On the relation between flows and performance in the mutual fund industry

Vijay Yadav[1]

Svetoslav Covachev[2]

Abstract

We investigate the impact of past performance of fund flows using a data set that contains the net flows of mutual funds, as well as the two main components of net flows viz. new sales and redemptions. Using the fractional flow model of Sirri and Tufano (1998), we find that the relation is convex in the mid to high performance range due to the high sensitivity of new sales to good performance and concave in the low to mid performance range mainly due to the high sensitivity of redemptions to poor performance. We find a similar nonlinear flow-performance relation using the "change in market share" measure of Spiegel and Zhang (2013), contrary to their finding of a linear relation. Finally, we find a similar nonlinear flow-performance relation using measures of fund flows that are independent of fund size, namely a fund's share of new sales to aggregate new sales and a fund's share of redemptions to aggregate redemptions.

JEL Classification:

Keywords: Mutual funds, Flow-performance relation, New sales, Redemptions

[1] ESSEC Business School, 2 One-North Gateway, Singapore 138502; Email: yadav@essec.edu.

[2] ESSEC Business School, Avenue Bernard Hirsch, 95021 Cergy-Pontoise Cedex, France; Email: svetoslav.covachev@essec.edu.

1. Introduction

Most of the finance literature on the relation between mutual fund flows and past performance uses the fractional flow model. The dollar flow in the current period is divided by the net assets under management at the beginning of the period to calculate the fractional flow (Chevalier and Ellison (1997), Sirri and Tufano (1998), Fant and O'Neal (2000), Coval and Stafford (2007), Huang, Wei, and Yan (2007), Ferreira, Keswani, Miguel, & Ramos (2012), Berggrun and Lizarzaburu (2015)). Many of these studies document a nonlinear relation between mutual fund flows and past performance. However, Spiegel and Zhang (2013) claim that the fractional flow model used to estimate the flow-performance relation is misspecified. They argue that estimating fund flows using net flows as a percentage of assets leads to spurious convexity in the flow-performance relation and suggest a new measure, "change in market share", to estimate the flow-performance relation. They find that the flow-performance is linear. In this paper, we estimate the flow-performance relation using both the fractional flow model and the "change in market share" model. We find that the flow-performance relation is nonlinear, even if we use the change in market share as a measure of flow. We also show mathematically that "change in market share" can be expressed as a linear function of fractional flows. Therefore, the two models must give the same result.

We improve upon the previous studies in this field by using data on actual amounts of new sales into and redemptions out of individual funds. The studies so far have estimated the net flows of an individual fund over a period by the change in its assets under management after accounting for the change in its assets due to return. By utilizing actual new sales and redemptions, we are able to provide new insights into the dynamics of fund flows in the mutual fund industry.

We start by measuring the performance sensitivity of flows across the mutual funds. We use a piecewise linear regression model by dividing funds into three terciles each month according to their style adjusted performance in the last 12 months. Consistent with the majority of previous literature, we find a nonlinear relation between net flows and past performance. However, our findings suggest that the sensitivity of net flows to performance is very high in the low performance and high performance terciles and relatively modest in the mid performance tercile. This result is different from some earlier studies that show that flow-performance relation is uniformly convex (for example, Sirri and Tufano (1998)). We also estimate the flow-performance relation for the two main components of net cashflows, viz. new sales and redemptions. Using separate analyses on new sales and redemptions, we find that the new sales increase at a very high rate as performance increases for funds in the lowest and highest performance terciles whereas redemptions increase at a very high rate as the performance decreases for funds in the lowest tercile. Overall, the results suggest that the flow-performance relation is nonlinear due to the combined effect of new sales into good funds and redemptions from bad funds, where good and bad are defined based on past one year style-adjusted performance. The flow-performance relation is concave in the low to mid performance range and convex in the mid to high performance range.

An important contribution of our paper is the investigation of the market share model of flow-performance relation proposed by Spiegel and Zhang (2013). Most of the finance literature uses the fractional flows in the flow-performance studies. The fractional flows are calculated by dividing the dollar flows in the current month by the net assets under management at the end of the previous month. Spiegel and Zhang (2013) claim that the fractional flow model to estimate flow-performance relation is misspecified. They claim that estimating fund flow using net flows as a percentage of assets leads to spurious convexity in the flow-performance relation. They suggest a new measure, change in market share, to estimate the flow-performance relation. We follow their formula to calculate the change in market share for each fund-month observation. We use the same regression specifications as in Spiegel and Zhang (2013) to estimate the flow-performance relation. Rather surprisingly, we find results contrary to their results. Our regressions show that the flow-performance relation is convex even if we use the change in market share as a measure of flow.

Spiegel and Zhang (2013) argue that the flow-performance relation using a fractional flow model is linear for hot funds (young and small funds) and cold funds (old or large funds) separately. The slope is higher in the case of hot funds. They claim that if we run regressions with both kinds of funds in a combined sample then the flow-performance relation appears convex. They suggest running separate analyses for hot funds and cold funds. Following Spiegel and Zhang (2013), we define hot funds as those with age less than five years and fund size below the median in the cross-section of funds each month. We run separate regressions for hot funds and cold funds. As predicted by Spiegel and Zhang (2013), the sensitivity of flows to performance is higher in the case of hot funds. However, we find that the flow-performance relation is nonlinear for hot funds as well as for cold funds. In both sub-samples, the relation is concave in the low to mid performance range and convex in the mid to high performance range.

Although our empirical results are contrary to those of Spiegel and Zhang (2013), we agree with their argument that an estimate of the flow-performance relation should not be dependent on whether large or small funds have recently performed relatively well. We propose a simple model based on the share of fund new sales and redemptions of the aggregate new sales and redemptions respectively. For each fund-month observation, we define NEWSALES_SHARE= (fund new sales)/(aggregate new sales) and REDEMPTIONS_SHARE= (fund redemptions)/(aggregate redemptions), where aggregate means the sum over all funds under consideration. In other words, we are simply measuring the fraction of aggregate new sales or redemptions accounted for by each fund. NEWSALES_SHARE and REDEMPTIONS_SHARE depend only on the dollar flows to funds and not on the fund size. Both of these flow share measures lie between 0 and 1 and the sum of either over all funds in a month equals one. Our 'flow share' measure directly measures the dollar flows to funds without being affected by the size of the fund, and therefore addresses a major criticism of the fractional flow model. We estimate the sensitivity of NEWSALES_SHARE and REDEMPTIONS_SHARE to past performance. Quite surprisingly, we find that the relation between NEWSALES_SHARE and past performance is nonlinear with higher slopes in the low performance tercile and high performance tercile. Also, REDEMPTIONS_SHARE is highly sensitive to performance in the low performance tercile and the sensitivity decreases monotonically as we move to the mid and high performance terciles. We conclude that the flow-performance relation is very robust across model specifications.

We perform several robustness checks for our results on the flow-performance relation based on the fractional flow model. In our first robustness test, we divide the sample into months with positive aggregate flows and months with negative aggregate flows. We find that the nonlinear flow-performance relation holds in both sub-samples. In our second robustness test, we want to confirm that the nonlinearity in flow-performance relation holds true in all years and is not concentrated in only a few years. We estimate the flow-performance relation for each year from 2000 to 2013. We find that the nonlinearity in the flow-performance relation holds every year separately. In these Fama-Macbeth regressions, we have only 12 monthly coefficients for each year. Therefore, it might be suspected that the coefficients are not statistically significant even if the shape of the flow-performance curve is nonlinear every year. However, we find that the coefficients of the performance rank variables in the low performance tercile and high performance tercile are statistically significant in all years. This is a very strong result because it suggests that the nonlinear relation between flows and performance is a permanent characteristic of the mutual fund industry that holds true in all kinds of market conditions. In our third robustness test, we estimate the flow-performance relation for each investment style separately. We find that it is nonlinear for funds within each investment style. Our battery of tests confirms that the relation between flows and performance is nonlinear and that this relation is very robust.

This is not the first paper to use the N-SAR data on new sales and redemptions of mutual funds. Cashman et al (2012) use this data to study the flow-performance relation in mutual funds but their analysis suffers from poor data quality. They match N-SAR data to CRSP data manually. However, in the absence of any common identifier in the N-SAR files and the CRSP data, their manual matching procedure leads to a large number of unmatched funds. Their sample contains only 88,910 fund-month observations from 3,214 funds from April 1997 through December 2007. In comparison, the sub-sample of our data for the period January 2000 to December 2007 contains 199,341 fund-month observations for 3,130 funds. The number of observations per fund in their data is much smaller which seems to suggest that there is a large number of missing observations in their data. Cashman et al (2012) report that the correlation between net flow measure using N-SAR data and the Sirri and Tufano (1998) net flow proxy is 0.501 in their sample. We find this correlation to be 0.888 in our sample. This gives us more confidence about the high quality of our data. O'Neal (2004), Bergstresser and Poterba (2002) use self-collected data on new sales and redemptions but their datasets are limited to less than 200 funds. Christoffersen et al (2013) and Christoffersen et al (2007) also use new sales and redemptions data from N-SAR filings, but they do not focus on the flow-performance relation in their studies.

There is very little evidence of performance persistence among mutual funds, when excluding the persistence of the poor performance of the worst mutual funds (Carhart (1997)). Even the mutual funds accept this fact by including a disclaimer in their prospectuses that reads something like "past performance is not an indication of future performance." It is, therefore, rather surprising that individuals tend to invest disproportionately in funds with relatively better performance in the recent past. Berk and Green (2004) resolve this apparent contradiction in their model by assuming that investors rationally interpret high past performance as evidence of the fund manager's superior ability and the superior performance of the fund does not persist because new money inflows from investors, accompanied with decreasing returns to scale, make the future excess returns of funds competitive. Del Guercio and Tkac (2002) propose that the previous year return chasing behavior may be peculiar to the clientele of mutual funds and may not apply to the sponsors of pension funds. Gruber (1996) argues that past performance is a useful indicator for predicting future performance and rational "sophisticated investors" exploit this signal to earn superior returns by allocating their cash accordingly. However, Gruber's (1996) explanation does not rule out that some investors are "unsophisticated" or may not be able to act as "sophisticated investors" due to the constraints they face. Lynch and Musto (2003) provide a rational explanation for a convex shape of the flow-performance relation based on mutual fund return predictability and persistence following good performance, but not after bad performance. However, the predictability and persistence of the returns of mutual funds in the studies of Gruber (1996) and Lynch and Musto (2003) contradict the findings of Carhart (1997). Furthermore, Jaiprakash and Kumar (2009) do not find a strong link between "predictive and macroeconomic variables" and fund flows. Ivkovic and Weisbenner (2009) show that "tax considerations" are important drivers of the sensitivity of mutual fund redemptions to past performance. Regardless of the factors that drive the flow-performance relation, redemptions from funds inflict a cost on the remaining investors in the funds and lower the expected profitability of the funds (Chordia (1996), Johnson (2004)). Furthermore, Greene and Hodges (2002) show that the short-term flows of international mutual funds “dilute the returns” to the longer horizon investors of those funds. Therefore, the flow-performance relation is of economic significance to multiple stakeholders. In this paper, we empirically estimate the flow-performance relation in the cross-section of mutual funds, but do not attempt to provide a rational explanation of the performance chasing behavior of investors.

