 (0.71)	-0.0035 (-1.54)
YEAR DUMMIES	Yes	Yes	Yes	Yes	Yes	Yes
CONTROL VARIABLES	Yes	Yes	Yes	Yes	Yes	Yes
CONSTANT	0.1617 (1.78)	-0.2946* (-2.05)	0.0581 (0.37)	-0.9532** (-2.93)	-0.0856* (-2.17)	-0.2366** (-3.86)
No. of obs.	6166	6803	6166	6803	6166	6803
R ² (Pseudo R ²)	0.1179	0.1155	0.0927	0.0962	0.3251	0.4767