Time-varying Skill

Managing Mutual Fund Returns to Scale

MAURICE McCOURT*, QI ZENG*

November 21, 2021

ABSTRACT

Actively managed mutual funds exhibit heterogeneous and time-varying returns to fund and industry scale. When a fund starts out, it exhibits *increasing* returns to scale (IRS) to industry size and *decreasing* returns to scale (DRS) to fund size. As funds get older and larger, industry size IRS turns negative to become DRS. The fund size (industry size) DRS coefficient is a concave (convex) function of fund size. The industry DRS component of fund performance is the main driver of the flow-performance relationship.

JEL classification: G23, G34.

*Department of Finance at the University of Melbourne. Qi Zeng can be contacted at qzeng@unimelb.edu.au, Maurice McCourt is at maurice.mccourt@unimelb.edu.au. We would like to thank Konark Saxena, Eric Tan, Neal Galpin, and seminar participants at FIRN 2021 for valuable feedback.

Ever since the seminal work showing active mutual funds decreasing return to scale (DRS) to fund size (Berk and Green, 2004) and to industry size (Pástor and Stambaugh, 2015), there have been ample indications that fund DRS sensitivities are heterogeneous. Earlier studies show that the sensitivities are related to fund characteristics (Bris, Gulen, Kadiyala, and Rau, 2007; Chen, Hong, Huang, and Kubik, 2004; Pollet and Wilson, 2008; Yan, 2008). Recent studies find that sensitivities also differ across industry characteristics (Feldman, Saxena, and Xu, 2020) and across time period (Harvey and Liu, 2021; Harvey, Liu, Tan, and Zhu, 2020).

In this paper, we assume that the sensitivities of DRS of an active mutual funds are of the following format:

$$R_{it} = a_{it-1} + b^S_{it-1} S_{t-1} + b^q_{it-1} q_{it-1} + \epsilon_{it}, \tag{1}$$

where R_{it} is the benchmark adjusted gross return, S_{t-1} is the industry size measured as the ratio of aggregate active mutual fund size to total stock market capitalization, q_{it-1} is the fund size (Pastor, Stambaugh, and Taylor, 2015, PST). Motivated by recent work on time-varying estimations (Harvey and Liu, 2021), we use a rolling-window approach to estimate coefficients at fund level. The main drawback with the rolling-window approach is the lack of power due to noise and measurement error, so a portfolio approach is usually adopted to mitigate these issues. Here we use a different approach by controlling for a set of variables that are likely to be correlated with the sensitivities and therefore likely to be associated with measurement error. Our goal here is to study whether there exist "residual" sensitivities that are not likely to be related to measurement error or to capacity restrictions such as liquidity proposed in the literature.

We find robust evidence that significant sensitivities exist for both industry size and fund size. Specifically, we find significant *increasing* returns to scale (IRS) to industry size and DRS to fund size. Since we control for fund size and other fund characteristics, this is the return to scale property when a fund just starts. Just as the fund fixed-effects after controlling for fund size is characterized in the literature as the return earned on the first dollar invested in the fund, the residual sensitivities that we examine are the sensitivities of the return on the first dollar invested in a fund when no other funds are present in the industry and fund, family, and industry characteristics are set to zero. We regard these residual sensitivities as the innate *ability or skill* of a fund when managing increases in fund and industry size. Since the sensitivities have a negative (and non-linear) relationship with size, funds will exhibit DRS for both industry and fund size when they become larger.

To estimate the sensitivity of funds' benchmark adjusted returns to industry and fund size, we follow the approach taken by PST and Zhu (2018). In this approach, fund fixed-effects captures the skill of the fund, and small-sample bias (Stambaugh, 1999) is addressed using a recursive-demeaning process. Our specific implementation follows Zhu (2018) by using fund size as an instrument for recursive-demeaned fund size and running a multiple 2SLS regression on industry size and fund size. We estimate the sensitivities twice, once using level fund size and again using log fund size.

After estimating b^S, b^q, we then project these coefficients onto a set of variables that are motivated by either the literature or from our preliminary examination. In fact, we use three sets of variables: Fund characteristics, fund family characteristics, and market conditions. Our main fund characteristics include fund age, turnover, expense ratio etc (Pastor et al., 2015), as well as the number of fund managers (Harvey et al., 2020). We pay close attention to fund size itself as prior studies already show that it is at least related to fund level DRS (Zhu, 2018). The fund family characterstics include the number of funds in the family and family TNA. Market-wide characterstics include the aggregate number of actively managed funds and fund families, as recent studies show that fund concentration can affect the competitiveness of the active mutual fund industry (Feldman et al., 2020). Finally, we include market factors such as the Pastor-Stambaugh liquidity factor and the Fama-French five factors. These factors capture market-wide liquidity, market conditions, and fund category effects.

While we are unable separate the true relationship between estimated b^S, b^q and these variables from measurement error, we can report that they broadly support the view that liquidity costs do indeed strengthen DRS effects at both the industry and fund level. Here by "strengthen" we mean the coefficients become more negative. In fact, fund turnover and expense ratio have a robust negative relation with both b^S, b^q no matter what model specification we use. Moreover, fund age also exhibits a robust negative relation with both b^S, b^q throughout, partially confirming our argument that when a fund gets older, its DRS effect becomes stronger. One of the most important findings in Pastor et al. (2015) is that a fund's innate skill, measured as the return earned on the first dollar invested in the fund, increases over its age. They test this by using the fund fixed-effects plus an age dummy. In our setting, we confirm this relation directly using the estimated a_{it}.

One of our most intriguing findings concerns the relation between b^S, b^q and fund size. Zhu (2018) notes that in her sample, a large proportion of funds exhibit positive b^q. She regards fund level estimates as noisy and adopts a portfolio approach by forming size-sorted portfolios. The result is a monotonic negative relationship between average fund size in a portfolio and the magnitude of b^q. However we are unable to find a robust monotonic relationship between fund size and either b^S or b^q, even though we recover the monotonic linear relationship in the cross-section using the full sample if we force a linear regression. This is because fund size changes over time: Average fund size over the life of the fund is not a good representation of its DRS sensitivities. Pastor, Stambaugh, and Taylor (2020) provide a theoretical model showing that b^q should in fact be non-linear. Using the estimated b^q_{it}, we can show that there exists a concave relation between b^q and fund size, implying that the power term γ in Pastor et al. (2020) should be between 2 and 3. Interestingly, the relation between b^S_{it} and fund size is convex. We are not aware of any existing theory linking b^S to fund size, but our result is one that any such theory has to explain. These relations further validate our conviction that there is time-varying interaction between coefficients b^S, b^q, and using the portfolio approach might obscure their true relationship.

We turn next to investors' response to fund performance under DRS. Having the monthly timeseries estimates of each fund's b^S_{it}, b^q_{it} enables us to study directly the relationship between fund flows and fund performance. Prior literature has documented a positive and convex relation between fund flows and lagged fund performance[1]. Under DRS, a fund's benchmark-adjusted return can be decomposed into three components: the return on the first dollar invested in the fund with no other funds present in the industry, and the two DRS components related to industry size and fund size. Under our methodology, we can also use the fund fixed-effects from the panel regression on fund, family and market characteristics to separate between fitted b^S, b^q and their residuals. Our goal here is to study which component(s) contribute to the flow-performance relationship. Our main results are twofold. In terms of the linear flow-performance relationship, the effects of the fitted b^S, b^q components on future flows are larger in magnitude than the effects of the first dollar returns, and of the residuals of b^S, b^q, even though they are all statistically significant. Relatively speaking, industry DRS, both fitted and the residuals, have much larger effects than fund DRS (and first dollar returns). In terms of the convex flow-performance relationship, it is only the residual industry DRS component that drives the convexity.

Methodologically, we use a two-step process as stated above. We first run the unbiased 2SLS RD estimation using rolling-windows, and then project onto a set of control variables. We do not use the portfolio approach to address the measurement error problem with short horizon estimation because there are so many dimensions that might cause the heterogeneity. One potential alternative is to incorporate the characteristics directly into the equation (1) through interaction terms and estimate them using the full sample. However, we are not sure how to account for the potential biases in the estimated coefficients. Instead we address the measurement error issue by using a panel regression projecting b^S, b^q onto variables that control for fund and time correlation among the variables.

Our paper makes three main contributions to the literature. First, we present a systematic empirical study of active mutual funds' heterogeneous and time-varying DRS for both

[1] See, for example, Ippolito (1992); Gruber (1996); Chevalier and Ellison (1997); Sirri and Tufano (1998); Ivković and Weisbenner (2009); Spiegel and Zhang (2013).

industry size and fund size. To the best of our knowledge, we are the first to document that a fund starts with industry size IRS which becomes DRS when it gets older and larger. Second, we confirm many prior findings on the relation between the DRS sensitivities and fund characteristics, with the caveat that we cannot separate the true relation or the measurement error. Among these, two new results stand out. We give an estimate of the power term in the model from Pastor et al. (2020). We find that there is a concave relation between b^S and fund size. Finally, we show that the industry DRS component of fund performance is the main driver of the convex flow-performance relationship.

The remainder of the paper is structured as follows. In the next section, we present some preliminary evidence from the full-sample estimation to motivate the study. Then we present the methodology and our first set of main results, the projection of the estimated b^S, b^q on the fund, family and market characteristics. Next we study the flow-performance relationship using the estimated b^S, b^q. Finally we conclude.

I. Decreasing Returns across Time and across Funds

We start our exercise by briefly reviewing some classic results in the decreasing returns to scale (DRS) literature. Given the debate on the robustness of the DRS results at both industry and fund level (Adams, Hayunga, and Mansi, forthcoming; Pastor, Stambaugh, Taylor, and Zhu, 2021), we strongly believe that DRS is based on such a sound economic reasoning that it should hold in almost any sample. Although it is not the main goal of this paper to add to this debate, we use a somewhat different starting point from other papers to create yet another sample of mutual fund data. Our sample is based on the intersection of CRSP and Morningstar mutual fund data, however instead of starting with the whole CRSP survivorship-bias-free mutual fund universe (Berk and Van Binsbergen, 2015; Pastor et al., 2015), we first select actively-managed open-end US equity funds according to CRSP style and objective codes, which are in turn derived from Wiesenberger, Strategic Insight and Lipper Objective codes. Specifically we select those funds with style and objective codes starting with the letter “E”. We then follow the steps described by PST to clean and merge with Morningstar mutual fund data. We exclude funds that do not belong to any of the Morningstar 3x3 US equity fund categories.

Our full sample period is January 1990 to June 2019, but since we will focus later on a three-year window in most of our estimations, the actual duration of the sample is January 1993 to June 2019. Benchmark-adjusted fund returns are estimated as the fund’s monthly gross return in excess of the fund’s Morningstar 3x3 category benchmark. Each fund’s TNA is estimated as the aggregate TNA of the fund’s share classes. Monthly TNA values are adjusted for inflation following PST by multiplying the TNA by the market value of all stocks in CRSP that month and dividing by the market value of all stocks in CRSP in December 2011. We estimate industry size each month as the ratio of the aggregate TNA (without adjusting for inflation) of all funds in the sample that month to the market value of all stocks in CRSP that month.

[Insert Table I here]

[Insert Table II here]

Tables I and II produce the results from PST, Zhu (2018), and Pastor et al. (2021) using our sample. In Table I, we run OLS regressions of benchmark-adjusted gross returns on lagged industry size, fund size and log fund size. We report the results for returns and winsorized returns, while all the independent variables are winsorized following PST. We can see that industry size DRS is robust while fund size DRS is not. If anything, the log fund size coefficient suggests increasing returns to scale. Table II reports results for both a panel regression with fund fixed-effects and a 2SLS panel regression. More specifically, the tables show the results from the following DRS relationship between funds’ benchmark-adjusted gross returns R_{it} on industry size S_{t-1} and (log) fund size q_{it-1}:

$$R_{it} = a_i + b^S S_{t-1} + b^q q_{it-1} + \epsilon_{it}, \tag{2}$$

where a_i is the fund fixed-effects, b^S and b^q are the sensitivities of benchmark-adjusted returns with respect to the industry size and (log) fund size. In the 2SLS regression, all variables are recursively forward-demeaned at the fund level with (natural) fund size as an instrument for forward-demeaned fund size, following Zhu (2018) (we denote this approach to estimating size and scale coefficients the 2SLS RD regression). The DRS for both industry size and fund size are significant and robust. We can compare our results from Table II with those of PST and Zhu (2018). Looking at the univariate 2SLS RD regression of benchmark-adjusted returns on lagged industry size (model (1)), our coefficient on lagged industry size is -0.0293, which is close to that reported by Zhu (2018) (-0.0291) and PST (-0.0326). In the univariate 2SLS RD regression of benchmark-adjusted returns on lagged log fund size (model (4)), we find a coefficient of -0.0014 which is roughly comparable with that found by Zhu (2018) of -0.0026. In unreported tests, we confirm that the results hold for the periods March 1993 to December 2011 and January 1995 to December 2014.