The paper proceeds as follows. In Section 2, we describe our data and discuss summary statistics. In Section 3, we explain the empirical results for the fractional flow-performance model and the market share based model. In Section 4, we propose a new flow measure based on the monthly new sales and redemptions of funds as a fraction of the monthly aggregate new sales and redemptions for all funds. A brief conclusion follows in Section 5.

2. Data

The primary data source is Morningstar Direct. Our data is free of survivorship bias because we are able to obtain data on surviving as well as non-surviving funds from Morningstar Direct. Following the literature, we focus on the US domestic equity mutual funds. We consider all the funds reported as “Open-end Funds” over the period 2000-2013. Specifically, we include a fund in our sample if its domicile is “United States” and its US_Broad_Asset_Class is ‘U.S. Stock’ and its Morningstar category is one of the following: 'Large Blend', 'Large Growth', 'Large Value', 'Mid Blend', 'Mid Growth', 'Mid Value', 'Small Blend', ‘Small Growth’ and ‘Small Value’. We exclude index funds. We exclude funds that have less than $10 million in assets under management. Finally, we also exclude funds that are less than one year old.

We aggregate data on different share classes of a fund in a month to create a single fund observation. For the new observation, fund size (FUNDSIZE) is defined as the sum of the assets under management of all share classes of the fund. In practice, we do not need to calculate fund TNA from its share classes because Morningstar data reports the aggregate fund size of the funds. Fund age (AGE) is defined as the number of years since inception of the oldest share class of the fund. Expense ratio (EXPRATIO), turnover ratio (TURNOVER) and monthly return (RETURN) for the new observation are calculated as the weighted average of the corresponding figures of all share classes, the weights being determined by the lagged TNA of the share classes. The family total size (FAMTOTAL) for a fund is calculated as the total TNA managed by all the funds in the family. The family size (FAMSIZE) for a fund is calculated as the total TNA managed by all the funds in the family of the fund except the fund itself. We use log of one plus the family size in the regressions.

In the finance literature, fund flow over a period is calculated as the change in fund size over the period adjusted for the change in size due to the fund return. The finance literature has focused mainly on the net flows of the funds without separating out the new sales and redemptions. This is mainly due to a lack of data on new sales and redemptions in the CRSP database which is the most popular database used by academics. However, net cashflows provide an incomplete picture of the flows in the mutual fund industry. For example, a net cashflow of USD 50,000 could be due to new sales of USD 70,000 accompanied by redemptions of USD 20,000 or alternatively it could be due to new sales of USD 170,000 accompanied by redemptions of USD 120,000. These two scenarios portray two completely different kinds of flow dynamics. A quick look at the 2013 ICI Factbook reveals that both the new sales and the redemptions are of significant magnitude. For example, the ratio of annual redemptions to annual new sales of the equity mutual funds has ranged from 84% in the year 2000 to 111% in the year 2012 during our sample period (2013 ICI Factbook, page 161, Table 20). Therefore, both the new sales and the redemptions are important components of net cash flows. Fortunately, funds report their monthly new sales and redemptions to SEC in their N-SAR filings and we are able to obtain this data through Morningstar Direct.

Investment companies report the monthly sales and purchases of their shares to the SEC in the N-SAR filings. Morningstar Direct data contains the monthly New Sales (total NAV of shares sold) and Redemptions (total NAV of shares redeemed) of funds as reported in their NSAR filings. We calculate fractional new sales and fractional redemptions for a fund in a month as the new sales and redemptions of the fund during the month, divided by the fund size at the end of the previous month. Morningstar defines Net Cash Flow as New Sales plus Other Sales minus Redemptions from the NSAR report. However, Other Sales is zero for all observations, therefore Net Cash Flow is simply New Sales minus Redemptions. The total assets under management of a fund may also be affected by the payment of dividends by funds and the reinvestment of dividends by investors. In this study, we have ignored the flows due to dividends, which are quite small in comparison to new sales and redemptions.

The traditional measure of monthly flows, as used in Sirri and Tufano (1998), is based on the change in the assets under management after adjusting for returns:

$$TNA\_FLOW_{it} = \frac{n_{i,t}}{n_{i,t-1}} - (1 + r_{i,t}) \tag{1}$$

where $n_{i,t}$ is the size of fund i at the end of month t, $n_{i,t-1}$ is the size at the end of month t-1, and $r_{i,t}$ is the fund's return for month t.

We define New Sales Flow and Redemptions Flow for fund i in month t as the dollar values of new sales and redemptions respectively divided by the net assets under management at the end of the previous month:

$$NEWSALES\_FLOW_{i,t} = \frac{NewSales_{i,t}}{n_{i,t-1}} \tag{2}$$

$$REDEMPTIONS\_FLOW_{i,t} = \frac{Redemptions_{i,t}}{n_{i,t-1}} \tag{3}$$

where $NewSales_{i,t}$ and $Redemptions_{i,t}$ are the dollar values of new sales and redemptions for fund i in month t.

Net Cash Flow is simply the difference between New Sales Flow and Redemptions Flow.

$$NET\_CASHFLOW_{i,t} = NEWSALES\_FLOW_{i,t} - REDEMPTIONS\_FLOW_{i,t} \tag{4}$$

We report some descriptive statistics of the key variables in our study in Table 1. We tabulate the mean, median, standard deviation, and 1st and 99th percentile values of the variables. The first variable is the percentage flow calculated using the formula $(TNA_{i,t}/TNA_{i,t-1})-(1+r_{i,t})$. The next three variables are the percentage net cashflow, percentage new sales and percentage redemptions. We note that the mean, median, standard deviation, and 1st and 99th percentiles of the two measures of net flows are close to each other. The mean monthly net sales and redemptions are 3.28% and 2.81% respectively resulting in a net cash flow of 0.45% per month. The new sales and redemptions are quite large compared to the net cash flows. Moreover, redemptions are almost as big as new sales. Therefore, an understanding of the dynamics of new sales and redemptions is important to fully understand the dynamics of net cash flows. The mean TNA of funds is USD 1,321 million, but the median TNA is only USD 286 million. This shows that the distribution of TNA is positively skewed with a large number of smaller funds and a small number of very large funds. The mean family total size is USD 33.5 billion. There are some very large fund families and a large number of smaller fund families. On average, a fund family includes around 31 actively managed equity funds. The mean gross expense ratio and net expense ratio are 1.31% per year and 1.20% per year respectively. The mean expense waiver by funds is 0.14% per year. The mean annual turnover, defined as the minimum of the total amount of new securities purchased and the total amount of securities sold divided by average fund TNA over the year, is equal to 81% per year. The funds in our sample are on average 14 years old. The mean monthly return is 0.71% per month. The volatility of fund returns is calculated as the 12-month rolling standard deviation of monthly returns, and averages 4.80%. We winsorize all variables below at the 0.5% level and above at the 99.5% level within each month for our analysis.

[Table 1 here]

3. Fund flows and past performance

In this section, we estimate the flow-performance relation using the fractional flow model of Sirri and Tufano (1998) and the market share based model of Spiegel and Zhang (2013). We also provide several robustness tests for the fractional flow model.

A. Fractional flow model

We start with an analysis of the effect of relative performance on net flows of individual mutual fund flows. We follow the same methodology as in Sirri and Tufano (1998) with some minor differences. The most important difference is that we use monthly flows whereas Sirri and Tufano (1998) use annual flows. Within each month, we rank funds within the same investment style based on their performance in the previous 12 months. We define the fund's performance rank variable RANK as the fractional rank of the fund in the cross-section of funds of the same style in a given month on a scale of 0 to 1. We define the following variables for our regressions: LOWPERF = Min(0.33, RANK), MIDPERF = Min(0.33, RANK-LOWPERF) and HIGHPERF = RANK-(LOWPERF+MIDPERF).

The finance literature has used several specifications to measure the flow-performance relation. Some authors, including Sirri and Tufano (1998), define LOWPERF, MIDPERF1, MIDPERF2, MIDPERF3, and HIGHPERF based on the performance rank cut-off points at 0.2, 0.4, 0.6, and 0.8. However, in most of their analysis, they combine the three mid-performance variables into one single variable MIDPERF and thereby use only three variables LOWPERF, MIDPERF and HIGHPERF with cut-off points at 0.2 and 0.8. Some authors use LOWPERF, MIDPERF and HIGHPERF with cut-off points at 0.33 and 0.67 (for example, Cashman et al (2012)). Some other authors, perhaps due to smaller sample size or because the flow-performance relation is not their focus, use only one cut-off point at 0.5 and use only LOWPERF and HIGHPERF (for example, Christoffersen et al (2013)). The other major variation used by some authors is to directly use a RANK variable and include the square of RANK in the regressions to estimate nonlinear characteristic of the flow-performance relation. We have defined LOWPERF, MIDPERF and HIGHPERF with cut-off points at 0.33 and 0.67. For most of our analyses, we present results of a piecewise linear model that includes LOWPERF, MIDPERF and HIGHPERF as the performance variables as well as a continuous model that includes RANK and RANK_SQUARE (square of RANK) as the performance variables.

We regress percentage NET_CASHFLOW and percentage TNA_FLOW on the performance rank variables and a set of control variables. The flows into a fund may be simply due to the fact that the investment style to which the fund belongs is popular among the investors in the given month. Therefore, we control for the percentage flows to all funds in the same investment category as the fund. We calculate category flow as the total dollar flow to all the funds in the category during the month divided by the total assets of all funds in the category at the beginning of the month. We include the lagged value of the dependent variable in the set of control variables. We also control for lagged values of the fund size, family size, annual turnover ratio, net expense ratio, and volatility of raw returns of the fund in the previous year. The coefficients of the performance variables are not affected much if we use the gross expense ratio in place of the net expense ratio. The output of several regression specifications is summarized in Table 2. The results in column (1) confirm a nonlinear relation between current month net cashflows and past performance. All three performance variables LOWPERF, MIDPERF, and HIGHPERF are statistically significant. The magnitudes of the slope coefficients of these variables suggest that the sensitivity of flows to performance is high in the low performance range, modest in the mid performance range and very high in the high performance range. The results in column (2) provide further evidence that the relation is nonlinear. We prefer to use the term nonlinear rather than convex to describe the flow-performance relation because the relation is concave in the low to mid performance range and convex only in the mid to high performance range. This shape of the flow-performance relation is different from that in Sirri and Tufano (1998) who do not find any significant flow-performance sensitivity in the low performance range and consistent with the nonlinear shape of the relation in Chevalier and Ellison (1997).

[Table 2 here]

We would like to confirm that the results based on our two measures of net fund flows, viz. TNA_FLOW (based on change in TNA adjusted for returns) and NET_CASHFLOW (based on data from N-SAR filings) are consistent. We run the same regressions with the dependent variable TNA_FLOW in place of NET_CASHFLOW. The fact that there is very little difference between the results in the column pairs (1)-(2) and (3)-(4) suggests that TNA_FLOW is a good estimate of NET_CASHFLOW, which is the actual flow to the fund.

[Table 3 here]

Next, we focus on the two components of net cashflows, i.e. new sales and redemptions. We acknowledge that previous attempts to study separately the sensitivities of new sales and redemptions of mutual funds to past performance have produced highly asymmetric results. At the extreme, Johnson (2007) finds no performance sensitivity in the case of redemptions, despite finding a strong performance sensitivity of new sales. However, that study is limited to just one mutual fund family, which makes the generalization of its results questionable. We regress percentage new sales (NEWSALES) and percentage redemptions (REDEMPTIONS) on the performance rank variables. The control variables in these regressions are similar to those in our earlier regressions with some minor modifications. First, we replace the lagged value of net cash flows with its two components, viz. the lagged value of new sales and the lagged value of redemptions. Second, we include as a control variable the average value of the dependent variable for all funds with the same investment objective.

We run these regressions on the set of observations for which all three flow variables (net cashflow, new sales and redemptions) are available. The results are reported in Table 3. Column (1) shows that new sales are very sensitive to performance in the low as well as in the high performance regions. The coefficients of HIGHPERF and LOWPERF are statistically significant. However, the coefficient of HIGHPERF is much higher in magnitude, suggesting a very strong performance sensitivity of new sales in the highest performance region. Column (2) suggests that the performance sensitivity of redemptions is the highest in the low performance region, although the coefficients are statistically significant in the mid and high performance regions as well. Net cash flow is equal to new sales minus redemptions. Column (3) represents the aggregate effect of performance on new sales and redemptions by regressing percentage net cash flows on the performance rank variables. The sensitivity of redemptions dominates that of new sales in the low and mid performance range whereas the sensitivity of new sales dominates that of redemptions in the high performance range. The net effect on new sales and redemptions is that the net cash flows are very sensitive to performance in the low as well as in the high performance regions. The sensitivity in the mid performance range, although statistically significant, is smaller in magnitude. It follows that the shape of the net cash flow to performance curve is concave in the low to mid performance range and convex in the mid to high performance range. We conclude that investors punish the funds with very poor past performance by withdrawing their money and reward the funds with very good past performance by investing heavily in these funds.

In models (4)-(6), we run continuous models to estimate the flow-performance relation by including RANK and RANK_SQUARE as the main explanatory variables. Since our regression equation is of the form Flow = a + b*RANK + c*RANK_SQUARE, the sensitivity of flow to performance rank is equal to b+2*c*RANK. The sensitivity of new sales to performance, as according to column (4), at a given value of RANK is equal to -1.28+2*3.07*RANK, which is increasing in RANK. We need to be careful in interpreting this result. The sensitivity is negative for values of RANK below 0.21, but this is just a shortcoming of the continuous model. The quadratic form of the continuous model cannot fully capture the fact that the actual flow-performance relation is concave in the lower range of RANK and convex in the higher range of RANK. Similarly, the sensitivity of redemptions to RANK in model (5) is given by -1.71+2*1.05*RANK, which is negative for values of RANK below 0.81, but positive for values of RANK above 0.81. It is highly negative in the low performance regions and decreases in magnitude as RANK increases. Therefore, the results of the continuous model broadly confirm the results of the piecewise linear model.

B. Market share based model

The sensitivity of flows to past performance of funds plays a critical role in managerial incentives and risk-taking behavior (Brown, Harlow, & Starks (1996), Chevalier and Ellison (1997), Taylor (2003), Basak and Makarov (2012), Koijen (2014)). Furthermore, Berk and Tonks (2007) propose that the performance persistence of mutual funds, as in Carhart (1997), could be linked to “an attenuation in the flow of funds relation”. The finance literature has provided extensive evidence in support of a nonlinear flow-performance relationship in the cross-section of mutual funds. Most of the papers use percentage flows as the measure of fund flows. However, the fractional flow model is not without its critics. Therefore, it is important to address some other important research on this issue. In particular, there have been studies that show that the flow-performance relation is actually linear. Arguably, the most important contribution in this direction is by Spiegel and Zhang (2013). They claim that the fractional flow model used to measure the flow-performance sensitivity is misspecified, and suggest the use of ‘change in market share’ to estimate the flow-performance sensitivity. They find that the flow-performance relation is linear and therefore not convex.

Spiegel and Zhang (2013) define the change in the market share of a fund in month t as follows:

$$\Delta m_{it} = \frac{n_{it}}{\widehat{N}_t} - \frac{n_{i,t-1}}{N_{t-1}} \tag{5}$$

The left-hand side of Eq. (5) captures the change in a fund's market share due to both the flows it receives and the returns generated to garner them. N_t and N_{t-1} are based on only the funds that are in existence in period t-1. The arc over the N_t term indicates that the term captures the aggregate assets under management of all the funds, that were present at time t-1, as of time t.

Using the change in market share to measure the fund flows, Spiegel and Zhang (2013) show that the flow-performance relationship is linear. This is in contrast to several other studies that use the fractional flow model to prove that the flow-performance relation is nonlinear. Given our focus on the shape of the flow-performance relationship, we estimate it using the measure of change in market share.

[Table 4 here]

In contrast to the findings of Spiegel and Zhang (2013), we find that the flow-performance relationship is nonlinear even if we use the change in market share as our measure of fund flows. For our sample, we utilize the Spiegel and Zhang (2013) methodology to estimate the flow-performance relation and report the results in Table 4. For the sake of comparison, we run similar regression models as in Spiegel and Zhang (2013). In regressions (1)-(4), we use a piecewise linear regression model in which the main independent variables are LOWPERF, MIDPERF and HIGHPERF. In regressions (5)-(8), we use a continuous model in which the main independent variables are RANK and RANK_SQUARE. Our results are starkly different from those in Spiegel and Zhang (2013). We find a nonlinear flow-performance relation even with the use of the 'change in market share' measure. In the piecewise linear regressions, the coefficients of LOWPERF and HIGHPERF are high and that of MIDPERF is relatively low, which means that the flow-performance curve is concave in the low to mid performance range and convex in the mid to high performance range. The coefficient of HIGHPERF is more than double that of MIDPERF in each regression, which leaves no doubt about the convexity of the flow-performance relation in the mid to high performance region. In case of the continuous model, the coefficients of RANK and RANK_SQUARE are positive and statistically significant, at the 5% and 1% levels respectively, in all regressions. The positive coefficient of RANK_SQUARE suggests a flow-performance relation that has a dominant convex characteristic. However, unlike the piecewise regressions, the continuous regressions are unable to separate out the regions of concave and convex relations.

In Appendix 1, we show mathematically that "change in market share" for a fund can be expressed as a linear function of its fractional flow. Our claim is that if the relation between fractional flows and past performance is nonlinear then the relation between "change in market share" and past performance must also be nonlinear. The regressions in Appendix Table 1 prove our claim.

[Table 5 here]

Spiegel and Zhang (2013) provide several arguments for why an estimate of the flow-performance relation based on fractional flows may be misspecified. They categorize funds into hot funds and cold funds. The hot funds are represented by young and small funds, while old and large funds are classified as cold funds. They argue that the flow-performance relation is linear for the hot funds as well as for the cold funds. However, the slope of the flow-performance relation for hot funds is much higher than that for cold funds. Therefore, if we include both types of funds in the same regression, the flow-performance relation may appear convex. They run separate regressions for young and small funds and the rest of the funds, and claim to find little evidence of nonlinearity for either set of funds. Based on their idea, we run separate regressions for the hot funds and cold funds in our sample. Hot funds are defined as those that are less than five years old and are below median in a ranking of funds according to their size within each month. The funds that are not hot are classified as cold funds. The results are reported in Table 5. Models (1)-(3) are for cold funds and models (4)-(6) are for hot funds. As predicted by Spiegel and Zhang (2013), the coefficients of the performance rank variables are indeed higher in the case of hot funds. However, contrary to their predictions, we find from columns (3) and (6) that the flow-performance relation is nonlinear for both the cold funds and the hot funds. These results are also confirmed by the continuous regressions in Panel B in which we have used RANK and RANK_SQUARE as the performance variables. The coefficients of RANK and RANK_SQUARE are higher in the case of hot funds. However, the coefficients of RANK_SQUARE in columns (3) and (6) are positive and statistically significant at the 1% level, which shows that the flow-performance relation is convex for both the cold funds and the hot funds.

We conclude that, despite the forceful logic of Spiegel and Zhang (2013), the empirical evidence is in favor of nonlinearity of the flow-performance relation. In our sample, the flow-performance relation appears to be nonlinear using the fractional flow model of Sirri and Tufano (1998) as well as the change in market share model of Spiegel and Zhang (2013).

C. Robustness Tests

[Table 6 here]

The flow-performance relation may change dynamically over a period of time. In particular, the flow-performance relation may depend on the aggregate flows in the mutual fund industry. The flow-performance dynamics during periods of negative aggregate flows may be very different from those during periods of positive aggregate flows. Fortunately, our sample period contains a large number of months with negative as well as positive aggregate flows. We define aggregate flow for a month as the total net dollar flows to all funds in our sample during the month. There are 94 months with negative aggregate flows and 74 months with positive aggregate flows during our sample period. We divide the 168 months in our sample into three equal groups – the 56 months with the most negative aggregate flows, the 56 months with the most positive aggregate flows, and the 56 months in the middle, of which 38 have negative and 18 have positive aggregate flows. We run separate regressions for the three subsamples and report the results in Table 6, Panel A for the piecewise linear model and Panel B for the continuous model. The flow-performance relation is nonlinear in each subsample. We find that the coefficients of the performance rank variables for net cash flows are the highest in the case of months with positive aggregate flows. In the case of the continuous model, the coefficient of RANK_SQUARE is positive and significant in all regressions. We conclude that the flow-performance relation remains nonlinear in all kinds of economic conditions.

[Table 7 here]

We provide further robustness tests on the dynamics of the flow-performance relation. Our sample period is from January 2000 to December 2013. This period includes the stock market peak during the dot-com bubble in early 2000, which was followed by the dot-com crash. The period also includes the credit crisis in 2007-08 and the subsequent recovery. It is possible that our aggregate results on the flow-performance relation are due to the influence of a small sub-period in the data. It is of interest to see the flow-performance relation during various sub-periods. We follow a simple approach. We run the flow-performance regressions separately for each year from 2000 to 2013. We use the full set of control variables in these regressions, but report only the coefficients of LOWPERF, MIDPERF and HIGHPERF in Table 7. The dependent variables are percentage net cashflows, new sales and redemptions in Panels A, B and C respectively. The coefficients of LOWPERF and HIGHPERF for net cashflows are significant and greater than the MIDPERF coefficient in each year. While the magnitude of the coefficients varies from year to year, a nonlinear relationship similar to that in the aggregate sample is observed in each individual year. In the case of new sales and redemptions, the results in each individual year are also very similar to the corresponding results in the aggregate sample. The flow-performance relation observed in the aggregate data is clearly a phenomenon that is consistent throughout the sample period.

[Table 8 here]

The flow-performance relation might differ across funds belonging to different investment categories. Although we have included the average flow to the category of the fund as a control variable in the regressions with aggregate data, a separate analysis of the flow-performance relation in each category may be interesting due to several reasons. Spiegel and Zhang (2013) argue that there might be some hot funds and some cold funds. They conjecture that the flow-performance relation may be linear in both subsets of funds, with the slope being lower in the case of cold funds. They claim that running a regression of percentage flows on performance might incorrectly reveal a nonlinear shape in the aggregate data, while the relation is linear in the two subsets of funds. We have shown earlier that the relation is nonlinear in the subsamples of hot funds and cold funds, where hot fund are defined as those that are less than five years old and are below median in the ranking of funds according to size. We now approach the issue of hot funds and cold funds from a different perspective. It can be argued that the funds that invest in small stocks and growth stocks are more likely to be hot funds, whereas the funds that invest in large stocks and value stocks are more likely to be cold funds. In order to operationalize our idea, we run separate flow-performance regressions for each category of funds. The results are reported in Table 8. We observe from Panel A that the relation between net cashflows and past performance for each category of funds is nonlinear in the same way as for the aggregate data. The flow-performance relation is concave in the low to mid performance range, whereas it is strongly convex in the mid to high performance range for all investment categories.