A. Rolling-Window Estimates

As discussed in the introduction, the literature has shown ample indications that DRS effects change over time and vary across funds. Here we illustrate the time-varying DRS in our sample. Using rolling-windows, we run a 2SLS RD panel regression for each window. In other words, we assume in this subsection that the benchmark-adjusted returns satisfy the following equation:

$$R_{it} = a_{it-1} + b^S_{t-1} S_{t-1} + b^q_{t-1} q_{it-1} + \epsilon_{it}, \tag{3}$$

where S_{t-1} is industry size, q_{it-1} is the (log) fund AUM. As we run the regressions for both fund size and log fund size, there are always two sets of the regression coefficients (a, b^S, b^q). We denote level coefficients if using fund size, and log coefficients if using log fund size.

Figures 1 and 2 show the results for three-year and five-year rolling windows, respectively. We can see that unlike the full sample results, there are significant periods during which DRS effects are simply not there. The most notable episodes occurred during the late 1990s and early 2000s, and again just before the 2008 GFC. During these periods, actively managed US equity funds exhibited significant increasing returns to scale (IRS) relative to industry size. Funds also exhibit IRS relative to fund size during the early and late periods of the sample. A further observation is that there are some extreme estimates of the regression coefficients, reflecting the effects of the short time horizons.

[Insert Figure 1 here]

[Insert Figure 2 here]

In the last rows of the two figures, we plot the average intercept a across funds. While the benchmark-adjusted gross return is the usual measure of fund performance studied in the literature, a is the return on the first dollar invested in the fund with no other funds present in the market. PST point out that DRS effects might cloud the true skills of the funds if one simply looks at the overall benchmark adjusted returns. Here we observe a seemingly inverse relationship between average a_{it} and the DRS coefficients. For example, during the late 1990s and early 2000s, funds exhibited IRS relative to industry size and fund size before crashing. In the meantime, average a_{it} dropped sharply during the same period. This suggests the following explanation for the performance behavior at that time: The main driver of performance then was the increasing fund size. There were ample profitable investment opportunities in the equity market and one just needed to attract capital inflow. At the same time, one does not need very highly skilled managers to seek out such opportunities. Of course, what was going on in the markets in that period was the dot-com bubble.

To formally establish whether the estimated full sample coefficients are different across time, we use two tests. We separate the sample time periods into three-year non-overlapping intervals. Our first test is a one-way ANOVA test to check whether each of (a, b^S, b^q) are different across the intervals. Table III shows the result. The first two columns show that with the exception of b^q in the level regressions, all the other coefficient means are significantly different from each other across the time periods. Our second test is the Jonckheere-Terpstra J^* test (Jonckheere, 1954; Terpstra, 1952). This tests whether there exists a trend across the sequence of time intervals. The last three columns of the table show the results. We can see that the tests reject the null that there is no trend for the three level estimates and industry b^S for the log estimate. In fact, there is an upward trend for a (consistent with the finding by PST of increasing aggregate skill over time) and b^q and a downward trend for industry DRS b^S in both the level and log estimations. These results are robust whether we use two or five year windows.

[Insert Table III here]

B. Fund-by-Fund Estimates

We now turn to heterogeneous DRS coefficients across funds. Zhu (2018) already noted that in her sample, many funds exhibit extreme or even positive fund DRS values. To overcome this, she uses size-sorted portfolios to impute a monotonic relationship between b^q and average size of the portfolio over the full sample period. Pastor et al. (2020) also establish that fund characteristics affect DRS. To examine further fund-specific DRS, we run 2SLS RD regressions for each fund using the full sample period. In other words, we assume the following relationship:

$$R_{it} = a_i + b_i^S S_{t-1} + b_i^q q_{it-1} + \epsilon_{it}. \tag{4}$$

We require that a fund must have at least 12 monthly observations. Then we plot the estimated (a_i, b_i^S, b_i^q) against log fund size. Figure (3) shows the result.

[Insert Figure 3 here]

In this figure, we plot the fractional polynomial (Royston and Altman, 1994) fitted curve and 95th percentile confidence interval of the three coefficients against log fund size. We first notice that at the fund level, we recover the DRS results documented by PST and Zhu (2018). For both industry and fund DRS, the majority of the funds exhibit DRS. In the rolling-window case, funds can have IRS during short episodes, however a fund simply cannot have IRS over the long run.

Furthermore, we observe that there seems to be a monotonic positive relationship between b_i^S and log fund size. Surprisingly however, the relationship between b_i^q and log fund size is not linear. It is not even monotonic for large funds, though it is very flat. Overall, we can safely observe that extremely small funds have lower (more negative) b_i^S and b_i^q, consistent with findings in Zhu (2018). When funds get larger, b_i^S becomes less negative. b_i^q, however, is non-monotonic. Pastor et al. (2020) shows that there are other fund characteristics that might play a role in determining the relationship between b_i^q and fund size. Using an equilibrium model, they shows that there is a rich cross-sectional relationship between fund size, expense ratio, liquidity and fund activeness.

To formally test whether there are indeed differences in the cross-section, we again run Anova and Jonckherre-Terpstra tests in the cross section. Here we show the heterogeneity across two dimensions, size and fund categories. Table IV Panel A presents the results of Anova and Jonckheere-Terpstra tests for variation in the coefficients across fund size deciles. For these tests, the coefficients are estimated fund-by-fund using 2SLS RD regressions over the full sample period. Funds are grouped from small to large into TNA deciles based on their mean TNA over the full sample period. Consistent with the graphs, there is little variation for level or log b_i^S across size deciles, however the Jonckheere-Terpstra J* statistic is positive suggesting a weak upward trend as fund size increases. The null of no variation across size deciles is strongly rejected for level and log b_i^q. The J* statistic is large and positive for b_i^q, indicating that b_i^q is increasing as size decile increases. The Anova F-statistic for the level a_i coefficient is not significant (meaning there is no significant variation across size deciles), but the log a_i coefficient is significant at the 5% level. The J* statistic is negative but not significant for both level and log a_i, suggesting a weak downward trend as fund size increases.

Table IV Panel B gives the results of Anova tests for variation in the fund-by-fund coefficients across 3x3 Morningstar categories (the mean coefficients for each category are depicted in Figure 4). The null of no significant variation in the category means for a_i and b_i^S is rejected as indicated by the high Anova F-statistic values for those coefficients. However the mean b_i^q coefficient does not differ significantly between categories.

[Insert Table IV here]

[Insert Figure 4 here]

To summarize this section, we reproduce results from prior literature for fund and industry DRS in the full sample, and examine DRS effects separately for the timesereis and for the cross-section. However we have not yet examined DRS effects jointly for the time-series and the cross-section together; we do this in the next section.

II. Skill in Managing Decreasing Returns to Scale

A. *Dynamic Fund-Specific DRS coefficients*

Motivated by the results from the previous section, we estimate the following relationship between funds' benchmark-adjusted gross returns on industry scale and (log) fund size that appeared in the introduction:

$$R_{it} = a_{it-1} + b_{it-1}^S S_{t-1} + b_{it-1}^q q_{it-1} + \epsilon_{it},$$

where $(a_{it-1}, b_{it-1}^S, b_{it-1}^q)$ are time-varying and fund-specific.

To estimate the fund-level time-varying heterogeneous size and scale coefficients $(a_{it}, b_{it}^S, b_{it}^q)$ we run fund-by-fund 2SLS RD regressions using rolling windows. Summary statistics for the data sample that we use for the regressions are given in Table V. The top panel shows the statistics for fund, family and market characteristic variables. Overall, our sample is more or less similar to that used by PST.

The main focus of our interest however is the estimated DRS coefficients in the bottom panel of Table V. We multiply b_{it}^q (level) by 1000 to facilitate reading. Compared with the full sample estimation from Table II, we can see that the average fund-level heterogeneous and time-varying b_{it}^q is much larger in magnitude compared to the full-sample estimation where all b^q are forced to be the same for all funds. In fact, the difference is about 1000 times for both the level and log fund size. If we look across the distribution, we note that at the 75th percentile, b_{it}^q is already positive, which is consistent with Zhu (2018)'s observation in her sample. A more significant issue is the industry DRS coefficient, b_{it}^S. The average b_{it}^S across the funds is *positive* for both the level and log regressions. If we look across the distribution, it is positive already at the median.

To take a closer look the details of coefficients $(a_{it}, b_{it}^S, b_{it}^q)$, we plot the time-series and cross-sectional graphs in the same way that we did for the full-sample case in the previous section. The time-series plot in Figure 5 shows the fractional polynomial fit and the 95th percentile confidence interval of the monthly coefficients from the log regression. The graph is quite similar to the one we observed in Figure 1, both in terms of pattern and in terms of magnitude. In fact, here we observe a downward trend in b_{it}^S and an upward trend in b_{it}^q. The consistency implies that market conditions are likely determinants of the DRS sensitivities. The time-series graph for a_{it} indicates that a decreases over time. This is a divergence from the full sample graph in Figure 1, where a appears to be increasing over time.

[Insert Figure 5 here]

To examine more systematically the differences and trends across time for the $(a_{it}, b_{it}^S, b_{it}^q)$ coefficients, we perform Anova F-tests and Jonckheere-Terpstra tests of coefficient means using sequential (non-overlapping) 3-year time intervals. We report the results in Table VII. The Anova F-statistics for all coefficients (both level and log) are large and significant, meaning the coefficient means differ significantly across time intervals. The J* statistic is positive and significant for b^q_{it} (level and log), indicating that there is a strong increasing time trend for fund size DRS. For the industry scale DRS coefficient (b^S_{it}) (level and log), the J* statistic is negative and significant, indicating that industry DRS is weakening over time. The J* statistic for a_{it} is negative (significantly so for log a_{it}) indicating that a_{it} does indeed decrease over time. As noted in the previous paragraph, the negative J* statistic for a_{it} contrasts with the positive and significant J* statistic for a_t found in the tests using the full sample reported previously in Table III.

[Insert Table VII here]

Turning to the cross-section, we plot the fitted coefficients $(a_{it}, b^S_{it}, b_{it})$ along with the 95th percentile confidence interval against log fund size in Figure 6. Starting with b^S_{it}, we observe a downward trend and somewhat convex pattern over log fund size. There are a large proportion of positive b^S_{it}. The graph of b^S_{it} differs significantly from the graph of the full-sample b^S_i given in Figure 3. In the full sample case, we have an almost linear monotonically increasing relationship between b^S_i and log fund size, and most of the b^S_i are negative. We will come back to this point in a moment.

[Insert Figure 6 here]

The cross-sectional pattern for the level fund size coefficient in the full-sample estimation (b^q_i) is similar to that in the rolling-window estimations (b^q_{it}). The log b^q graphs have slightly differing patterns: In both graphs, the relationship increases sharply as very small funds get larger, but then declines slowly in the full sample graph, while in the rolling-window graph there is a relatively sharp decline as mid-sized funds get larger. All graphs exhibit a concave pattern, and the b_q values are almost always negative. This is consistent with the prediction made in Pastor et al. (2020). In their model, they produce an explicit relationship between b^q and q (see equation (43)), and it is negative. We will present more concrete evidence on the magnitude of the parameters in their expression in the next section. Finally when we compare a_{it} and a_i, the log regression results are similar: There is a convex relationship between a and log fund size, and the magnitudes are similar. They are both almost always positive.

We now return to the significantly different cross-sectional pattern for b^S between the full-sample estimation and the rolling-window estimation. Note that the only difference between the two estimation approaches is the way that (log) fund size is calculated. In the full-sample estimation, the average fund size over the entirety of the fund's life-time in the sample is used, while in the rolling-window case, we use the average fund size over the three-year windows. The significant difference between the two approaches implies that b^S changes over the life-time of the funds. If we consider the two graphs together, the following story emerges: When funds are small, they have *increasing* returns to industry scale. When those funds grow larger, their returns to industry scale is *decreasing*. In the full sample graph, this dynamic behavior is obscured because b^S is estimated using the full sample means of each fund's size. The similarity in the pattern for level b_i^q and b_{it}^q implies that there is no such life-cycle pattern for returns to fund size. Finally, the difference in the patterns for a_i and a_{it} implies that returns to industry scale have a significant effect when estimating the return on the first dollar invested in a fund when no other funds are present in the industry.

We formally test the above hypothesis by comparing equal-weight and value-weight coefficients. Table VI shows the results. In Panel A, we calculate the equal-weighted average of the three coefficients over fund size in the three-year rolling window estimation and, as a sanity check, we also run univariate regressions with industry scale S and (log) fund size q. In the univariate regressions (columns (1-3)), the coefficients are all negative and significant. Thus dynamic DRS holds quite well when fund size and industry scale are considered independently of each other. In fact, we plot the univariate results across funds in Figure 7. One can see that the univariate coefficient patterns are similar to the multiple regression coefficients graphed in Figure 6, but most b_{it}^S become negative in the univariate case. However, the discussion in the previous paragraph implies that the univariate results potentially conceal the correlation between the two regressors. Indeed, in the multiple regressions, the average coefficients for b_{it}^q are also negative. However, the (equal-weighted) average coefficients for b_{it}^S are positive and significant. To show that it is fund size that contributes to the difference in the b^S, in Table VI Panel B we present value-weighted mean coefficient values. All coefficients including b_{it}^S are now negative. This switch in sign for b_{it}^S is evidence that larger funds have smaller (more negative) industry scale coefficients than smaller funds.