The results for new sales in Panel B and redemptions in Panel C also confirm that the flow-performance relation in individual categories is similar to that in the aggregate data in most of the regressions. These results reaffirm our previous results on hot and cold funds. While the shape of the flow-performance relation may vary somewhat across various categories of funds, the relation remains nonlinear in all categories of funds.

4. A new model: Share of new sales and redemptions

The fractional flow model of Sirri and Tufano (1998) and the market share model of Spiegel and Zhang (2013) are based on the notion that the flows to funds should be somehow related to the size of the fund. Consider two funds, one small and the other large. The fractional flow model divides dollar flows by fund size. This effectively means that the percentage flow of the smaller fund will be higher for the same level of dollar flows. This approach seems unsatisfactory. There is no compelling reason for the scaling of fund flows by fund size. The change in market share model implies that the market share of the two funds should not change if the two funds have equal performances. This effectively means that more dollars should flow to the larger fund in the case of no performance difference between the two funds. In reality, fund managers are more likely to care about the dollar flows rather than the fractional flows or the change in market share. The investors too, assuming that they are attracted by past performance, are more likely to invest their dollars into funds with better past performance and fund size is likely to be of only secondary importance to the investors.

We can think of an economy in which the aggregate new sales and aggregate redemptions are determined by the state of the economy. Given a fixed amount of aggregate new sales (or aggregate redemptions), the investors have to determine how to allocate their dollars to the funds. Assuming that investors are enticed by the past performance of funds, it seems reasonable to argue that they will invest more dollars into funds with better past performance, irrespective of the funds' sizes.

In this section, we propose a measure of flows that focuses on the dollar flows and is independent of the fund size. We define NEWSALES_SHARE for a fund in a month as the dollar new sales of the fund as a fraction of the total new sales of all the funds in the month. Similarly, REDEMPTIONS_SHARE is defined as the dollar redemptions as a fraction of total redemptions from all funds during the month. New sales and redemptions are non-negative amounts for all the funds. Therefore, NEWSALES_SHARE and REDEMPTIONS_SHARE are meaningful variables. Aggregate net cashflows in a month can be negative, zero or positive. Therefore, the share of NET_CASHFLOW of a fund is difficult to interpret. If the aggregate net cashflow in a month is 0, then it is not even possible to define the NET_CASHFLOW share of funds in that month. On the contrary, aggregate new sales and aggregate redemptions are always positive. Therefore, NEWSALES_SHARE and REDEMPTIONS_SHARE are always well-defined and have a nice interpretation. Still, aggregate new sales and (or) aggregate redemptions can be very low in some months and very high in some other months. We divide the 168 months of our sample into two groups according to whether the aggregate new sales in the month are below median or above median. Similarly we divide the 168 months into two groups according to whether the aggregate redemptions in the month are below median or above median. Finally, we divide the sample into four subsamples – months with low aggregate new sales and low aggregate redemptions (LL, 61 months), months with low aggregate new sales and high aggregate redemptions (LH, 23 months), months with high aggregate new sales and low aggregate redemptions (HL, 23 months), and months with high aggregate new sales and high aggregate redemptions (HH, 61 months). For each subsample, we run separate regressions of NEWSALES_SHARE and REDEMPTIONS_SHARE on performance rank variables and control variables. We include dummy variables for investment styles to control for any fund style effect on flows.

Table 9 reports the results for piecewise linear models in Panel A and continuous models in Panel B. The coefficients have been multiplied by 1,000 for better readability. We first focus on the piecewise linear models. Models (1)-(4) are the NEWSALES_SHARE regressions for the four subsamples. We learn the following from these regressions: (a) the dollar flows are sensitive to past performance; (b) the sensitivity is the lowest in the mid performance range; higher in the low performance range, and the highest in the high performance range for all sub-periods; and (c) a visual inspection of the magnitudes of the coefficients suggests that the shape of the flow-performance curve is similar in all sub-periods.

[Table 9 here]

Models (5)-(8) are the REDEMPTIONS_SHARE regressions for the four subsamples. We learn the following from these regressions: (a) investors react strongly to poor performance by withdrawing their money from funds with relatively poorer past performance in all sub-periods; (b) in the months with above median aggregate new sales, HL and HH, redemptions are sensitive to performance even in the highest performance range.

The continuous models in Panel B confirm the previous results. For example, the sensitivity of the share of new sales to performance in model (1) is -14.3+2*37.8*RANK, which is increasing in RANK and is positive for RANK above 0.19. Similarly, the sensitivity of the share of redemptions to RANK in model (5) is -24.6+2*15.8*RANK, which is increasing in RANK, but is decreasing in absolute value and is negative for RANK below 0.78. Unfortunately, as in the case of fractional flow regressions, the continuous model is not able to capture the dynamic nature of the flow-performance sensitivity as accurately as the piecewise linear model.

The results based on the sensitivity of NEWSALES_SHARE and REDEMPTIONS_SHARE to performance can be thought of as just another robustness check on our earlier flow-performance relation results. However, we believe that the 'flow shares' measure is an important concept by itself. The fractional flow measure used by earlier studies suffers from the criticism that it produces higher measures for smaller funds for the same dollar flows. Our 'flow share' measures, NEWSALES_SHARE and REDEMPTIONS_SHARE, directly measure the dollar flows to and from funds respectively, without being affected by the sizes of the funds, and therefore address a major criticism of the fractional flow model.

Spiegel and Zhang (2013) criticize the fractional flow model for its dependence on fund size. Their criticism goes as follows: "Consider an economy with two funds: one has $100 under management, the other, $10. The fractional flow model states that the better performing fund will see an inflow of 10%, the other, 0%. If the large fund does better, the aggregate flow equals $10, and if the small one does better, the aggregate flow is $1. This relation between fund returns across different size groups and aggregate flows is, in fact, a general implication of the standard fractional flow model: When large funds do relatively well, aggregate flows should be larger. However, our tests yield little evidence that this is the case. Aggregate flows are seemingly determined by economy-wide events, such as the overall market return, not by whether large or small funds have recently done relatively well." They propose the market share model to overcome this problem. However, their market share model is not without its own misspecification problems. In the same example, consider the scenario in which both funds have the same performance. According to Spiegel and Zhang (2013), the market share of funds should not change. Therefore, if the aggregate flows are $11, then the distribution of flows must be $1 to the smaller fund and $10 to the larger fund. There is no obvious reason why investors will buy more shares of the larger fund in the absence of any performance differential. Our measures of flows, NEWSALES_SHARE and REDEMPTIONS_SHARE, overcome the misspecification problems of both the fractional flow model and the market share model. They are independent of the fund size and directly capture the effect of performance on new sales and redemptions. Any effect of fund size on flows is captured by including fund size as a control variable in the regressions.

5. Conclusion

We study the relation between flows and past performance in the mutual fund industry. We use the actual dollar flows into and out of individual funds which is a significant improvement over the majority of previous studies that use only an estimate of net flows and have no information on inflows and outflows. We find that net flows are very sensitive to performance for funds in the high performance range as well as for those in the low performance range. New sales are very sensitive to performance in the low as well as high performance ranges, whereas redemptions are highly sensitive to performance in the low performance range.

We also consider the alternative methodology suggested by Spiegel and Zhang (2013) to estimate the flow-performance relation. We find that the flow-performance relation is similarly nonlinear even if we use the market share based measure of flows. We also the estimate flow-performance relation for the subsample of young and small funds (hot funds) and the subsample of the rest of the funds (cold funds). The results confirm that both hot funds and cold funds face a nonlinear sensitivity of flows to performance. The assertion of Spiegel and Zhang (2013) that the flow-performance relation is actually linear puts a question mark on the studies that explain managerial actions based on the incentives that might be generated by a nonlinear flow-performance relation. The results in this paper show that the flow-performance is nonlinear, even if we measure flows using the "change in market share" measure suggested by Spiegel and Zhang (2013). We conclude that incentives based explanations of the risk-taking behavior exhibited by managers cannot be discarded.

We perform several robustness tests on the fractional flow model. We find that the nonlinearity in the flow-performance relation holds every year separately. Naturally, it holds in periods with positive aggregate flows as well as in periods with negative aggregate flows. We also find that it holds for funds within each investment style. Our battery of tests confirms that the relation between flows and performance is nonlinear and that this relation is very robust.

Finally, we propose measures of fund flows that are independent of fund size. The share of a fund's new sales of the aggregate new sales and the share of a fund's redemptions of the aggregate redemptions during the month depend only on the dollar flows and not on the fund size. We find that the relation between flows and performance remains nonlinear with a shape similar to that in the case of fractional flows or changes in market share. We conclude that the flow-performance relation is nonlinear with a concave shape in the low to mid performance range and a convex shape in the mid to high performance range. This result is robust over time and across model specifications.

References

Basak, S., and D. Makarov, (2012). Difference in interim performance and risk taking with short-sale constraints. *Journal of Financial Economics* 103, 377–392.

Berggrun, L., and E. Lizarzaburu, (2015). Fund flows and performance in Brazil. *Journal of Business Research,* 68(2), 199-207.

Bergstresser, D. and J. Poterba, (2002). Do after-tax returns affect mutual fund inflows? *Journal of Financial Economics* 63, 381–414.

Berk, J.B. and R.C. Green, (2004). Mutual fund flows and performance in rational markets, *Journal of Political Economy* 112, 1269–1295.

Berk, J.B. and I. Tonks, (2007). Return persistence and fund flows in the worst performing mutual funds. *Working paper*, National Bureau of Economic Research.

Carhart, M. M. (1997). On persistence in mutual fund performance. *The Journal of Finance*, 52(1), 57-82.

Brown, K., W.V. Harlow, and L.T. Starks, (1996). Of tournaments and temptations: An analysis of managerial incentives in the mutual fund industry. *Journal of Finance* 51, 85–110.

Cashman, George D., Daniel N. Deli, Federico Nardari, and Sriram V. Villupuram, (2012). Investors do respond to poor mutual fund performance: Evidence from inflows and outflows, *Financial Review* 47, 719-739.

Chevalier, J. and G. Ellison, (1997). Risk taking by mutual funds as a response to incentives, *Journal of Political Economy* 105, 1167–1200.

Chordia, T., (1996). The structure of mutual fund charges, *Journal of Financial Economics* 41, 3–39.

Christoffersen, S., R. Evans, and D. Musto, (2013). What do consumers' fund flows maximize? Evidence from their brokers' incentives. *The Journal of Finance*, 68(1), 201-235.

Christoffersen, S., R. Evans, and D. Musto, (2007). Fund flows vs. family flows: Evidence from the cross section of brokers. *Working paper*, University of Pennsylvania.

Coval, J. and E. Stafford, (2007). Asset fire sales (and purchases) in equity markets, *Journal of Financial Economics* 86, 479–512.

Del Guercio, D. and P.A. Tkac, (2002). The determinants of the flow of funds of managed portfolios: Mutual funds vs. pension funds, *Journal of Financial and Quantitative Analysis* 37, 523–557.

Fant, L.F., and E.S. O'Neal, (2000). Temporal changes in the determinants of mutual fund flows. *Journal of Financial Research* 23, 353–371.