[Insert Table VI here]

[Insert Figure 6 here]

Finally, we check for variation in the dynamic coefficients across the 3x3 Morningstar categories. The average sensitivities to fund size and industry scale for funds in each category are graphed in Figure 8. The average sensitivities to fund size (b_{it}^q) are negative and large in magnitude for large-cap and mid-cap funds, and noticeably smaller in magnitude for small-cap funds. The industry scale coefficients (b_{it}^S) are negative for large-cap funds, but positive for mid-cap and small-cap funds. The average intercept (a_{it}) is positive across all categories, but is larger for large-cap funds and smaller for small-cap funds.

Anova F-test results for differences in coefficients across the Morningstar categories are presented in Table VIII Panel B. The F-statistics for level and log a_{it} and b_{it}^S are large, suggesting significant differences in category means for these coefficients. For level and log b_{it}^q however the F-statistic is not significant, thus the mean values for b_{it}^q do not vary significantly across categories.

[Insert Figure 8 here]

[Insert Table VIII here]

B. Determinants of Dynamic DRS coefficients and Fund Skill

In the previous section, we document the dynamic patterns of (a, b^S, b^q). As we have discussed in the introduction, the literature argues that the sensitivities (b_{it-1}^S, b_{it-1}^q) are functions of fund liquidity, teamwork and other characteristics. In this section, we examine the relationship between the estimated coefficients for a set of determinants. More specifically, we run a panel regression of the following format:

$$b_{it} = c_i + d^x x_{it} + d^f f_{it} + d^m m_t + u_{it}, \tag{5}$$

where b_{it} are either b^S or b^q for level or log regressions, x_{it}, f_{it}, m_t are a set of fund characteristics, family characteristics and market conditions respectively. Specifically, our fund characteristics include fund size, age, turnover, expense ratio, number of managers per fund, and the proportion of fund TNA sourced from institutional investors. Fund age, turnover, expense ratio have been used in the literature as proxies for the liquidity of the funds. The number of fund managers have been shown to affect the DRS of funds as well (Harvey et al., 2020). The fund family characteristics include the number of funds in the family and family TNA. Variables capturing market conditions include the risk-free rate, Fama-French 5-factors (market premium, HML, SMB, RMW, CMA), and the Pastor and Stambaugh liquidity factor LIQv. We also include the aggregate number of actively managed funds and the number of families (fund management firms) of actively managed funds each year as proxies for industry concentration (Feldman et al. (2020)). Variable definitions are given in Table A1. Table (V) gives summary statistics for the fund characteristics, family characteristics and market conditions variables.

One might wonder why we use a two-step process in the estimation. If we substitute equation (5) into the DRS equation (1), it is simply a panel regression with interaction terms. However, the panel regression introduces a small-sample bias similar to that pointed out in PST and Zhu (2018), and in our case, correcting the bias for the interaction terms would require unwieldy complexity, so instead we adopt the two-step process. As long as the estimators described in PST and Zhu (2018) are unbiased and the coefficients do not change much over the window period, then our estimated coefficients are also unbiased (though they may be a little noisy given the short windows).

We focus our results on three-year rolling windows and require that a fund has at least 12 observations in each window. We use 2SLS RD estimators studied in Zhu (2018). Table (IX) and Table (X) present the first set of our main results. In Table (IX), the dependent variables are level coefficients while in Table (X) they are log coefficients.

[Insert Table IX here]

[Insert Table X here]

Columns (1) - (3) show the effects on b_{it}^S, (4) - (6) show the effects on b_{it}^q. Generally speaking, the coefficients do not change sign or significance when we include family and market variables, suggesting that these are relatively speaking independent factors. They are correlated of course, resulting in weakening but still significant sensitivity.

Let's first look at the fund characteristics. Fund age is negatively associated with b_{it}^S. That is, the longer a fund exists, the stronger (more negative) the industry DRS. This is consistent with our previous argument that the longer a fund survives, the more likely it is to exhibit (stronger) industry DRS effect. This holds regardless whether we use level b_{it}^S or log b_{it}^S. The effect of age on fund size DRS b_{it}^q, however, is different. It is positively associated with level b_{it}^q, but negatively associated with log b_{it}^q. This suggests a possible non-linear relationship between fund age and b_{it}^q.

Turnover and fund expense ratio, two measures of liquidity, uniformly strengthen DRS effects for b_{it}^S and b_{it}^q for both level and log coefficients. This confirms the liquidity argument for DRS at both the industry level and fund level.

The number of managers for a fund strengthens industry scale DRS b_{it}^S, but weakens fund level DRS b_{it}^q. This is consistent with the arguments of Harvey et al. (2020) for the ability of teamwork to weaken fund DRS. However, teamwork does not moderate industry scale DRS. The proportion of a fund's TNA that is sourced from institutional investors captures the views of professional investors about fund skill - skilled institutional investors are likely to invest more in funds they perceive as skilled. Higher institutional investment is associated with stronger industry DRS, but weaker fund size DRS.

Our final fund level characteristic is fund size. As we have seen in Figure 6, the relationships between fund size and (b_{it}^S, b_{it}^q) are likely not to be linear. Here we include both a linear and a quadratic term for fund size. We can see that b_{it}^S is convex in log fund size with a downward trend: The linear term has much larger (negative) magnitude than the (positive) magnitude of the quadratic term. This holds for both level and log b_{it}^S. The downward trend is consistent with what we observe from Figure 6, but b_{it}^S there looks to be concave. We have to keep in mind that the graphs are plotted for univariate relationships, while here we have a large number of controls. Overall, the relationship between b_{it}^S and fund size is that larger funds have *stronger* industry DRS. Small funds have *increasing* returns to industry scale, while larger funds exhibit decreasing returns to industry scale. This holds whether we run multiple regressions or univariate regressions on fund size only.

The relationship between log size and b_{it}^q however is quite different. Zhu (2018) argues that there is a monotonically increasing relationship. This is consistent with Panel C in Figure 7, where level b_{it}^q is likely to have an upward trend in a univariate situation. However, we provide ample evidence to see that this view is not robust. In Tables IX and X, we can see that there is a concave (negative quadratic coefficient) relationship. Furthermore, the magnitudes of the linear term and quadratic term are similiar. The pattern holds for both level and log b_{it}^q. This suggests that very small and very large funds exhibit stronger fund DRS b_{it}^q than medium-sized funds. This empirical result gives a hint of the relationship between b_{it}^q and q_{it}. For example, if we take the model in Pastor et al. (2020) literally, their equation (43) gives an explicit relationship between b^q and q. Our results suggests that, other things being equal, the second order derivative of b^q ($\partial\alpha/\partial A$ in their paper), $(\gamma-2)(\gamma-3)<0$. Or $\gamma \in (2,3)$.

The number of funds per family and family TNA have opposite effects. Our argument to explain these findings is linked to the growth of current funds in the family instead of creating new funds to explore new investment opportunities in the industry. A fund in a family with more funds has weaker industry DRS b^S, but stronger fund DRS b^q. This first finding makes sense if the family is creating many small funds to take advantage of the weak industry DRS for small funds documented in the previous sections. The second finding is also easy to understand: Presumably a family with more funds, other things equal, is likely to have worse fund level DRS, otherwise it would prefer to grow its current funds instead of creating new funds.

Finally we turn to the market conditions. The number of concurrently active funds and fund families have no significant relationship with the industry size coefficient b_{it}^S. However, the number of active families strengthens b_{it}^q. Higher numbers of active fund families may reflect over-crowding in the market. A family will ensure that their own funds are set up so that they minimize competition with other funds in the same family. Thus, when industry scale is held constant, competition must come from funds in other families, so when the number of other families increases, true fund-level competition increases. Feldman et al. (2020) argue that increased competition strengthens fund size DRS by reducing incentives for fund managers to exert the effort required to find profitable investment opportunities as their funds get larger, which is consistent with our finding of stronger b_{it}^q when more families are active in the industry. After controlling for family-level competition, higher numbers of active funds may reflect the fact that market-wide conditions associated with weaker fund size DRS b_{it}^q (such as higher liquidity) encourages new funds to enter the industry.

Most market factors significantly affect industry and fund level DRS, and the effects are generally consistent between the level and log of b_{it}^S and b_{it}^q. One thing that immediately comes to view is the effect of the market-wide liquidity factor, LIQv of Pastor and Stambaugh (2003). When market-wide liquidity increases, DRS (both industry and fund) decreases. When the market-risk premium increases, industry DRS also decreases, as presumably there are more attractive investment opportunities in the industry. However the market-risk premium does not much affect the fund level DRS. The effect of the risk-free rate on industry DRS is not robust, but it significantly strengthens fund level DRS. Presumably higher interest rates are associated with higher borrowing costs, lowering market liquidity. The four factors (value HML, size SMB, profitability RMW, investment CMA) try to capture the effect of specific types of investment opportunities, similar to the category heterogeneity that we showed in the previous section (Figure 4). Using factors generalizes those results. We can see that HML and RMW both strengthen DRS at industry and fund levels, while CMA has the opposite effect, weakening DRS at both the industry and fund levels. SMB strengthens fund level DRS but weakens industry DRS. This means that during periods when small-cap funds perform well, there are ample industry-wide investment opportunities, but the existing funds are hard to grow. Note that these results have a different interpretation compared to the arguments of Pastor et al. (2015) when they interact fund characteristics with fund size and industry size. Their goal is to see whether proxies for fund level liquidity are associated with DRS. Here we are looking at how market-wide conditions are associated with DRS.

Having discussed the DRS coefficients, we now turn to results for the intercept a_{it} given in column (7). We note that there is a large negative correlation (-.82) between a_{it} and b_{it}^q (see Table A3), so variables that have positive effect on b_{it}^q have a generally negative effect on a_{it}. In the literature, a is the return earned on the first dollar in the fund with no other funds present in the industry (PST). One of the most important findings in PST is that fund performance decreases with fund age, but this relationship disappears after controlling for industry scale. They find a positive but marginally significant relationship between performance and age. In our setting, we obtain a time-varying and fund specific intercept, therefore we can directly test this relationship. Our results unambiguously confirm their findings. Specifically, fund skill, measured by a_{it}, is significantly positively associated with fund age. This result is robust, regardless of level or log and with all the controls. We further show that fund skill is also positively associated with fund turnover and expense ratio . The remaining fund level characteristics are either insignificant, or not robust (for example, the results for level and log are conflicting). At the family level, a fund belonging to a family with more funds is likely to have lower skill, but if it belongs to a family with higher family TNA, it is likely to have higher skill. Both of these findings are easy to understand: A family will create more funds if it cannot grow current funds, other things being equal. This is because of the lower skill of current funds. A family is likely to have higher family TNA if the funds in the family have higher skill.

Turning to market conditions, periods when market factor premia are at levels where it is easy to find profitable opportunities (e.g. reduced DRS), funds do not require higher skilled managers (higher a). Such periods include when market liquidity is higher, when market risk premium is higher, and when there is a higher premium for firms that invest conservatively. On the other hand, periods where finding profitable investment opportunities requires higher skill include when interest rates are higher, when value and size premia are higher, and when the premium for firms with robust profitability is higher.

In the above we discuss the relationship between (a, b^S, b^q) and the determinant variables as the true relation. But as noted in the introduction, this relation might be affected by measurement error introduced by our methodological approach. To the extent that a correlated variable is also likely to have measurement error, the effect that we observe might not be robust. However we believe our approach to controlling for measurement error is reasonable. This belief is underpinned by ex-ante arguments that measurement error is associated with and absorbed by control variables included in our model, and ex-post by the fact that we are able to robustly confirm many findings of prior studies.

C. Managing DRS as Skill

After controlling for as many as possible of the observable determinants (such as liquidity, market conditions etc) that might affect industry and fund DRS, we now examine the fixed-effects intercepts, the constant terms in Tables IX and X. Similar to skill measured by the fund fixed-effect coefficients a_i (PST), the intercepts here represent the sensitivities of fund performance with respect to changes in industry scale and fund size on the first unit change in funds' characteristics, family characteristics, and market conditions. Since we already incorporate in our model as many exogenous economic sources of DRS as possible, we regard the intercepts of b_{it}^S, b_{it}^q as the unobserved *skill* of funds in managing industry and fund level DRS.

Our null hypothesis is that the fund fixed-effects c_i in Equation 5 should be zero, or more specifically, the time-varying heterogeneous DRS effects should be fully captured by the fund and family characteristics and the market conditions. However, Table (IX) and Table (X) show that the average c_i^S, c_i^q are significantly different from zero. More specifically, $c_i^S > 0$, namely when a fund starts investing, it has *increasing* return to industry scale. This result is robust across different controls, and holds for both level and log. The result for c_i^q is not quite robust. It is negative for both level and log, but not significant in the log. That is, when a fund starts investing, it has marginally *decreasing* return to scale for fund size.

Note that we cannot say much about the fixed-effect intercept of fund DRS with respect to a, as the results on the controls are not robust, as discussed above.

To further confirm that c_i^S, c_i^q are measures of skill, we check whether adjusted $c_i^S + u_i^S, c_i^q + u_i^q$ are persistent. Table XI shows this result. In this table, we calculate the Markov transition matrix of coefficient persistence over consecutive 1-year and 3-year periods. The one-year horizon is our focus of interest, while we consider the 3-year horizon as a sanity check: Our results are estimated using 3-year rolling window with the underlying assumptions that (a, b^S, b^q) do not change over three-year window. If they are still persistent beyond the three year horizon, it means that our 3-year window assumption is not justified. To give the comparison more context, we also show the transition matrix for benchmark-adjusted returns across funds. We can see that, at the one-year horizon, adjusted (a, b^S, b^q) are strongly persistent, but this is not the case at the 3-year horizon. In unreported tests, we find almost identical result using unadjusted (a, b^S, b^q).