Ferreira, M.A., Keswani, A., Miguel, A.F. and Ramos, S.B. (2012). The flow–performance relationship around the world. *Journal of Banking & Finance,* 36(6), 1759-1780.

Greene, J. and C. Hodges, (2002). The dilution impact of daily fund flows on open-end mutual funds, *Journal of Financial Economics* 65, 131–158.

Gruber, M.J., (1996). Another puzzle: The growth in actively managed mutual funds. *Journal of Finance* 51, 783–810.

Huang, J., K.D. Wei, and H. Yan, (2007). Participation costs and the sensitivity of fund flows to past performance, *Journal of Finance* 62, 1273–1311.

Investment Company Institute, (2013). *Fact Book* (Investment Company Institute, Washington, DC).

Ivkovic, Z. and S. Weisbenner, (2009). Individual investor mutual fund flows, *Journal of Financial Economics* 92, 223–237.

Jaiprakash, P. and R.Kumar, (2009). A close look at the aggregate fund flows and market returns relationship. *Working paper*, Virginia Tech.

Johnson,W.T., (2004). Predictable investment horizons and wealth transfers among mutual fund shareholders, *Journal of Finance* 59, 1979–2012.

Johnson, W.T., (2007). Who monitors the mutual fund manager, new or old shareholders? *Unpublished working paper*, Securities and Exchange Commission.

Koijen, R. S. (2014). The Cross-Section of Managerial Ability, Incentives, and Risk Preferences. *The Journal of Finance*, 69(3), 1051-1098.

Lynch, A.W. and D.K. Musto, (2003). How investors interpret past fund returns, *Journal of Finance* 58, 2033–2058.

O'Neal, E., (2004). Purchase and redemption patterns of U.S. equity mutual funds, *Financial Management* 33.1, 63–90.

Sirri, E.R. and P. Tufano, (1998). Costly search and mutual fund flows, *Journal of Finance* 53, 1589–1622.

Spiegel, M. and H. Zhang, (2013). Mutual fund risk and market share-adjusted fund flows, *Journal of Financial Economics* 108, 506-528.

Taylor, J., (2003). Risk-taking behavior in mutual fund tournaments. *Journal of Economic Behavior and Organization* 50, 373–383.

Table 1: Summary Statistics

We report summary statistics for the mutual funds in our sample. The sample covers monthly observations from January 2000 to December 2013. The number of distinct mutual funds in the sample is 3791 and the total number of fund-month observations is 352875. Fund total net assets is the sum of the net assets of the different share classes of the same fund. Family size is the total assets under management of the other funds in the family that the fund belongs to, excluding the assets of the fund itself. Log of Family Total Net Assets is the logarithm of (one plus family size). Net expense ratio is the ratio of total investment that shareholders paid for the fund's operating expenses in the previous year. Turnover is the minimum of aggregated sales or aggregated purchases of securities, divided by the average 12-month total net assets of the fund in the previous year. Age is the number of years since the fund was first offered. Monthly Return is the net return of the fund in a month. Return Volatility for a fund in a month is calculated as the standard deviation of monthly returns over 12 months immediately preceding the current month.

Variable	Mean	Median	Std Dev	1st Pctl	99th Pctl
Flow (% of TNA)	0.45	-0.26	6.07	-13.69	23.83
Net Cash Flow (% of TNA)	0.45	-0.22	5.42	-12.28	22.22
New Sales (% of TNA)	3.28	1.62	6.29	0.00	30.40
Redemptions (% of TNA)	2.81	1.90	3.99	0.00	20.75
Total Net Assets ($ million)	1321	286	3597	12	19884
Family Size ($ million)	32075	7306	78755	0	431201
Family Total Size ($ million)	33514	8224	80566	15	441874
Number of Funds in the Family	30.79	24.00	30.01	1.00	128.00
Log Total Net Assets ($ million)	20.51	20.47	1.72	17.27	24.71
Log Family Size ($ million)	20.55	22.71	6.78	0.00	26.79
Gross Expense Ratio (% per year)	1.31	1.22	0.69	0.22	3.63
Net Expense Ratio (% per year)	1.20	1.15	0.56	0.18	2.83
Expense Waiver	0.14	0.04	0.26	0.00	1.37
Turnover (% per year)	80.81	61.00	76.47	2.00	400.00
Age (years)	13.81	10.59	12.62	1.25	70.42
Monthly Return	0.0071	0.0125	0.0548	-0.1535	0.1359
Return Volatility	0.0480	0.0426	0.0283	0.0166	0.1265

Table 2: The effect of relative performance on net flows of mutual fund

We estimate the sensitivity of net flows to the past performance of funds. The two measures of net flows are NET_CASHFLOW_PCT=100*(new sales – redemptions)/(lag fund size) and TNA_FLOW_PCT= (fund size / lag fund size)-(1+return). We define a fund's performance rank variable RANK as the fractional rank of the fund in the cross-section of funds of same style in a given month on a scale of 0 to 1. We define the following variables for our regressions: LOWPERF = Min(0.33, RANK), MIDPERF = Min(0.33, RANK-LOWPERF) and HIGHPERF = RANK-(LOWPERF+MIDPERF). We regress the two measures of monthly percentage fund flows on performance rank variables and other fund characteristics. We use Fama-Macbeth regressions – cross-sectional regressions are run every month. The standard errors are adjusted for serial correlation using Newey-West lags of order three. The sample consists of monthly observations from January 2000 to December 2013, a total of 168 months. P-values are reported in parentheses.

	(1) Net	(2) Net	(3) flow_pct	(4) flow_pct
LOWPERF	2.525		2.494	
	(0.00)		(0.00)	
MIDPERF	0.449		0.939	
	(0.00)		(0.00)	
HIGHPERF	4.835		5.262	
	(0.00)		(0.00)	
RANK		0.161		0.139
		(0.37)		(0.57)
RANK_SQUARE		1.969		2.334
		(0.00)		(0.00)
CAT_DEPVAR	57.781	57.722	64.992	64.657
	(0.00)	(0.00)	(0.00)	(0.00)
LAG_DEPVAR	0.400	0.402	0.269	0.270
	(0.00)	(0.00)	(0.00)	(0.00)
LOGTNA	-0.087	-0.086	-0.110	-0.108
	(0.00)	(0.00)	(0.00)	(0.00)
LOGFAMSIZE	0.001	0.001	0.003	0.003
	(0.67)	(0.69)	(0.30)	(0.27)
TURNOVER	-0.001	-0.001	-0.001	-0.001
	(0.01)	(0.01)	(0.03)	(0.03)
EXPRATIO	-0.161	-0.159	-0.106	-0.105
	(0.00)	(0.00)	(0.14)	(0.14)
AGE	-0.016	-0.016	-0.018	-0.018
	(0.00)	(0.00)	(0.00)	(0.00)
VOLATILITY	-15.738	-15.903	-12.924	-13.168
	(0.00)	(0.00)	(0.00)	(0.00)
CONSTANT	1.981	2.286	2.206	2.486
	(0.00)	(0.00)	(0.00)	(0.00)
Observations	208632	208632	229726	229726
R^2	0.23	0.23	0.17	0.16

Table 3: The effect of relative performance on new sales and redemptions

We estimate the sensitivity of new sales and redemptions to the past performance of funds. We define NEWSALES_PCT= 100*(new sales)/(lag fund size), REDEMPTIONS_PCT= 100*(redemptions)/(lag fund size), and NET_CASHFLOW_PCT=100*(new sales – redemptions)/(lag fund size). We regress the monthly flow measures on performance rank variables and other fund characteristics. In columns (1)-(3), we estimate a piecewise linear model with LOWPERF, MIDPERF, and HIGHPERF as the performance variables. In columns (4)-(6), we estimate a continuous model with RANK and RANK_SQUARE as the performance variables. We use Fama-Macbeth regressions – cross-sectional regressions are run every month. The standard errors are adjusted for serial correlation using Newey-West lags of order three. The sample consists of monthly observations from January 2000 to December 2013, a total of 168 months. P-values are reported in parentheses.

	(1) New	(2) Red	(3) Net	(4) New	(5) Red	(6) Net
LOWPERF	1.490	-1.557	3.019			
	(0.00)	(0.00)	(0.00)			
MIDPERF	0.104	-0.427	0.472			
	(0.49)	(0.00)	(0.00)			
HIGHPERF	5.215	-0.228	5.550			
	(0.00)	(0.06)	(0.00)			
RANK				-1.275	-1.709	0.304
				(0.00)	(0.00)	(0.19)
RANK_SQUARE				3.065	1.051	2.133
				(0.00)	(0.00)	(0.00)
CAT_DEPVAR	48.623	31.560	57.795	48.388	31.745	57.371
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
LAG_NEWSALES_PCT	0.449	0.048	0.381	0.451	0.047	0.383
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
LAG_REDEMPTIONS_PCT	0.091	0.378	-0.298	0.089	0.378	-0.300
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
LOGTNA	-0.139	-0.008	-0.123	-0.137	-0.009	-0.121
	(0.00)	(0.34)	(0.00)	(0.00)	(0.32)	(0.00)
LOGFAMSIZE	0.018	0.017	-0.000	0.018	0.017	-0.000
	(0.00)	(0.00)	(0.97)	(0.00)	(0.00)	(1.00)
TURNOVER	0.002	0.002	-0.000	0.002	0.002	-0.000
	(0.00)	(0.00)	(0.18)	(0.00)	(0.00)	(0.19)
EXPRATIO	0.030	0.243	-0.198	0.034	0.241	-0.193
	(0.52)	(0.00)	(0.00)	(0.47)	(0.00)	(0.00)
AGE	-0.021	-0.003	-0.018	-0.021	-0.003	-0.019
	(0.00)	(0.02)	(0.00)	(0.00)	(0.02)	(0.00)
VOLATILITY	-0.694	20.625	-19.107	-0.521	20.624	-18.990
	(0.83)	(0.00)	(0.00)	(0.87)	(0.00)	(0.00)
CONSTANT	1.879	-0.230	2.314	2.196	-0.240	2.649
	(0.00)	(0.19)	(0.00)	(0.00)	(0.18)	(0.00)
Observations	207876	207876	207876	207876	207876	207876
R^2	0.31	0.28	0.23	0.31	0.28	0.23

Table 4: The effect of relative performance on change in market share (dms)

We estimate the sensitivity of flows to the past performance of funds using the market share model of Spiegel and Zhang (2013). We calculate the change in market share as explained in the text. We regress the change in market share on performance rank variables and other fund characteristics. Columns (1)-(4) report the outputs from piecewise linear models with LOWPERF, MIDPERF, and HIGHPERF as the performance variables. Columns (5)-(8) report the outputs from continuous models with RANK and RANK_SQUARE as the performance variables. We use Fama-Macbeth regressions – cross-sectional regressions are run every month. The standard errors are adjusted for serial correlation using Newey-West lags of order three. The sample consists of monthly observations from January 2000 to December 2013, a total of 168 months. P-values are reported in parentheses.

	(1) dms	(2) dms	(3) dms	(4) dms	(5) dms	(6) dms	(7) dms	(8) dms
LOWPERF	6.830	6.910	3.148	3.379				
	(0.00)	(0.00)	(0.00)	(0.00)				
MIDPERF	4.849	4.931	2.433	2.663				
	(0.00)	(0.00)	(0.00)	(0.00)				
HIGHPERF	10.552	10.819	6.100	6.955				
	(0.00)	(0.00)	(0.00)	(0.00)				
RANK					3.702	3.674	1.171	1.043
					(0.00)	(0.00)	(0.02)	(0.04)
RANK_SQUARE					3.035	3.190	2.330	2.839
					(0.00)	(0.00)	(0.00)	(0.00)
CAT_DMS		1.170		0.853		1.169		0.853
		(0.00)		(0.00)		(0.00)		(0.00)
LAG_DMS			0.347	0.341			0.347	0.341
			(0.00)	(0.00)			(0.00)	(0.00)
LOGTNA			-0.209	-0.181			-0.207	-0.179
			(0.00)	(0.00)			(0.00)	(0.00)
LOGFAMSIZE			0.004	-0.003			0.004	-0.003
			(0.32)	(0.43)			(0.30)	(0.46)
TURNOVER			-0.002	-0.001			-0.002	-0.001
			(0.01)	(0.01)			(0.01)	(0.01)
EXPRATIO			-0.111	-0.336			-0.112	-0.338
			(0.25)	(0.00)			(0.25)	(0.00)
AGE			-0.054	-0.049			-0.054	-0.049
			(0.00)	(0.00)			(0.00)	(0.00)
VOLATILITY			-19.008	-24.439			-19.238	-24.680
			(0.15)	(0.01)			(0.15)	(0.01)
CONSTANT	-3.438	-3.750	3.976	3.837	-3.005	-3.303	4.210	4.111
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Observations	241150	241150	190974	190974	241150	241150	190974	190974
R^2	0.03	0.09	0.22	0.25	0.03	0.08	0.22	0.25

Table 5: The flow-performance relation for Hot and Cold Funds

We estimate the sensitivity of new sales, redemptions, and net cashflows to the past performance of funds for hot funds and cold funds separately. Hot funds are defined as those that are less than five years old and are below median in a ranking of funds according to their size within each month. Cold funds are the funds that are not hot, i.e. old or large funds. In Panel A, we estimate a piecewise linear model with LOWPERF, MIDPERF, and HIGHPERF as the performance variables. In Panel B, we estimate a continuous model with RANK and RANK_SQUARE as the performance variables. Columns (1)-(3) are for cold funds and columns (4)-(6) are for hot funds. We use Fama-Macbeth regressions – cross-sectional regressions are run every month. The standard errors are adjusted for serial correlation using Newey-West lags of order three. The sample consists of monthly observations from January 2000 to December 2013, a total of 168 months. P-values are reported in parentheses.

Panel A: Piecewise linear model

	(1) New	(2) Red	(3) Net	(4) New	(5) Red	(6) Net
LOWPERF	1.346	-1.483	2.856	2.230	-1.535	3.658
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
MIDPERF	0.155	-0.439	0.514	0.381	-0.589	0.971
	(0.32)	(0.00)	(0.00)	(0.50)	(0.03)	(0.04)
HIGHPERF	4.670	-0.199	5.013	8.540	0.685	7.558
	(0.00)	(0.16)	(0.00)	(0.00)	(0.02)	(0.00)
CAT_DEPVAR	41.894	29.567	52.333	109.612	39.314	113.462
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
LAG_NEWSALES_PCT	0.446	0.049	0.373	0.422	0.038	0.369
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
LAG_REDEMPTIONS_PCT	0.087	0.385	-0.309	0.123	0.355	-0.219
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
LOGLAGFUNDSIZE	-0.089	-0.022	-0.060	-0.411	0.096	-0.455
	(0.00)	(0.03)	(0.00)	(0.00)	(0.01)	(0.00)
LOGLAGFAMSIZE	0.017	0.018	-0.001	0.026	0.021	0.004
	(0.00)	(0.00)	(0.66)	(0.00)	(0.00)	(0.50)
LAG_TURNOVER	0.001	0.002	-0.001	0.003	0.003	-0.001
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.37)
LAG_NETEXPRATIO	0.077	0.256	-0.174	-0.209	0.323	-0.477
	(0.05)	(0.00)	(0.00)	(0.09)	(0.00)	(0.00)
AGE	-0.018	-0.003	-0.015	-0.423	0.133	-0.533
	(0.00)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)
RETVOL	-2.144	19.173	-19.439	5.980	17.282	-9.137
	(0.47)	(0.00)	(0.00)	(0.23)	(0.00)	(0.05)
CONSTANT	1.070	0.059	1.105	6.276	-2.946	9.832
	(0.00)	(0.77)	(0.00)	(0.00)	(0.00)	(0.00)
Observations	189321	189321	189321	29349	29349	29349
R^2	0.31	0.28	0.23	0.37	0.33	0.32

Panel B: Continuous model

	(1) New	(2) Red	(3) Net	(4) New	(5) Red	(6) Net
RANK	-1.109	-1.669	0.444	-2.150	-2.195	0.128
	(0.00)	(0.00)	(0.04)	(0.00)	(0.00)	(0.85)
RANK_SQUARE	2.745	1.032	1.832	5.187	1.718	3.283
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
CAT_DEPVAR	41.674	29.717	51.999	109.188	39.411	113.451
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
LAG_NEWSALES_PCT	0.448	0.049	0.375	0.424	0.038	0.372
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
LAG_REDEMPTIONS_PCT	0.085	0.385	-0.310	0.121	0.354	-0.220
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
LOGLAGFUNDSIZE	-0.088	-0.022	-0.058	-0.404	0.095	-0.447
	(0.00)	(0.03)	(0.00)	(0.00)	(0.01)	(0.00)
LOGLAGFAMSIZE	0.017	0.018	-0.001	0.026	0.021	0.004
	(0.00)	(0.00)	(0.66)	(0.00)	(0.00)	(0.46)
LAG_TURNOVER	0.001	0.002	-0.001	0.003	0.003	-0.001
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.35)
LAG_NETEXPRATIO	0.081	0.254	-0.170	-0.232	0.323	-0.497
	(0.04)	(0.00)	(0.00)	(0.06)	(0.00)	(0.00)
AGE	-0.018	-0.003	-0.015	-0.430	0.130	-0.537
	(0.00)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)
RETVOL	-1.995	19.128	-19.326	5.704	17.256	-9.392
	(0.49)	(0.00)	(0.00)	(0.26)	(0.00)	(0.04)
CONSTANT	1.346	0.053	1.405	6.723	-2.875	10.183
	(0.00)	(0.80)	(0.00)	(0.00)	(0.00)	(0.00)
Observations	189321	189321	189321	29349	29349	29349
R^2	0.31	0.28	0.23	0.37	0.32	0.31

Table 6: The flow-performance relation for months with positive (negative) aggregate flows

We estimate the sensitivity of new sales, redemptions, and net cashflows to the past performance of funds for periods of low aggregate net cash flows (columns (1)-(3)), medium aggregate net cash flows (columns (4)-(6)), and high aggregate net cash flows (columns (7)-(9)). In Panel A, we estimate a piecewise linear model with LOWPERF, MIDPERF, and HIGHPERF as the performance variables. In Panel B, we estimate a continuous model with RANK and RANK_SQUARE as the performance variables. We use Fama-Macbeth regressions – cross-sectional regressions are run every month. The standard errors are adjusted for serial correlation using Newey-West lags of order three. The sample consists of monthly observations from January 2000 to December 2013, a total of 168 months. P-values are reported in parentheses.

Panel A: Piecewise linear model

	(1) New	(2) Red	(3) Net	(4) New	(5) Red	(6) Net	(7) New	(8) Red	(9) Net
LOWPERF	0.987	-1.704	2.755	1.758	-1.521	3.268	1.975	-1.405	3.296
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
MIDPERF	0.139	-0.486	0.566	-0.145	-0.412	0.228	0.524	-0.399	0.838
	(0.51)	(0.00)	(0.01)	(0.33)	(0.00)	(0.17)	(0.30)	(0.02)	(0.07)
HIGHPERF	4.379	0.222	4.168	5.807	-0.086	5.914	6.173	-0.364	6.733
	(0.00)	(0.31)	(0.00)	(0.00)	(0.71)	(0.00)	(0.00)	(0.12)	(0.00)
CAT_DEPVAR	63.219	44.224	67.056	50.347	28.988	63.670	39.247	23.639	54.759
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
LAG_NEWSALES_PCT	0.422	0.039	0.355	0.432	0.040	0.372	0.476	0.053	0.404
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
LAG_REDEMPTIONS_PCT	0.054	0.357	-0.317	0.126	0.405	-0.292	0.111	0.393	-0.279
	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)	(0.00)
LOGLAGFUNDSIZE	-0.086	0.028	-0.114	-0.188	-0.002	-0.173	-0.263	-0.042	-0.195
	(0.00)	(0.09)	(0.00)	(0.00)	(0.82)	(0.00)	(0.00)	(0.02)	(0.00)
LOGLAGFAMSIZE	0.015	0.019	-0.003	0.020	0.015	0.006	0.031	0.021	0.005
	(0.00)	(0.00)	(0.46)	(0.00)	(0.00)	(0.21)	(0.00)	(0.00)	(0.39)
LAG_TURNOVER	0.002	0.003	-0.001	0.002	0.003	-0.001	0.002	0.002	0.000
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)	(0.00)	(0.48)
LAG_NETEXPRATIO	-0.019	0.382	-0.394	-0.082	0.237	-0.247	0.078	0.203	-0.118
	(0.74)	(0.00)	(0.00)	(0.27)	(0.00)	(0.00)	(0.54)	(0.00)	(0.23)
AGE	-0.022	-0.008	-0.015	-0.021	-0.001	-0.021	-0.021	0.003	-0.025
	(0.00)	(0.00)	(0.00)	(0.00)	(0.74)	(0.00)	(0.00)	(0.34)	(0.00)
RETVOL	-6.059	17.789	-21.020	-1.145	17.881	-16.374	5.521	19.437	-14.036
	(0.19)	(0.00)	(0.00)	(0.75)	(0.00)	(0.00)	(0.44)	(0.00)	(0.02)
CONSTANT	1.176	-1.160	2.848	2.723	-0.285	3.124	3.770	0.312	3.202
	(0.01)	(0.00)	(0.00)	(0.00)	(0.23)	(0.00)	(0.00)	(0.25)	(0.00)
Observations	94430	94430	94430	75838	75838	75838	48402	48402	48402
R^2	0.26	0.20	0.20	0.32	0.27	0.25	0.37	0.35	0.28