[Insert Table XI here]

III. Flow-Performance Relationship

Having a dynamic panel of estimated (a, b^S, b^q) from rolling-window approach enables a fresh examination of the flow-performance relationship from a DRS perspective. Following Chevalier and Ellison (1997); Sirri and Tufano (1998), we calculate the three-month flow for each fund as follows:

$$flow_{it+3}^{3m} = \frac{TNA_{it+3} - TNA_{it}(1 + R_{it+3}^{r3})}{TNA_{it}} \tag{6}$$

where TNA_{it} is the TNA of fund i at the end of month t, and R_{it}^{r3} is fund i's cumulative raw return in the 3 months up to and including month t.

We start by performing an analysis similar to that by Harvey and Liu (2021) of the relationship between loadings and future flows. Table XII shows this result. In this table, we run a panel regression of three-month fund flows on estimated DRS coefficients. Panel A shows the results for (a, b^S, b^q) and Panel B shows those for the residuals (a_r, b_r^S, b_r^q) from the panel regressions of DRS coefficients on fund, family and market characteristics described in the previous section. Control variables include gross returns over the past 1-month, 12-month and 24 -month periods, and TNA, expense ratio, flow and log fund age during the prior month. We include Morningstar 3x3 category $\times$ month fixed-effects and use robust standard errors clustered by fund and month.

We can see that the most robust relationship comes from the industry DRS sensitivity, for both b^S itself and the residual b_r^S. There is a positive relationship between current month industry DRS sensitivity and the fund flow over the following 3 months. This is consistent with the results from Harvey and Liu (2021). In terms of the first dollar return a, there is a negative relation with the 3-month flow in the univariate regressions (Table XII Panel A columns (1&5)), but it disappears in the multiple regression (columns (4&8)). The residual a_r is significantly positive in the residual regression without controls (XII Panel B column (1)), but turns insignificant when controls are included (Panel B column (8)).

A similar situation arises for the fund level DRS. The flow-performance relation is weakly significant for b^q, and significantly negative for b_r^q in the univariate regressions (columns (2) and (6)), but the effect disappears (or turns weakly negative) in the multiple regressions (columns (4) and (8)). This confirms that it is important to consider the correlation between the three variables simultaneously. We also show the results without the controls. In the multiple regressions using the residuals (Panel B), all three variables (a_r, b_r^S, b_r^q) show a strong positive relation with flows without controls (column (4)), but only industry DRS remains significant after including controls (column (8)). We want to emphasize here that our fund level results are not comparable to those in Harvey and Liu (2021) as they use different definitions of fund size and fund flow.

Our main focus here is to study the flow-performance relationship in more detail in light of the DRS, at both industry and fund level. Specifically, we adopt the view that there are essentially three components to the benchmark-adjusted return performance measure used in prior studies: The industry scale DRS (b^S), fund size DRS (b^q), and the return on the first dollar (a). Table XIII shows the same regressions as Table XII, but we use the components of the benchmark-adjusted return $(a, b^S S, b^q q)$ as independent variables instead of the coefficients (a, b^S, b^q). The question we then ask is, to what extent do flows respond to the three components of performance. This way we can explicitly compare the relative sensitivity among the three as they are returns.

In the results for the multiple regressions using plain coefficients (Table XIII Panel A column (8)), we can see that all three components are positively related to flows. In terms of magnitude, fund level DRS and first dollar returns are more or less similar, while the industry DRS component is much larger. The same results hold for the residuals (Panel B). We conclude from these results that investors do respond to fund skills measured by both the first dollar earned *and* the ability to manage DRS. Investors seem to be much more sensitive to funds' ability to manage industry DRS relative to the other two abilities.

At this stage, it is natural to create the full picture. We decompose the benchmark-adjusted return as follows:

$$E_{t-1}(R_{it}) = a_{it-1} + b^S_{it-1}S_{t-1} + b^q_{it-1}q_{it-1} \tag{7}$$

$$= a_{it-1} + b^S_{it-1,p}S_{t-1}b^S_{it-1,r}S_{t-1} + b^q_{it-1,p}q_{it-1} + b^q_{it-1,r}q_{it-1}, \tag{8}$$

where $b_{it-1,p}, b_{it-1,r}$ are fitted and residual values from the panel regression (Equation 5). We then project the flows on the five components of performance.

Table XIV shows the results. In this table, we first confirm that there is a positive relation between flow and benchmark-adjusted return itself and then project the flow onto the five components. Columns (1) and (2) show the results without controls, (3) and (4) give the results with controls. As in the previous two tables, we run panel regressions with category $\times$ month fixed-effects with robust standard errors clustered by fund and month.

From the five components, we can see that the fitted DRS related variables are the dominant performance components that investors respond to (column (4)). To the extent that these components signify liquidity and market related performance, investors choose funds that are much more liquid, or suffer less DRS. Of the two DRS components, investors are much more sensitive to industry DRS: Its sensitivity is more than double that of the fund DRS. Of the three remaining components, investors respond only to the funds' unobserved ability to manage industry DRS; the effects of the return on the first dollar and the unobserved ability to manage fund DRS are not significantly different from zero.

Finally, we examine the convex relationship between fund flows and past performance[2]. Figure 9 shows the results. In this figure, we plot the 3-month flows against fund performance components. Specifically, we rank each component of each fund's one-month performance into 20 (vigintile) bins ((Spiegel and Zhang, 2013)). Panel A plots flows against benchmark-adjusted returns. There is an obvious convex pattern that has been well-documented in the literature. There are two differences here. One is that the average flows are negative.

[2]See, for example, Ippolito (1992); Gruber (1996); Chevalier and Ellison (1997); Sirri and Tufano (1998); Ivković and Weisbenner (2009); Spiegel and Zhang (2013); Starks and Sun (2016).

Note that here we use the inflation adjusted AUM in calculating the flows. This reflects the negative flows out of the active managed mutual fund industry over recent decades. Presumably a large proportion of that flow went to the ETF industry. The other difference is an obvious sharp drop on the left tail. This actually reflects the extreme negative performance of the worst funds (bottom vigintile). In unreported tests, we observe an almost linear relationship between average performance (benchmark-adjusted returns) and performance rank, with the exception of the lowest rank: The returns of that vigintile are much lower than a linear relationship would predict. This does not show up in a coarser grouping such as quintile rank ((Harvey and Liu, 2021)).

To see which of the three components contributes to the convex relationship, we plot the same flow-performance relationship, but use $\{a, b^S S, b^q q\}$ respectively. Panel B shows this result. Here we can see that only the industry DRS component of performance exhibits a convex relationship, while fund level DRS and the first dollar returns both have a concave pattern. Given this finding, we further decompose b^S, b^q into fitted and residual components from a regression on fund, famiy and market characteristics (Equation 5). Figure 9 Panel C shows the flow-performance relationship using fitted b^S, b^q and their residuals. With this further decomposition, we can see that the convexity comes from the fitted b^S, b^q and the residual b^S_r. In summary, the graphs indicate that the main driver of the convexity comes from industry DRS. The fitted fund level DRS also exhibits a convex pattern, but not the residuals. The first dollar returns has no such pattern.

Figure 9 illustrates the univariate relationship betwen flows and performance components. To establish the robustness of the relationships, we run the following regressions (Spiegel and Zhang, 2013):

$$
\begin{aligned}
flow^{3m}_{t+3} &= c0 + c1perf_t + controls + \epsilon_t \\
flow^{3m}_{t+3} &= d0 + d1perf_t + d2perf_t^2 + controls + \epsilon_t \\
flow^{3m}_{t+3} &= e0 + e1perf_t + e2perf_t^2 + e3perf_t^3 + controls + \epsilon_t
\end{aligned}
\tag{9}
$$

where $perf_t$ is the vigintile rank of 1-month performance measured using either benchmark adjusted returns Ret_adj_t, first dollar returns a_t, or decomposed fitted values and residuals of fund and industy DRS components from regressions in fund, family and market characteristics. Control variables include gross returns over the prior 1-month, 12-month and 24 -month periods, prior month TNA, expense ratio, flows and log fund age.

Table XV gives the results. We first confirm that there exists a strong convex relation between flow and ranked benchmark-adjusted returns as indicated by the positive quadratic coefficient in column (2). However the fund DRS convexity observed in the graph disappears once we include controls, but the convex relationship for industry DRS still holds. If we further decompose b^S, b^q into their fitted and residual components, we can see that the ranked fitted and residual industry level DRS components still exhibit the convex relationship, while fund level DRS components now have a concave relationship. In short, the most robust driver of the convex relationship between flow and performance comes from the fitted industry DRS irrespective of whether we use raw values or include controls.

IV. Conclusions and Further Discussions

Using a dynamic approach adapted and refined from prior literature, this study confirms many existing findings on returns to fund and industry size in the mutual fund industry, and on the flow-performance relationship, and establishes a number of significant new facts.

We first address the idea that DRS effects are fixed across time and across funds, and show empirically that this is not the case. By taking a time-varying cross-sectional panel approach, we establish that DRS varies significantly in the time-series and in the cross-section. We uncover new trends and associations that were previously obscured in traditional time- and fund-invariant models (and in some dynamic models). In particular we identify significant non-linearity in the relation between fund and industry DRS and fund size.

Second, we show that time-varying cross-sectional DRS effects are linked to a battery of fund, family and market-level characteristics. These tests support the liquidity hypothesis for fund size and industry size DRS, and show that family and market characteristics also significantly impact funds' DRS. Furthermore we find that after controlling for exogenous characteristics, significant residuals remain which we attribute to fund skill in managing DRS effects.

Finally we show that fund performance can be decomposed into fund and industry DRS components, and of these components, the industry component has the greatest impact on fund flows. In particular, we find that the convexity of the flow-performance relationship is driven by a fund's unobserved ability to manage industry size DRS.

We acknowledge that there are potential drawbacks to our methodology relating to measurement error due to parameter estimation over short horizons. However we believe our approach to controlling for measurement error is reasonable. This belief is underpinned by ex-ante arguments that measurement error is associated with and absorbed by control variables included in our model, and ex-post by the fact that we are able to robustly confirm many findings of prior studies.

Our study opens up a number of interesting avenues for future theoretical and empirical research. We find a convex relation between industry DRS and fund size. This is an area where further theoretical modelling is required. We identify a relation between DRS and a range of exogenous variables. While some of these variables have been examined in recent literature examining mutual fund DRS (eg team management, industry concentration), others have not and are therefore candidates for further exploration.

REFERENCES

Adams, John C., Darren K. Hayunga, and Sattar Mansi, forthcoming, Scale and performance in active management are not negatively related, *Critical Finance Review* .

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

Berk, Jonathan B, and Jules H Van Binsbergen, 2015, Measuring skill in the mutual fund industry, *Journal of Financial Economics* 118, 1–20.

Bris, Arturo, Huseyin Gulen, Padma Kadiyala, and P Raghavendra Rau, 2007, Good stewards, cheap talkers, or family men? the impact of mutual fund closures on fund managers, flows, fees, and performance, *The Review of Financial Studies* 20, 953–982.

Chen, Joseph, Harrison Hong, Ming Huang, and Jeffrey D Kubik, 2004, Does fund size erode mutual fund performance? the role of liquidity and organization, *American Economic Review* 94, 1276–1302.

Chevalier, Judith, and Glenn Ellison, 1997, Risk taking by mutual funds as a response to incentives, *Journal of political economy* 105, 1167–1200.

Feldman, David, Konark Saxena, and Jingrui Xu, 2020, Is the active fund management industry concentrated enough?, *Journal of Financial Economics* 136, 23–43.

Gruber, Martin J, 1996, Another puzzle: The growth in actively managed mutual funds, *The Journal Of Finance* 51, 783–810.

Harvey, Campbell R., and Yan Liu, 2021, Decreasing returns to scale, fund flows and performance, *Working Paper* .

Harvey, Campbell R., Yan Liu, Eric K. M. Tan, and Min Zhu, 2020, Crowding: Evidence from Fund Managerial Structure, *Working paper* .

Ippolito, Richard A, 1992, Consumer reaction to measures of poor quality: Evidence from the mutual fund industry, *The Journal of Law and Economics* 35, 45–70.

Ivković, Zoran, and Scott Weisbenner, 2009, Individual investor mutual fund flows, *Journal of Financial Economics* 92, 223–237.

Jonckheere, A. R., 1954, A distribution-free k-Sample test against ordered alternatives, *Biometrika* 41, 133–145.

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

Pástor, Luboš, and Robert F. Stambaugh, 2015, On the Size of the Active Management Industry, *Journal of Political Economy* .

Pastor, Lubos, Robert F Stambaugh, and Lucian A Taylor, 2015, Scale and skill in active management, *Journal of Financial Economics* 116, 23–45.

Pastor, Lubos, Robert F Stambaugh, and Lucian A Taylor, 2020, Fund tradeoffs, *Journal of Financial Economics* 138, 614–634.

Pastor, Lubos, Robert F Stambaugh, Lucian A Taylor, and Min Zhu, 2021, Diseconomies of Scale in Active Management: Robust Evidence, *Working Paper* 21.

Pollet, Joshua M, and Mungo Wilson, 2008, How does size affect mutual fund behavior?, *The Journal of Finance* 63, 2941–2969.

Royston, Patrick, and Douglas G. Altman, 1994, Regression using fractional polynomials of continuous covariates: Parsimonoious parametric modelling, *Journal of the Royal Statistical Society: Series C* 43, 429–453.

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

Spiegel, Matthew, and Hong Zhang, 2013, Mutual fund risk and market share-adjusted fund flows, *Journal of Financial Economics* 108, 506–528.

Stambaugh, Robert F, 1999, Predictive regressions, *Journal of Financial Economics* 54, 375–421.

Starks, Laura T, and Sophia Yue Sun, 2016, Economic policy uncertainty, learning and incentives: Theory and evidence on mutual funds, *Unpublished working paper* .

Terpstra, T. J., 1952, The asymptotic normality and consistency of Kendall's test against trend, when ties are present in one ranking, *Indagationes Mathematicae* 14, 327–333.

Yan, Xuemin (Sterling), 2008, Liquidity, Investment Style, and the Relation between Fund Size and Fund Performance, *The Journal of Financial and Quantitative Analysis* 43, 741–767.

Zhu, Min, 2018, Informative fund size, managerial skill, and investor rationality, *Journal of Financial Economics* 130, 114–134.

Appendix

Variable Definitions

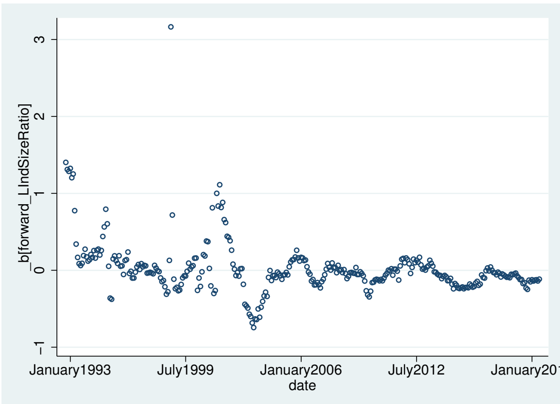
[Insert Table (A1) here]

Coefficient Correlation Matrix

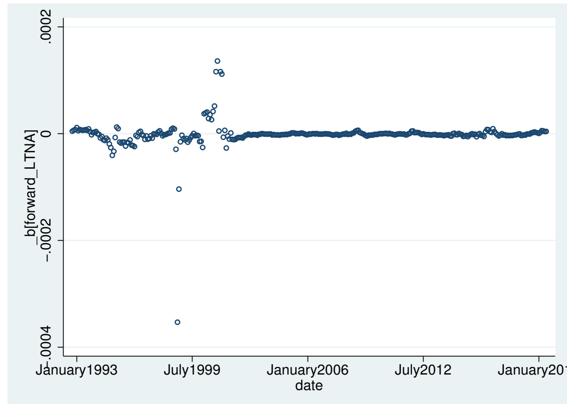
[Insert Table (A3) here]

Figure 1. Time-Varying DRS: Full Sample 3-year Rolling Window

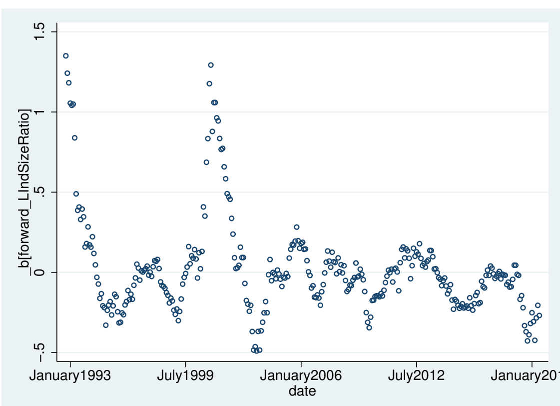
This figure plots the dynamics of DRS results using three-year rolling window. The first row shows the results of regression with fund size. The second row shows the results of regression with log fund size. The third row shows the average FE in the two regressions.



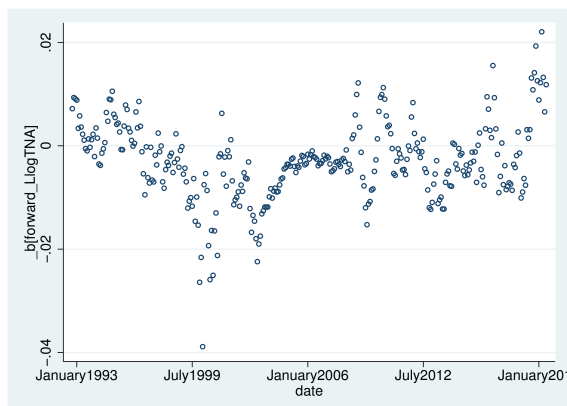
(a) b_t^S (level)



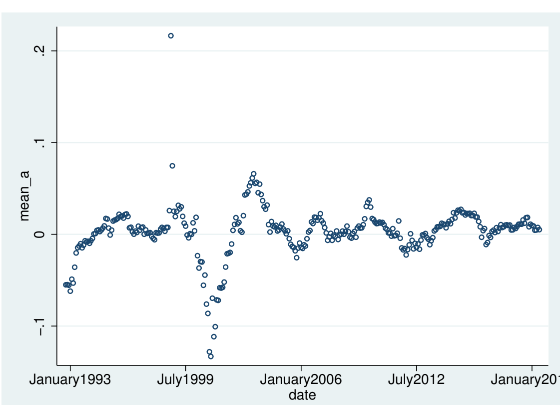
(b) b_t^q (level)



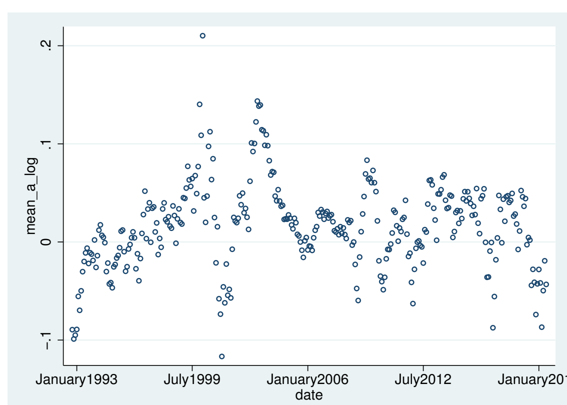
(c) b_t^S (log)



(d) b_t^q (log)



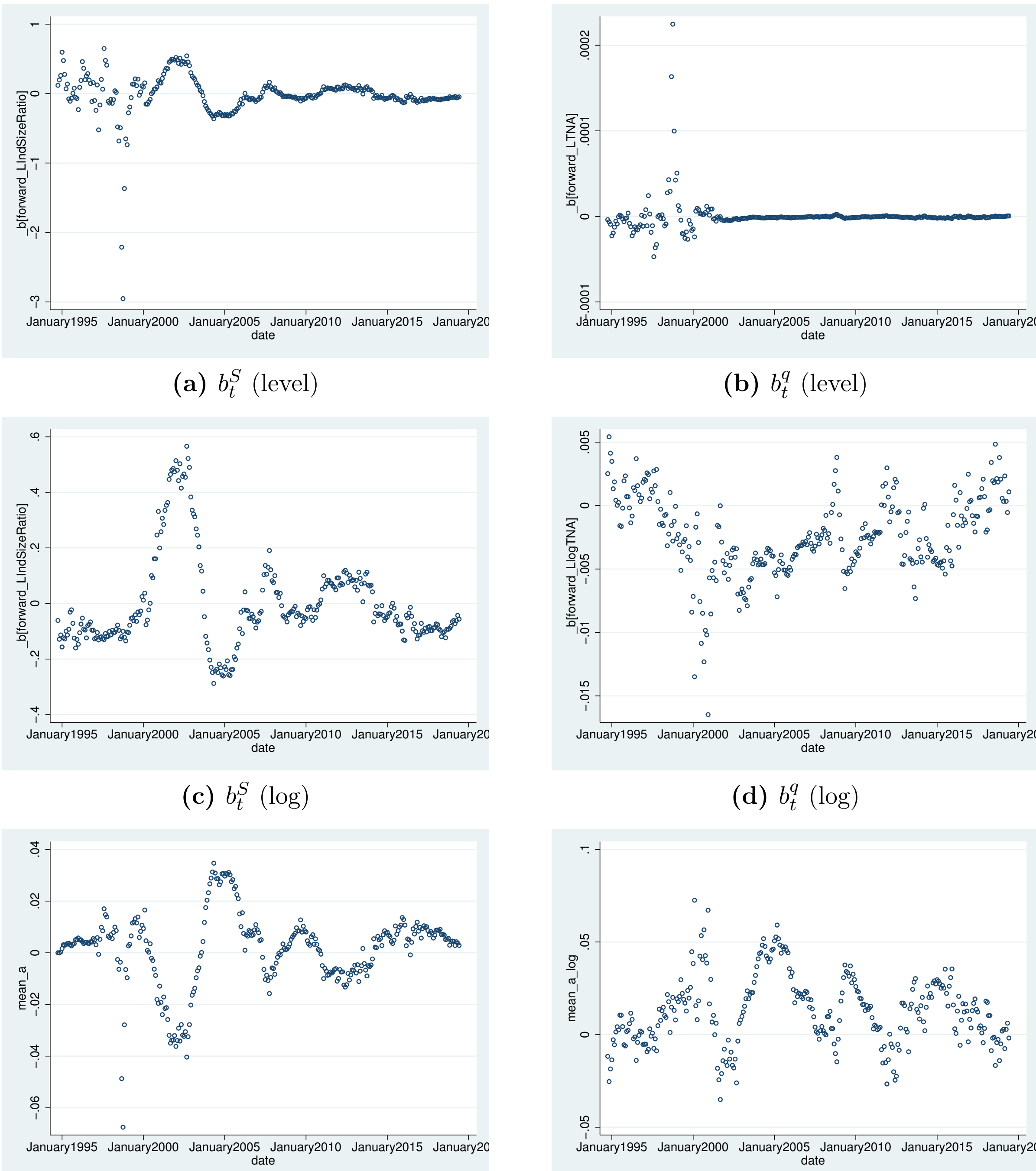
(e) average a_t (level)



(f) average a_t (log)

Figure 2. Time-Varying DRS: Full Sample 5-year Rolling Window

This figure plots the dynamics of DRS results using five-year rolling window. The first row shows the results of regression with fund size. The second row shows the results of regression with log fund size. The third row shows the average FE in the two regressions.



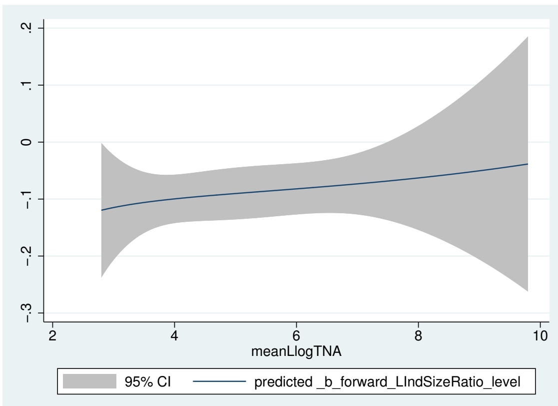
(a) b_t^S (level) **(b)** b_t^q (level)

(c) b_t^S (log) **(d)** b_t^q (log)

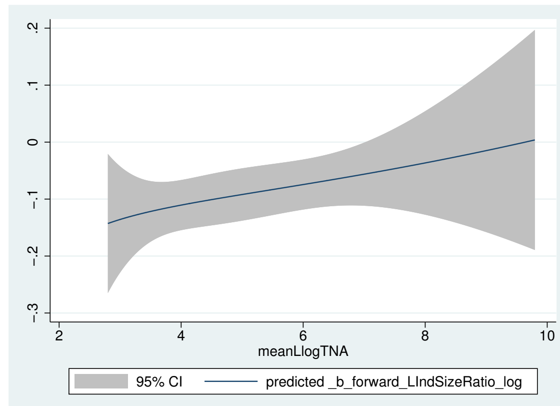
(e) average a_t (level) **(f)** average a_t (log)

Figure 3. Cross-Section DRS: Full Sample Fund Size

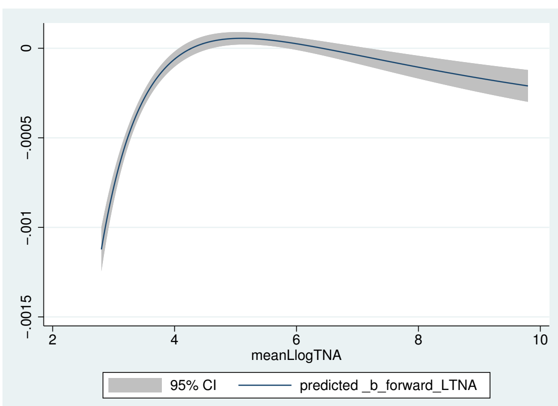
This figure plots the DRS coefficients against log fund size. We run fund by fund RD regression of benchmark adjusted returns against industry size and (log) fund size. For the resulting three coefficients (two regression coefficients and the constant), we winsorize at 1 and 99 percent. Then we plot the fractional polynomial fit over log fund size with 95 confidence interval. The first row is industrial sensitivity b_i^S. The second row is the fund size sensitivity b_i^q. The third row is the average intercepts of the two regressions.