Panel B: Continuous model

	(1) New	(2) Red	(3) Net	(4) New	(5) Red	(6) Net	(7) New	(8) Red	(9) Net
RANK	-1.288	-2.118	0.846	-1.459	-1.728	0.204	-1.147	-1.515	0.133
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.55)	(0.04)	(0.00)	(0.79)
RANK_SQUARE	2.760	1.514	1.220	3.350	1.122	2.275	3.527	0.873	2.876
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
CAT_DEPVAR	63.050	44.135	66.539	50.187	29.202	63.343	39.037	23.919	54.658
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
LAG_NEWSALES_PCT	0.423	0.038	0.357	0.434	0.040	0.374	0.477	0.052	0.405
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
LAG_REDEMPTIONS_PCT	0.053	0.358	-0.318	0.124	0.405	-0.293	0.110	0.393	-0.280
	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)	(0.00)
LOGLAGFUNDSIZE	-0.085	0.028	-0.113	-0.186	-0.002	-0.170	-0.259	-0.041	-0.192
	(0.00)	(0.09)	(0.00)	(0.00)	(0.80)	(0.00)	(0.00)	(0.02)	(0.00)
LOGLAGFAMSIZE	0.015	0.019	-0.003	0.021	0.016	0.006	0.030	0.021	0.005
	(0.00)	(0.00)	(0.50)	(0.00)	(0.00)	(0.18)	(0.00)	(0.00)	(0.41)
LAG_TURNOVER	0.002	0.003	-0.001	0.002	0.003	-0.001	0.002	0.002	0.000
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)	(0.00)	(0.41)
LAG_NETEXPRATIO	-0.017	0.381	-0.393	-0.078	0.236	-0.243	0.076	0.200	-0.117
	(0.76)	(0.00)	(0.00)	(0.29)	(0.00)	(0.00)	(0.55)	(0.00)	(0.24)
AGE	-0.022	-0.008	-0.015	-0.021	-0.001	-0.022	-0.021	0.003	-0.025
	(0.00)	(0.00)	(0.00)	(0.00)	(0.74)	(0.00)	(0.00)	(0.35)	(0.00)
RETVOL	-6.068	17.923	-21.237	-0.865	17.863	-16.099	5.900	19.177	-13.465
	(0.18)	(0.00)	(0.00)	(0.80)	(0.00)	(0.00)	(0.40)	(0.00)	(0.02)
CONSTANT	1.448	-1.158	3.133	3.086	-0.289	3.494	4.086	0.299	3.539
	(0.00)	(0.00)	(0.00)	(0.00)	(0.23)	(0.00)	(0.00)	(0.28)	(0.00)
Observations	94430	94430	94430	75838	75838	75838	48402	48402	48402
R^2	0.26	0.20	0.19	0.32	0.27	0.25	0.37	0.35	0.28

Table 7: Year-wise coefficients for the flow-performance relation

We estimate the sensitivity of net cashflows (Panel A), new sales (Panel B), and redemptions (Panel C) to the past performance of funds for each year from 2000 to 2013. Columns (1) to (14) represent years 2000 to 2013 in the same order. Regressions are run with the full set of control variables as in Table 3. For the sake of brevity, only the coefficients of LOWPERF, MIDPERF, and HIGHPERF are reported. We use Fama-Macbeth regressions – cross-sectional regressions are run every month. The standard errors are adjusted for serial correlation using Newey-West lags of order three. P-values are reported in parentheses.

Panel A: Net Cash Flow

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
LOWPERF	4.345	3.923	2.982	2.253	2.570	3.476	3.502	3.908	3.873	2.910	2.296	2.499	2.759	2.754
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
MIDPERF	0.680	-0.600	1.460	2.139	0.712	1.042	0.058	0.238	0.493	0.127	0.225	1.024	0.334	-0.025
	(0.07)	(0.15)	(0.08)	(0.01)	(0.07)	(0.04)	(0.91)	(0.45)	(0.20)	(0.67)	(0.20)	(0.22)	(0.22)	(0.95)
HIGHPERF	5.289	10.533	5.979	3.747	4.267	6.528	5.443	6.083	6.000	5.185	4.113	3.680	4.401	5.461
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)	(0.00)

Panel B: New Sales

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
LOWPERF	2.367	2.593	1.716	0.902	1.989	1.948	1.665	2.334	0.544	0.386	1.605	1.305	1.176	1.519
	(0.00)	(0.01)	(0.13)	(0.23)	(0.00)	(0.04)	(0.02)	(0.05)	(0.28)	(0.28)	(0.00)	(0.00)	(0.00)	(0.00)
MIDPERF	-0.001	-0.859	0.977	1.601	0.249	0.414	-0.546	-0.325	0.634	0.007	-0.324	0.833	-0.225	-0.294
	(1.00)	(0.06)	(0.43)	(0.05)	(0.61)	(0.41)	(0.19)	(0.35)	(0.14)	(0.98)	(0.00)	(0.21)	(0.26)	(0.36)
HIGHPERF	3.008	9.134	5.467	4.048	3.975	6.874	5.795	6.229	4.872	5.300	4.814	4.002	4.691	5.322
	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)

Panel C: Redemptions

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
LOWPERF	-1.854	-1.369	-2.065	-1.075	-0.541	-1.630	-1.763	-1.837	-2.765	-2.379	-0.595	-1.070	-1.584	-1.335
	(0.01)	(0.04)	(0.01)	(0.18)	(0.10)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.06)	(0.14)	(0.00)	(0.00)
MIDPERF	-0.667	-0.241	-0.773	-0.746	-0.497	-0.666	-0.525	-0.367	-0.317	-0.091	-0.612	-0.397	-0.501	-0.084
	(0.12)	(0.07)	(0.07)	(0.10)	(0.01)	(0.01)	(0.00)	(0.22)	(0.23)	(0.68)	(0.00)	(0.25)	(0.19)	(0.64)
HIGHPERF	-2.027	-0.634	-0.030	0.257	-0.128	0.149	0.121	-0.143	-0.488	0.079	0.683	0.092	0.135	-0.279
	(0.00)	(0.04)	(0.97)	(0.69)	(0.64)	(0.55)	(0.41)	(0.71)	(0.12)	(0.66)	(0.16)	(0.67)	(0.63)	(0.51)

Table 8: BY category

We estimate the sensitivity of net cashflows (Panel A), new sales (Panel B), and redemptions (Panel C) to the past performance of funds for each investment style. Columns (1) to (9) represent the investment styles 'Large Blend', 'Large Growth', 'Large Value', 'Mid Blend', 'Mid Growth', 'Mid Value', 'Small Blend', ‘Small Growth’ and ‘Small Value’ respectively. Regressions are run with the full set of control variables as in Table 3. For the sake of brevity, only the coefficients of LOWPERF, MIDPERF, and HIGHPERF are reported. We use Fama-Macbeth regressions – cross-sectional regressions are run every month. The standard errors are adjusted for serial correlation using Newey-West lags of order three. P-values are reported in parentheses.

Panel A: Net Cash Flows

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
LOWPERF	2.599	2.834	2.128	1.586	3.348	4.216	3.155	3.834	1.925
	(0.00)	(0.00)	(0.00)	(0.09)	(0.00)	(0.00)	(0.00)	(0.00)	(0.02)
MIDPERF	0.478	0.602	0.903	0.439	0.192	0.283	1.021	-0.330	1.874
	(0.11)	(0.02)	(0.02)	(0.45)	(0.65)	(0.62)	(0.08)	(0.46)	(0.01)
HIGHPERF	5.147	4.436	4.560	6.685	5.110	6.644	5.594	6.150	4.464
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)

Panel B: New Sales

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
LOWPERF	2.150	1.145	0.678	-0.421	2.406	2.901	1.760	1.755	0.627
	(0.00)	(0.00)	(0.10)	(0.57)	(0.00)	(0.00)	(0.00)	(0.02)	(0.31)
MIDPERF	-0.159	0.053	0.653	0.432	0.058	0.485	0.126	-0.689	1.783
	(0.64)	(0.89)	(0.09)	(0.39)	(0.90)	(0.38)	(0.80)	(0.12)	(0.02)
HIGHPERF	4.332	4.333	4.766	6.855	4.832	6.334	6.391	5.860	3.095
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)

Panel C: Redemptions

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
LOWPERF	-0.504	-1.610	-1.597	-1.840	-1.156	-1.151	-1.544	-2.658	-1.450
	(0.14)	(0.00)	(0.00)	(0.00)	(0.00)	(0.11)	(0.00)	(0.00)	(0.01)
MIDPERF	-0.772	-0.599	-0.377	0.095	-0.160	0.124	-0.616	-0.377	-0.235
	(0.00)	(0.00)	(0.10)	(0.78)	(0.57)	(0.73)	(0.04)	(0.18)	(0.55)
HIGHPERF	-0.494	-0.065	0.352	-0.001	-0.258	-0.380	0.205	-0.663	-1.291
	(0.13)	(0.76)	(0.19)	(1.00)	(0.42)	(0.36)	(0.58)	(0.10)	(0.05)

Table 9: Flow share

We estimate the sensitivity of new sales and redemptions to the past performance of funds using the 'share of flow' method. We define NEWSALES_SHARE for a fund in a month as the dollar new sales of the fund as a fraction of the total dollar new sales of all of the funds in the month. Similarly, REDEMPTIONS_SHARE is defined as the dollar redemptions from the fund as a fraction of the total dollar redemptions from all of the funds during the month. We divide the sample into four subsamples – months with low aggregate new sales and low aggregate redemptions (LL, 61 months), months with low aggregate new sales and high aggregate redemptions (LH, 23 months), months with high aggregate new sales and low aggregate redemptions (HL, 23 months), and months with high aggregate new sales and high aggregate redemptions (HH, 61 months). For each subsample, we run separate regressions of NEWSALES_SHARE and REDEMPTIONS_SHARE on performance rank variables and control variables. In Panel A, we estimate a piecewise linear model with LOWPERF, MIDPERF, and HIGHPERF as the performance variables. In Panel B, we estimate a continuous model with RANK and RANK_SQUARE as the performance variables. We include dummy variables for investment styles. The coefficients have been multiplied by 1000. The dependent variable is NEWSALES_SHARE in models (1)-(4) and REDEMPTIONS_SHARE in columns (5)-(8). We use Fama-Macbeth regressions – cross-sectional regressions are run every month. The sample consists of monthly observations from January 2000 to December 2013, a total of 168 months. P-values are reported in parentheses.

Panel A: Piecewise linear model

	NEWSALES_SHARE				REDEMPTIONS_SHARE			
	(1) LL	(2) LH	(3) HL	(4) HH	(5) LL	(6) LH	(7) HL	(8) HH
LOWPERF	16.074	2.728	9.950	12.344	-19.521	-31.876	-20.610	-24.761
	(0.00)	(0.28)	(0.03)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
MIDPERF	7.782	3.887	12.526	5.694	-7.931	-6.341	-6.709	-9.657
	(0.00)	(0.09)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)
HIGHPERF	61.516	55.918	59.053	47.539	0.488	1.184	-11.014	-6.998
	(0.00)	(0.00)	(0.00)	(0.00)	(0.80)	(0.60)	(0.01)	(0.01)
LOGLAGFUNDSIZE	20.702	19.241	20.452	18.821	21.642	20.328	21.004	19.741
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
LOGLAGFAMSIZE	-0.313	-0.252	-0.315	-0.274	-0.220	-0.262	-0.191	-0.263
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
LAG_TURNOVER	-0.006	-0.003	-0.002	0.006	0.001	-0.005	0.005	0.004
	(0.00)	(0.19)	(0.54)	(0.02)	(0.59)	(0.06)	(0.06)	(0.02)
LAG_NETEXPRATIO	2.029	5.218	0.956	2.825	4.316	6.327	3.039	4.293
	(0.00)	(0.00)	(0.22)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
AGE	-0.298	-0.316	-0.292	-0.279	0.113	0.048	0.109	0.106
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)	(0.00)
RETVOL	19.664	-181.643	-72.873	-69.680	176.187	100.470	223.864	168.900
	(0.64)	(0.02)	(0.33)	(0.09)	(0.00)	(0.02)	(0.00)	(0.00)
CONSTANT	-364.080	-332.941	-355.510	-333.083	-389.906	-361.252	-379.907	-353.792
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Observations	61133	30224	21436	70131	61133	30224	21436	70131
R^2	0.39	0.38	0.39	0.38	0.49	0.48	0.49	0.47