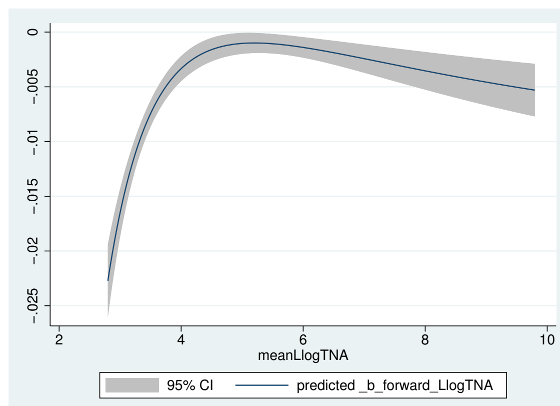
(a) b_i^S (level)



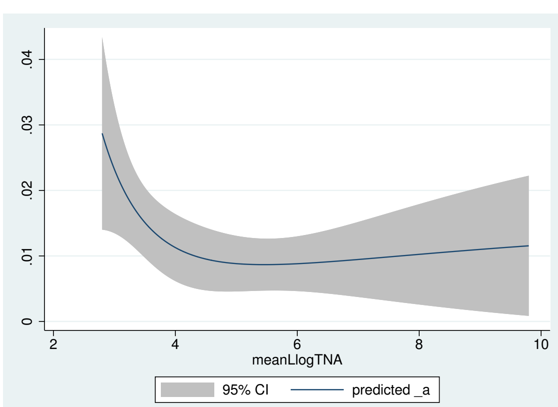
(b) b_i^S (log)



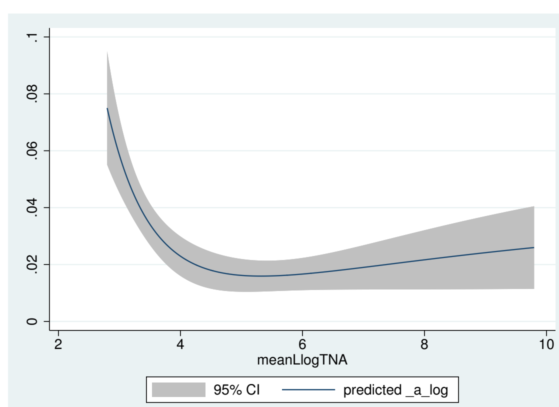
(c) b_i^q (level)



(d) b_i^q (log)



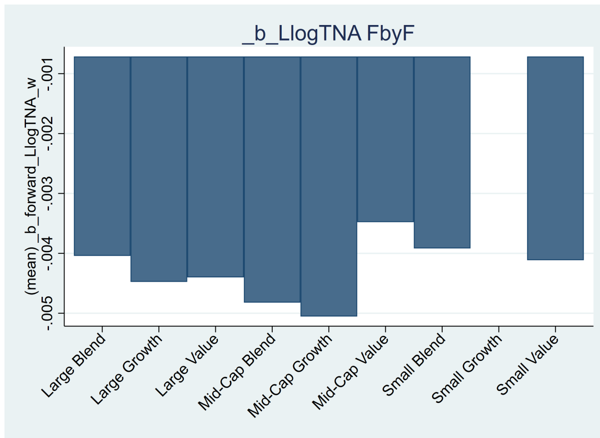
(e) Average a_i (level)



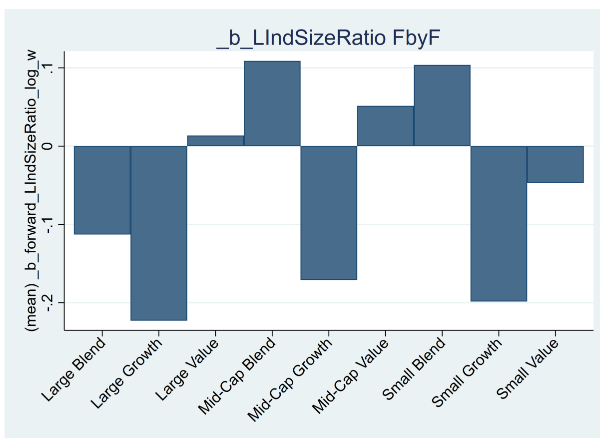
(f) Average a_i (log)

Figure 4. Sensitivity to Fund Size and Industry Scale by Category: Fund-by-Fund

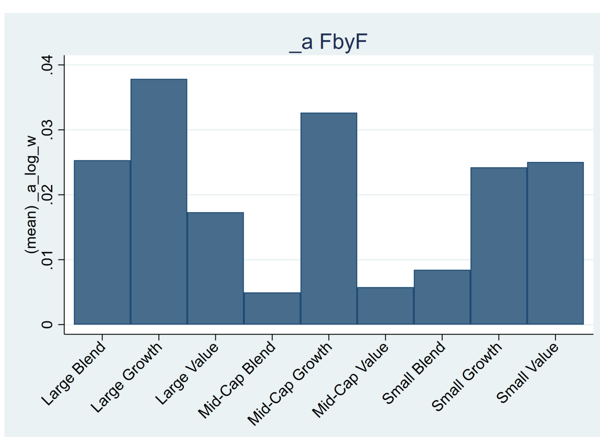
This figure graphs the mean Morningstar 3x3 category values of the coefficients and intercepts from 2SLS RD regressions. The coefficients for lagged log fund size b_i^q (top) and lagged industry size ratio b_i^S (middle), are estimated for each fund for the full sample period (the mean intercept values a_i are depicted in the bottom graph).



(a) $b_i^q(log)$



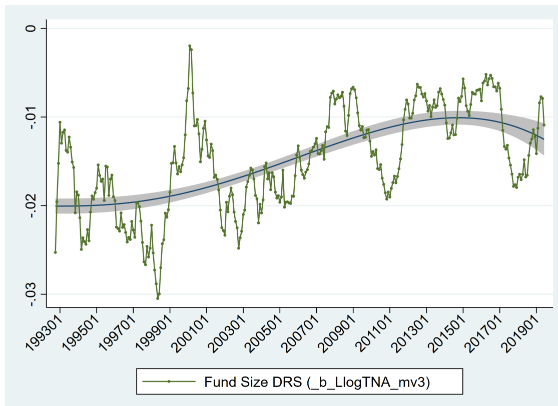
(b) $b_i^S(log)$



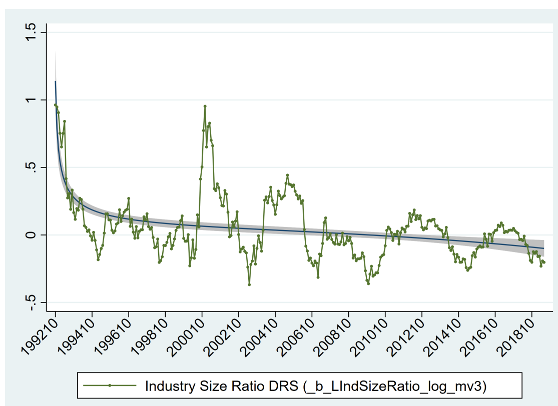
(c) a_i (log)

Figure 5. Time-varying Sensitivity to Scale

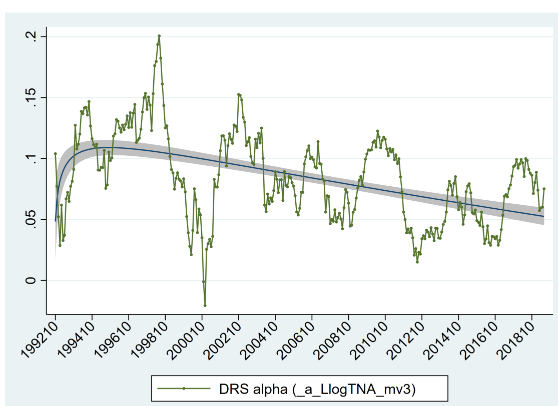
This figure depicts the mean monthly values of the coefficients from fund-by-fund 2SLS RD regressions of benchmark-adjusted returns on lagged log fund size b^q_{it} (top) and lagged industry scale b^S_{it} (middle), estimated monthly over rolling 3 year windows (the intercept a_{it} is depicted in the bottom graph). The graphs include fitted values and 95% confidence intervals estimated using fractional polynomials.



(a) $b^q_{it}(log)$



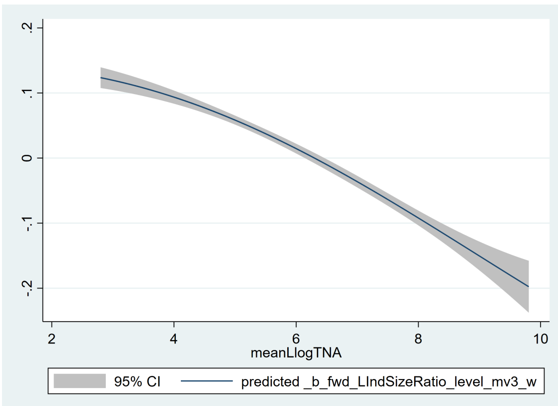
(b) $b^S_{it}(log)$



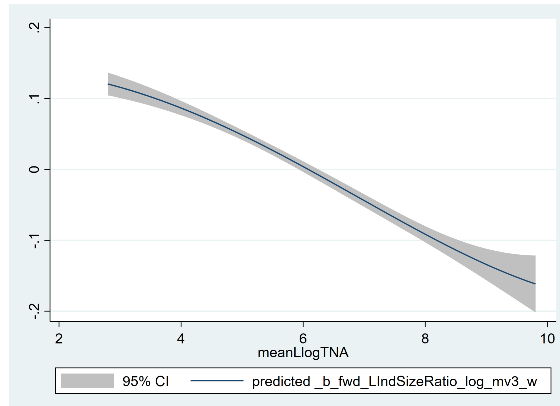
(c) $a^S_{it}(log)$

Figure 6. Cross-Section DRS: Fund-by-Fund 3-year Rolling Window Fund Size

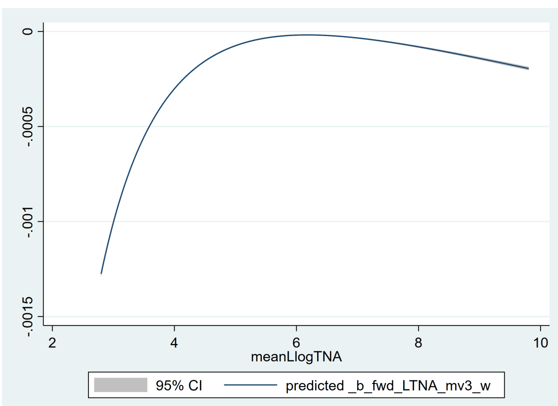
This figure plots the DRS coefficients against log fund size. We run fund by fund 2SLS RD regressions monthly using 3 year rolling windows. For the resulting three coefficients (two regression coefficients and the constant), we winsorize at 1 and 99 percent. Then we plot the fractional polynomial fit over log fund size with 95 confidence interval.



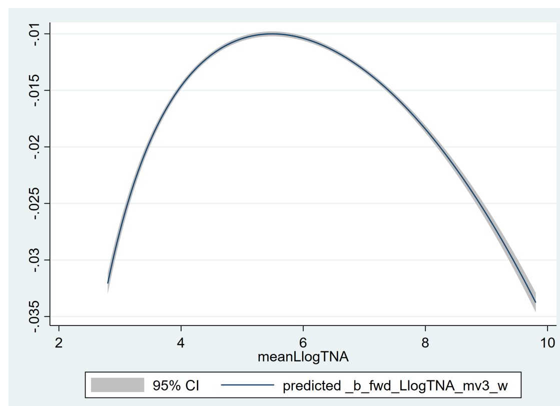
(a) b_{it}^S (level)



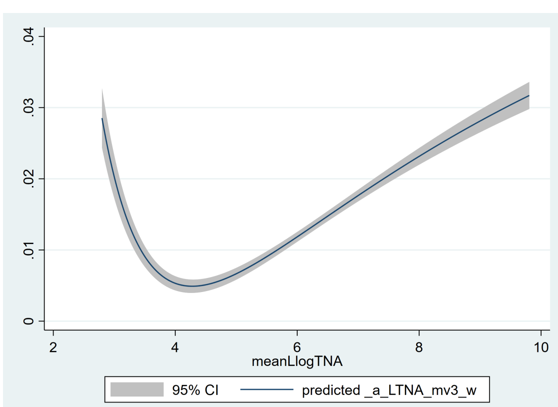
(b) b_{it}^S (log)



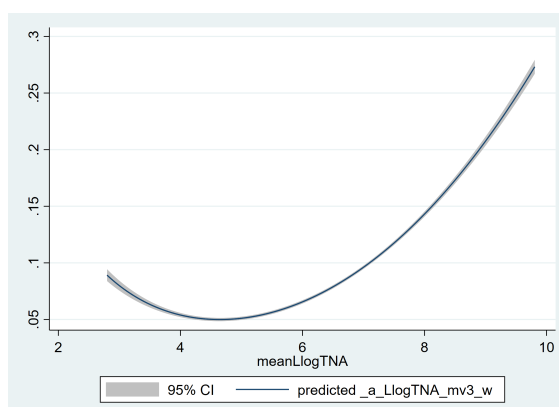
(c) b_{it}^q (level)