Panel B: Continuous model

	(1) newk LL	(2) newk LH	(3) newk HL	(4) newk HH	(5) redk LL	(6) redk LH	(7) redk HL	(8) redk HH
RANK	-14.305	-25.690	-15.265	-10.286	-24.618	-36.543	-18.488	-26.789
	(0.00)	(0.00)	(0.02)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
RANK_SQUARE	37.792	42.244	38.876	28.293	15.823	25.601	7.095	13.885
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.05)	(0.00)
LOGLAGFUNDSIZE	20.715	19.257	20.456	18.825	21.653	20.329	20.998	19.732
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
LOGLAGFAMSIZE	-0.313	-0.252	-0.315	-0.270	-0.221	-0.259	-0.187	-0.261
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
LAG_TURNOVER	-0.006	-0.004	-0.002	0.006	0.001	-0.005	0.005	0.004
	(0.01)	(0.10)	(0.53)	(0.02)	(0.65)	(0.06)	(0.07)	(0.02)
LAG_NETEXPRATIO	1.989	5.235	0.945	2.813	4.327	6.295	3.041	4.262
	(0.00)	(0.00)	(0.23)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
AGE	-0.299	-0.317	-0.291	-0.280	0.112	0.048	0.110	0.107
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)	(0.00)
RETVOL	11.394	-193.956	-63.452	-68.258	173.259	102.625	222.651	167.230
	(0.78)	(0.01)	(0.38)	(0.09)	(0.00)	(0.02)	(0.00)	(0.00)
CONSTANT	-368.679	-334.651	-363.437	-336.558	-389.539	-362.535	-379.979	-355.781
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Observations	61133	30224	21436	70131	61133	30224	21436	70131
R^2	0.39	0.38	0.39	0.38	0.49	0.48	0.49	0.47

Appendix A

Spiegel and Zhang (2013) define their 'change in market share' measure for fund flows as follows:

$$\Delta m_{it} = \frac{n_{it}}{\widehat{N}_t} - \frac{n_{i,t-1}}{N_{t-1}} \tag{A.1}$$

The formula for fractional flows is:

$$\frac{f_{i,t}}{n_{i,t-1}} = \left(\frac{n_{i,t}}{n_{i,t-1}}\right) - \left(1 + r_{i,t}\right) \tag{A.2}$$

We can rearrange the fractional flows formula as follows:

$$n_{i,t} = n_{i,t-1} * \left(1 + r_{i,t}\right) + f_{i,t} \tag{A.3}$$

Substituting $n_{i,t}$ from equation (A.3) in equation (A.1) and rearranging the terms, we can express fractional flows as a function of 'change in market share':

$$\frac{f_{i,t}}{n_{i,t-1}} = \left(\frac{\widehat{N}_t}{N_{t-1}}\right) - (1 + r_{i,t}) + \left(\frac{\widehat{N}_t}{n_{i,t-1}}\right)\Delta m_{i,t} \tag{A.4}$$

Or equivalently, we can express 'change in market share' as a function of fractional flows:

$$\Delta m_{i,t} = \left(\frac{n_{i,t-1}}{\widehat{N}_t}\right)\left[\frac{f_{i,t}}{n_{i,t-1}} + r_{i,t} - \left(\frac{\widehat{N}_t - N_{t-1}}{N_{t-1}}\right)\right] \tag{A.5}$$

In the square bracket in the right hand side of equation (A.5), there are three terms. The first term is simply the fractional flow which is known to be a nonlinear function of performance rank. The second term is current month return of the fund. The third term is negative of the fractional change of the aggregate assets of all funds, which is a constant across funds and therefore not related to the past performance rank of funds. Therefore, the shape of the relation between the term in the square bracket and past performance rank depends mainly on the first term, i.e. fractional flows.

In columns (1) and (2) of Appendix Table 1, we regress 'change in market share' calculated using equations (A.1) and (A.5) respectively. We note that the results in columns (1) and (2) are almost identical. Next, we use the alternative expression for 'change in market share' in the right hand side of equation (A.5) and regress its various components on performance rank variables and other fund characteristics. The results are reported in columns (3)-(6). The dependent variable in column (3) is the first term in square brackets which is simply the fractional flows of funds. The coefficients of LOWPERF, MIDPERF and HIGHPERF are close to zero in column (5) and quite small in column (4) compared to those in column (3) for fractional flows. In other words, it is the first term inside the square brackets in equation (A.5), i.e. fractional flows, that dictates the shape of the relation between 'change in market share' and past performance. The term in the first parentheses in the RHS of equation (A.5) is too small to significantly alter the effect of fractional flows on 'change in market share'. As a result, the relation between performance and 'change in market share' in column (1) is very similar to the relation between performance and fractional flows in column (3). These results are also confirmed by the results of regressions based on a continuous model, reported in Panel B. We conclude that the two measures of flows, viz. fractional flows and ‘change in market share’, are closely related and the flow-performance relation is nonlinear using both these measures of fund flows.

Appendix Table 1

We estimate the sensitivity of flows to past performance of funds using the market share model of Spiegel and Zhang (2013). In columns (1) and (2), we regress 'change in market share' calculated using equations (A.1) and (A.5) respectively. We use the alternative expression of 'change in market share' in the right hand side of equation (A.5) and regress its various components on performance rank variables and other fund characteristics in columns (3)-(6). Panel A reports output from piecewise linear models with LOWPERF, MIDPERF, and HIGHPERF as the performance variables. Panel B reports output from continuous models with RANK and RANK_SQUARE as the performance variables. We use Fama-Macbeth regressions – cross-sectional regressions are run every month. The standard errors are adjusted for serial correlation using Newey-West lags of order three. The sample is monthly observations from January 2000 to December 2013, a total of 168 months. P-values are reported in parentheses.

Panel A: Piecewise Linear Model

	(1) dms $\frac{n_{i,t}}{N_t} - \frac{n_{i,t-1}}{N_{t-1}}$	(2) dmsnew (4)* [(1)+(2)+(3)]	(3) dmspart2a $\frac{f_{i,t}}{n_{i,t-1}}$	(4) dmspart2b $r_{i,t}$	(5) dmspart2c $\left(\frac{N_t - N_{t-1}}{N_{t-1}}\right)$	(6) dmspart1 $\left(\frac{n_{i,t-1}}{N_t}\right)$
LOWPERFX	5.193	5.180	4.502	0.557	0.000	0.267
	(0.00)	(0.00)	(0.00)	(0.00)	(0.42)	(0.53)
MIDPERFX	2.398	2.718	0.536	-0.022	-0.000	0.087
	(0.00)	(0.00)	(0.00)	(0.80)	(0.51)	(0.67)
HIGHPERFX	10.015	10.097	8.093	0.757	-0.000	0.738
	(0.00)	(0.00)	(0.00)	(0.00)	(0.89)	(0.02)
CAT_DMS	0.930	0.942	0.084	0.268	-0.000	-0.002
	(0.00)	(0.00)	(0.00)	(0.00)	(0.74)	(0.90)
LAG_DMSNEW	0.261	0.261	0.034	-0.003	0.000	-0.016
	(0.00)	(0.00)	(0.00)	(0.00)	(0.97)	(0.00)
LOGLAGFUNDSIZE	-0.137	0.000	-0.174	-0.021	-0.000	3.539
	(0.02)	(1.00)	(0.00)	(0.01)	(0.98)	(0.00)
LOGLAGFAMSIZE	-0.009	-0.012	0.006	0.002	0.000	-0.075
	(0.10)	(0.03)	(0.09)	(0.01)	(0.64)	(0.00)
LAG_TURNOVER	-0.001	-0.001	-0.000	0.000	0.000	-0.002
	(0.03)	(0.09)	(0.89)	(0.23)	(0.65)	(0.00)
LAG_NETEXPRATIO	-0.332	-0.370	-0.227	-0.016	0.000	-0.278
	(0.00)	(0.00)	(0.00)	(0.47)	(0.20)	(0.00)
AGE	-0.064	-0.062	-0.027	0.001	0.000	0.054
	(0.00)	(0.00)	(0.00)	(0.04)	(0.95)	(0.00)
RETVOL	-28.548	-27.752	-21.081	1.384	-0.000	-13.230
	(0.01)	(0.01)	(0.00)	(0.80)	(0.34)	(0.00)
CONSTANT	2.798	0.168	3.247	0.595	0.488	-62.910
	(0.03)	(0.91)	(0.00)	(0.04)	(0.26)	(0.00)
Observations	186318	186318	186318	186318	186318	186318
R^2	0.24	0.24	0.12	0.39	0.00	0.49

Panel B: Continuous Model

	(1) dms $\frac{n_{i,t}}{N_t} - \frac{n_{i,t-1}}{N_{t-1}}$	(2) dmsnew (4)* [(1)+(2)+(3)]	(3) dmspart2a $\frac{f_{i,t}}{n_{i,t-1}}$	(4) dmspart2b $r_{i,t}$	(5) dmspart2c $\left(\frac{N_t - N_{t-1}}{N_{t-1}}\right)$	(6) dmspart1 $\left(\frac{n_{i,t-1}}{N_t}\right)$
RANK	1.096	1.180	0.295	0.121	0.000	-0.240
	(0.06)	(0.06)	(0.39)	(0.29)	(0.12)	(0.59)
RANK_SQUARE	3.952	4.037	3.169	0.204	-0.000	0.535
	(0.00)	(0.00)	(0.00)	(0.04)	(0.12)	(0.16)
CAT_DMS	0.929	0.941	0.083	0.267	-0.000	-0.002
	(0.00)	(0.00)	(0.00)	(0.00)	(0.76)	(0.85)
LAG_DMSNEW	0.261	0.261	0.034	-0.003	0.000	-0.016
	(0.00)	(0.00)	(0.00)	(0.00)	(0.94)	(0.00)
LOGLAGFUNDSIZE	-0.134	0.002	-0.172	-0.021	-0.000	3.541
	(0.03)	(0.97)	(0.00)	(0.01)	(0.98)	(0.00)
LOGLAGFAMSIZE	-0.009	-0.012	0.006	0.002	0.000	-0.075
	(0.10)	(0.03)	(0.08)	(0.01)	(0.77)	(0.00)
LAG_TURNOVER	-0.001	-0.001	-0.000	0.000	0.000	-0.002
	(0.02)	(0.08)	(0.77)	(0.23)	(0.59)	(0.00)
LAG_NETEXPRATIO	-0.328	-0.371	-0.224	-0.015	0.000	-0.282
	(0.00)	(0.00)	(0.00)	(0.50)	(0.20)	(0.00)
AGE	-0.064	-0.062	-0.027	0.001	0.000	0.054
	(0.00)	(0.00)	(0.00)	(0.04)	(0.94)	(0.00)
RETVOL	-29.152	-28.357	-21.257	1.384	-0.000	-12.993
	(0.01)	(0.01)	(0.00)	(0.80)	(0.52)	(0.00)
CONSTANT	3.300	0.691	3.826	0.654	0.488	-62.871
	(0.01)	(0.63)	(0.00)	(0.02)	(0.26)	(0.00)
Observations	186318	186318	186318	186318	186318	186318
R^2	0.12	0.39	0.00	0.49	0.23	0.24