(d) b_{it}^q (log)



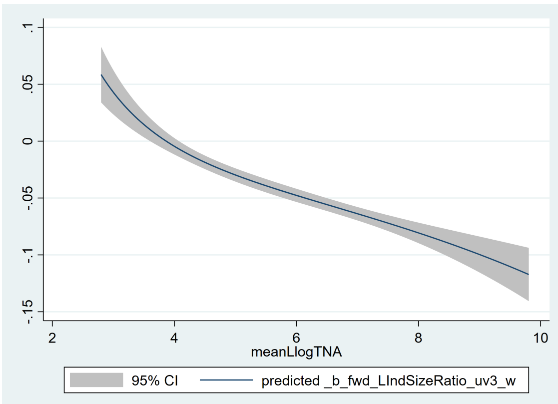
(e) Average a_{it} (level)



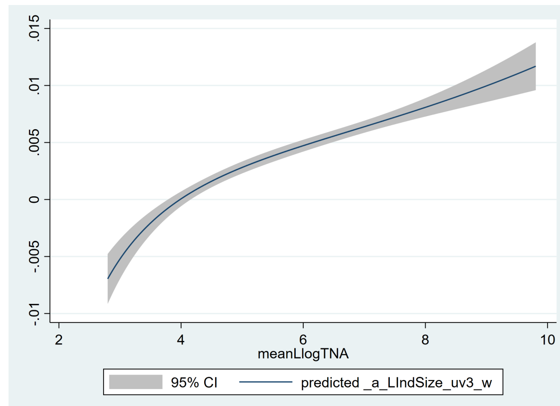
(f) Average a_{it} (log)

Figure 7. Cross-Section DRS: Univariate Fund-by-Fund 3-year Rolling Window Fund Size

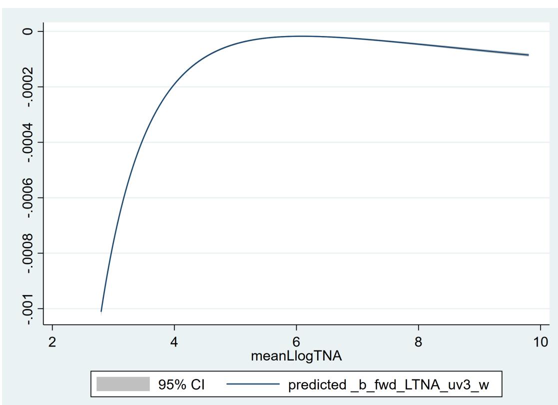
This figure plots the univariate DRS coefficients against log fund size. We run univariate fund-by-fund 2SLS RD regressions monthly using 3 year rolling windows. For the resulting coefficients (regression coefficients and the constant), we winsorize at 1 and 99 percent. Then we plot the fractional polynomial fit over log fund size with 95 confidence interval.



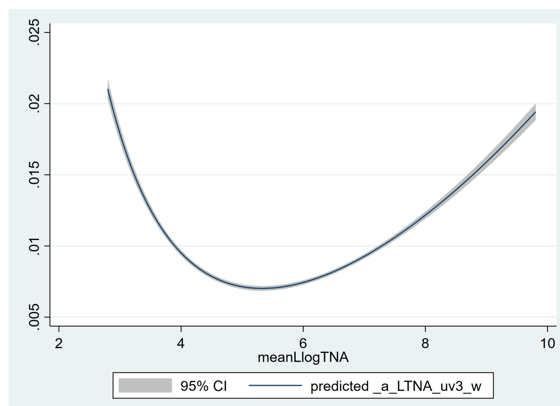
(a) $b^S_{UV,it}$



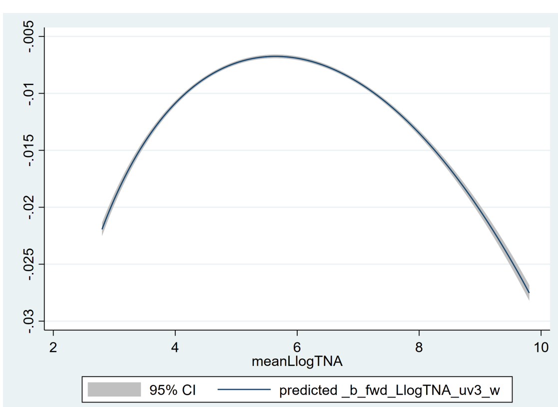
(b) Average $a_{UV,it}$ (Industry Scale)



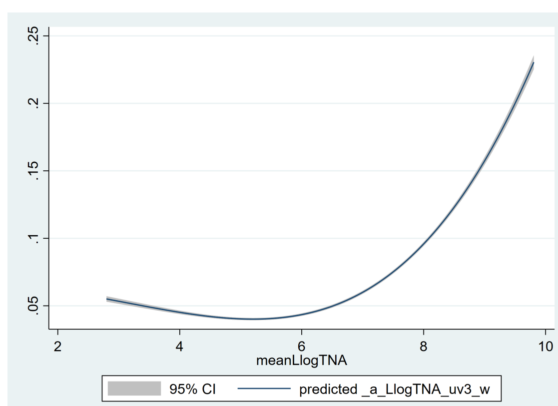
(c) $b^q_{UV,it}$ (level)



(d) Average $a_{UV,it}$ (Fund Size, level)



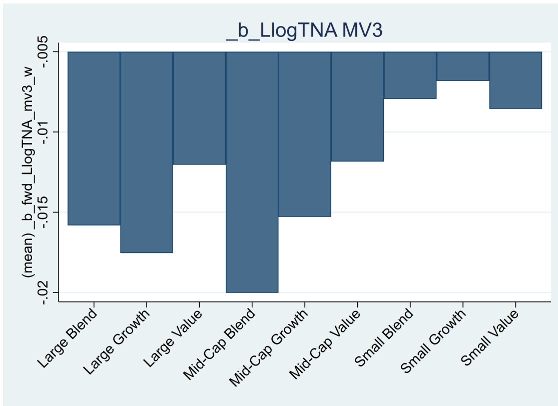
(e) $b^q_{UV,it}$ (log)



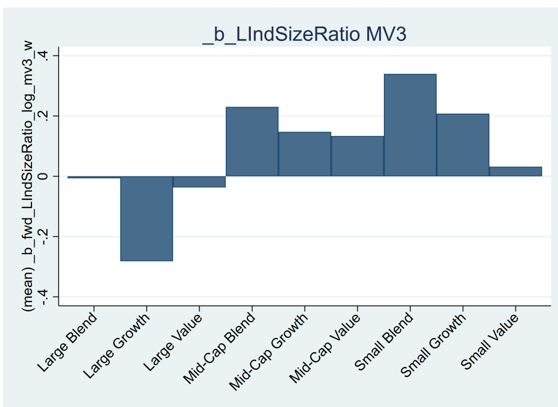
(f) Average $a_{UV,it}$ (Fund Size, log)

Figure 8. Sensitivity to Fund Size and Industry Scale by Category: Fund-by-Fund Rolling Window

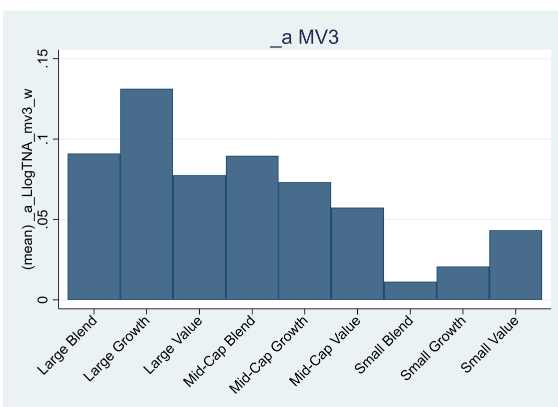
This figure graphs the mean Morningstar 3x3 category values of the coefficients and intercepts from 2SLS RD regressions. The coefficients for lagged log fund size b_{it}^{q} (top) and lagged industry size ratio b_{it}^{S} (middle), are estimated monthly for each fund over a rolling 3 year window (the mean intercept values a_{it} are depicted in the bottom graph).



(a) b_{it}^{q} (log)



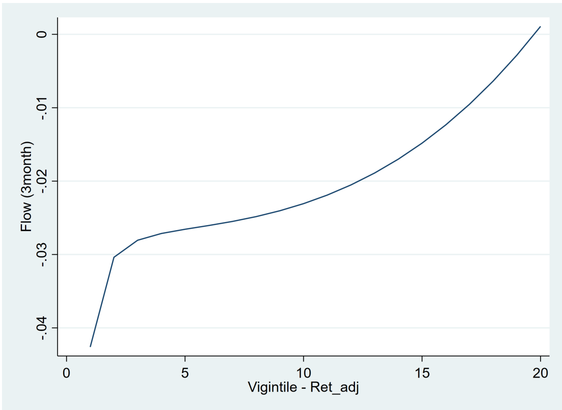
(b) b_{it}^{S}



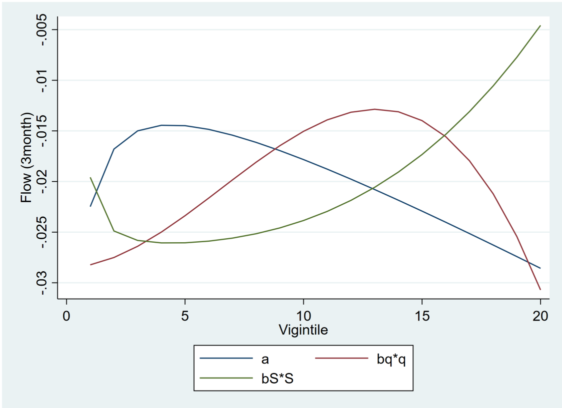
(c) a_{it}

Figure 9. Flows and Sensitivity to Fund Size and Industry Scale

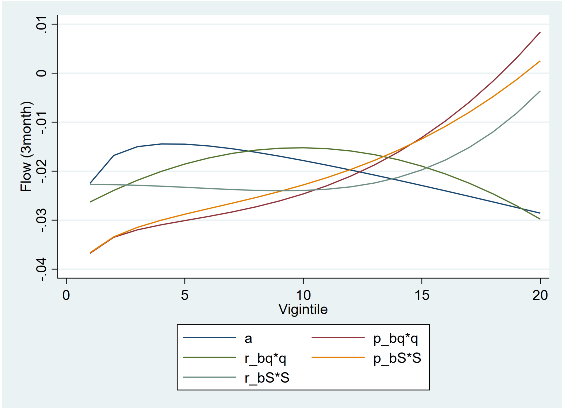
This figure graphs 3-month flows against funds' decomposed performance vigintile rank one month prior. Funds are ranked into 20 vigintiles each month by their benchmark-adjusted return (Ret_adj); a_{it} coefficient; raw, predicted and residual industry scale coefficient × industry scale ($b^S_{it} * S, b^S_{p,it} * S, b^S_{r,it} * S$); raw, fitted, and residual fund size coefficient × (log) fund size ($b^q_{it} * q, b^q_{p,it} * q, b^q_{r,it} * q$). Coefficients are estimated from 2SLS RD regressions of benchmark-adjusted return on industry scale and fund size, estimated monthly using 3-year rolling windows. Fitted coefficients and their residuals are estimated from regressions on fund, family and market characteristics.



(a) Flows vs Benchmark-adjusted return rank



(b) Flows vs raw $a_{it}, b^S_{it} * S, b^q_{it} * q$ rank



(c) Flows vs a_{it}, $b^S_{p,it} * S, bq_{p,it} * q$, $bS_r * S, and b_{r,it} q * q$ rank.

Table I OLS Regression of DRS: Full Sample

This table shows the OLS regression of benchmark adjusted fund returns on fund size, and industry size.

	(1)	(2)	(3)	(4)	(5)
	Ret_adj	Ret_adj	Ret_adj	Ret_adj	Ret_adj
b^S	-0.0151***		-0.0151***		-0.0146***
	(0.0026)		(0.0027)		(0.0027)
b^q (level)		0.0000	0.0000		
		(0.0000)	(0.0000)		
b^q (log)				0.0001***	0.0001***
				(0.0000)	(0.0000)
a	0.0007***	-0.0007***	0.0007***	-0.0010***	0.0004
	(0.0002)	(0.0000)	(0.0002)	(0.0001)	(0.0003)
Observations	361079	361079	361079	361079	361079

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

(a) Raw benchmark adjusted returns

	(1)	(2)	(3)	(4)	(5)
	Ret_adj	Ret_adj	Ret_adj	Ret_adj	Ret_adj
b^S	-0.0120***		-0.0119***		-0.0115***
	(0.0023)		(0.0023)		(0.0023)
b^q (level)		0.0000	0.0000		
		(0.0000)	(0.0000)		
b^q (log)				0.0001***	0.0001***
				(0.0000)	(0.0000)
Constant	0.0004*	-0.0007***	0.0004*	-0.0011***	-0.0000
	(0.0002)	(0.0000)	(0.0002)	(0.0001)	(0.0002)
Observations	361079	361079	361079	361079	361079

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

(b) Winsorized returns

Table II Panel Regression and 2SLS RD Regression: Full Sample

This table shows the Panel regression and 2SLS RD regression of benchmark adjusted fund returns on fund size, and industry size with fund fixed effect clustering month and sector.

	(1)	(2)	(3)	(4)	(5)
	Ret_adj	Ret_adj	Ret_adj	Ret_adj	Ret_adj
b^S	-0.02904***		-0.02965***		-0.03013***
	(-2.5834)		(-2.6377)		(-2.6741)
b^q (level)		-4.511e-07***	-4.541e-07***		
		(-9.0036)	(-9.0864)		
b^q (log)				-0.001381***	-0.001387***
				(-12.132)	(-12.162)
a	0.001990*	-0.0001665	0.002586**	0.007254***	0.01008***
	(1.8438)	(-1.0269)	(2.4044)	(10.711)	(7.7069)
Observations	355523	355523	355523	355523	355523
Adjusted R^2	0.003	0.004	0.004	0.005	0.006

t statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

(a) Panel regression with fund fixed effect

	(1)	(2)	(3)	(4)
	FRet_adj	FRet_adj	FRet_adj	FRet_adj
b^q (level)	-2.210e-07***		-2.580e-07***	
	(-3.6473)		(-4.3760)	
b^q (log)		-0.0001681		-0.0003527**
		(-1.0081)		(-2.1601)
b^S			-0.02930***	-0.02923***
			(-7.9703)	(-8.0524)
Observations	355534	355534	355534	355534

t statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

(b) 2SLS RD

Table III Time-Varying F and J* tests: Full Sample

This table presents results of Anova and Jonckheere-Terpstra tests of size and scale coefficients across time periods. The dependent variables are coefficients estimated using a 2SLS RD regression of benchmark-adjusted returns on lagged (log) fund size and lagged industry size ratio. For the tests, the sample is divided into nine sequential groups of 3-year periods, and the coefficients are estimated for each period. The Anova F-statistic is significant if the coefficient means are different across time periods. The Jonckheere-Terpstra J* statistic is significant if the coefficient means are trending over time; the p-values that the trend is increasing or decreasing are given.

	F-stat	p-value	J*-stat	p-down	p-up
a(level)	2.646	0.006	3.095	0.999	0.001
b^S (level)	8.034	0.000	-6.974	0.000	1.000
b^q (level)	0.453	0.905	1.956	0.975	0.025
a(log)	14.514	0.000	1.523	0.936	0.064
b^S (log)	10.871	0.000	-4.496	0.000	1.000
b^q (log)	22.352	0.000	0.441	0.670	0.330

Table IV Cross-Sectional F and J* tests: Full Sample

This table presents results of Anova and Jonckheere-Terpstra tests of size and scale coefficients across size decile (Panel A) and category (Panel B). The dependent variables are coefficients estimated using a 2SLS regression of benchmark-adjusted returns on lagged (log) fund size and lagged industry size ratio, estimated once for each fund using the full sample period. For the size decile tests, funds are grouped from low to high into deciles by their mean TNA over the full sample period. For the category tests, funds are grouped by their Morningstar category. The Anova F-statistic is significant if the means are different across groups. The Jonckheere-Terpstra J* statistic is significant if the means are trending across sequential groups; the p-values that the trend is increasing or decreasing are given.

	F-stat	p-value	J*-stat	p-down	p-up
			Panel A - TNA decile		
a_i (level)	0.428	0.921	-0.581	0.281	0.719
b_i^S (level)	0.397	0.937	0.130	0.552	0.448
b_i^q (level)	29.493	0.000	15.927	1.000	0.000
a(log)	2.088	0.027	-0.867	0.193	0.807
b_i^S (log)	0.764	0.650	1.486	0.931	0.069
b_i^q (log)	14.437	0.000	4.886	1.000	0.000
			Panel B - Category		
a_i (level)	4.537	0.000			
b_i^S (level)	4.866	0.000			
b_i^q (level)	0.949	0.475			
a(log)	2.532	0.010			
b_i^S (log)	5.511	0.000			
b_i^q (log)	1.051	0.395			

Table V Summary Statistics

This table shows summary statistics.

	count	mean	sd	p1	p25	p50	p75	p99
Adjusted Ret	355,534	-0.0007	0.0194	-0.0544	-0.0091	-0.0008	0.0074	0.0563
TNA	355,534	1,186.7862	2,650.1632	17.3276	92.0905	296.6476	972.1053	18,153.4258
Prop Inst	355,534	0.2832	0.3772	0.0000	0.0000	0.0292	0.5798	1.0000
Industry Size	355,534	0.0927	0.0123	0.0443	0.0875	0.0967	0.1009	0.1074
Fund Age	355,295	10.7825	10.3800	0.0000	4.0000	8.0000	14.0000	57.0000
Expense Ratio	355,534	0.0966	0.0429	0.0000	0.0758	0.0958	0.1186	0.2121
Turnover	332,761	6.1773	5.2417	0.2175	2.5833	4.7517	8.0833	27.9167
Managers	345,331	2.6969	2.6368	1.0000	1.0000	2.0000	3.0000	14.0000
Funds per Fam	355,534	16.0584	17.2561	1.0000	3.0000	9.0000	22.0000	77.0000
Family TNA	355,534	69,691.7021	167,920.6674	41.4935	1,544.8826	12,909.6455	45,240.0508	897,762.7500
Active Funds	355,534	1,371.3282	323.4999	330.0000	1,244.0000	1,484.0000	1,571.0000	1,694.0000
Active Fam	355,534	344.2443	48.5197	153.0000	340.0000	361.0000	366.0000	394.0000
Benchmark Ret	355,534	0.0080	0.0490	-0.1431	-0.0173	0.0131	0.0370	0.1132
RF	355,534	0.1558	0.1651	0.0000	0.0100	0.1000	0.3100	0.5100
Mkt-RF	355,534	0.6335	4.3169	-10.3500	-1.8800	1.1700	3.3700	9.5400
HML	355,534	0.0665	2.9030	-8.3200	-1.6500	-0.0900	1.5400	8.2100
SMB	355,534	0.1683	2.9030	-5.7100	-1.7100	0.1600	2.0700	7.0400
MOM	355,534	0.2831	4.9453	-12.4900	-1.4300	0.4500	2.9300	12.7500

(a) Original Variables

	mean	sd	p1	p25	p50	p75	p99	count
a (level)	0.0124	0.1832	-0.6416	-0.0667	0.0073	0.0871	0.7175	355,534
b^S (level)	0.0183	1.7886	-5.9699	-0.7990	0.0136	0.8217	6.1951	355,534
b^q (level)	-0.1332	0.7968	-5.2572	-0.0964	-0.0101	0.0145	2.6243	355,534
a (log)	0.0785	0.3494	-0.9214	-0.0730	0.0356	0.1755	1.7149	355,534
b^S (log)	0.0114	1.7997	-6.0849	-0.8062	0.0103	0.8212	6.2039	355,534
b^q (log)	-0.0136	0.0511	-0.2639	-0.0252	-0.0058	0.0067	0.1335	355,534

(b) Estimated sensitivities

Table VI 2SLS RD Regressions: 3-year Rolling Window

This table gives the mean size and scale coefficients estimated fund-by-fund over 3-year rolling windows using 2SLS RD regressions. Columns (1-3) give univariate regression results, and columns (4-5) give multiple regressions results.

	(1)	(2)	(3)	(4)	(5)
	Ret_adj	Ret_adj	Ret_adj	Ret_adj	Ret_adj
			Panel A - Equal-weight		
b^q_{it} (level)	-8.91e-05***			-0.000133***	
	(-89.95)			(-99.71)	
b^q_{it} (log)		-0.00969***			-0.0136***
		(-148.8)			(-159.1)
b^S_{it}			-0.0401***	0.0183***	0.0114***
			(-17.04)	(6.100)	(3.783)
a_{it}	0.00921***	0.0557***	0.00385***	0.0124***	0.0785***
	(141.0)	(144.5)	(18.33)	(40.36)	(134.0)
Observations	355,534	355,534	355,534	355,534	355,534
			Panel B - Value-weight		
b^q_{it} (level)	-0.000162***			-0.000287***	
	(-11.13)			(-14.40)	
b^q_{it} (log)		-0.0223***			-0.0288***
		(-28.20)			(-28.89)
b^S_{it}			-0.0731***	-0.0673***	-0.0421***
			(-11.36)	(-6.162)	(-3.425)
a_{it}	0.0163***	0.156***	0.00694***	0.0208***	0.203***
	(21.33)	(32.15)	(13.13)	(6.240)	(30.55)
Observations	355,534	355,534	355,534	355,534	355,534

Table VII Time-Varying F and J* tests: Fund-by-fund Rolling Window

This table presents results of Anova and Jonckheere-Terpstra tests of size and scale coefficients across time periods. The dependent variables are coefficients estimated using fund-by-fund 2SLS RD regressions over rolling 3-year windows. For the tests, the sample is divided into nine sequential (non-overlapping) 3-year periods. The Anova F-statistic is significant if the coefficient means are different across time periods. The Jonckheere-Terpstra J* statistic is significant if the coefficient means are trending over time; the p-values that the trend is increasing or decreasing are given.

	F	Pr(>f)	J*	p(down)	p(up)
a_{it} (level)	10.429	0.000	-0.168	0.433	0.567
b_{it}^q (level)	3.063	0.004	10.215	1.000	0.000
b_{it}^S (level)	19.597	0.000	-4.636	0.000	1.000
a_{it} (log)	30.649	0.000	-8.462	0.000	1.000
b_{it}^q (log)	25.926	0.000	8.928	1.000	0.000
b_{it}^S (log)	20.347	0.000	-4.949	0.000	1.000

Table VIII Cross-Sectional F and J* tests: Fund-by-fund Rolling Window

This table presents results of Anova and Jonckheere-Terpstra tests of log size and scale coefficients across size deciles (Panel A) and categories (Panel B). Coefficients estimated monthly fund-by-fund using 2SLS RD regressions over 3-year rolling windows. For the size decile tests, funds are grouped from low to high into deciles by their mean TNA over the full sample period. For the category tests, funds are grouped by their Morningstar category. The Anova F-statistic is significant if the means are different across groups. The Jonckheere-Terpstra J* statistic is significant if the means are trending across sequential groups; the p-values that the trend is increasing or decreasing are given.

	F-stat	p-value	J*-stat	p-down	p-up
			Panel A - TNA decile		
a_{it} (level)	4.041	0.003	4.738	1.000	0.000
b_{it}^q (level)	123.065	0.000	21.674	1.000	0.000
b_{it}^S (level)	5.640	0.000	-4.433	0.000	1.000
a_{it} (log)	40.629	0.000	11.463	1.000	0.000
b_{it}^q (log)	6.657	0.000	-0.467	0.320	0.680
b_{it}^S (log)	6.469	0.000	-4.499	0.000	1.000
			Panel B - Category		
a_{it} (level)	541.757	0.000			
b_{it}^q (level)	108.571	0.000			
b_{it}^S (level)	513.108	0.000			
a_{it} (log)	530.473	0.000			
b_{it}^q (log)	265.317	0.000			
b_{it}^S (log)	514.887	0.000			

Table IX Panel Regressions of Size and Scale Coefficients

This table presents results of a regression of log size and scale coefficients on Fama-French 5 factors, the Pastor-Stambaugh Liquidity factor, and a range of fund-level characteristics.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	b^S (level)	b^S (level)	b^S (level)	b^q (level)	b^q (level)	b^q (level)	a (level)
Fund Age	-0.1650***	-0.1626***	-0.07133***	0.00003072***	0.00002888***	0.000007661*	0.007766***
	(-13.067)	(-13.037)	(-5.6743)	(9.4251)	(8.2864)	(1.7501)	(6.0697)
Fund Turnover	-0.007867***	-0.007636***	-0.008708***	-0.000001260***	-0.000001376***	-0.000001352***	0.001292***
	(-3.8745)	(-3.7658)	(-4.5910)	(-2.9411)	(-3.2231)	(-3.0871)	(7.0299)
Fund Expense	-1.2284***	-1.1635***	-0.8664***	-0.00008174	-0.0001280*	-0.0002394**	0.1216***
	(-4.7577)	(-4.7181)	(-3.1270)	(-1.1720)	(-1.6617)	(-2.2714)	(4.7813)
N Manager	-0.06367***	-0.05849***	-0.02696***	0.00001460***	0.00001153***	0.000001093	0.003213***
	(-7.1222)	(-6.6899)	(-3.1696)	(4.8828)	(3.8505)	(0.3518)	(3.7841)
Inst. Proportion	-0.2511***	-0.2592***	-0.1742***	0.00007103***	0.00007472***	0.00006289***	0.007088**
	(-8.7102)	(-8.8841)	(-6.5569)	(6.3668)	(6.7554)	(5.5741)	(2.5717)
Fund Size	-0.4560***	-0.4124***	-0.3789***	0.0006235***	0.0005996***	0.0005932***	0.02935***
	(-16.415)	(-14.873)	(-13.969)	(28.646)	(27.499)	(27.197)	(7.8796)
Sqr Fund Size	0.02375***	0.02253***	0.02013***	-0.00004261***	-0.00004192***	-0.00004148***	-0.001307***
	(12.196)	(11.672)	(10.704)	(-26.589)	(-26.151)	(-25.915)	(-4.7161)
Family NFund		0.01735	0.09807***		-0.000005848	-0.00003642***	-0.007392***
		(0.9223)	(6.0035)		(-0.9722)	(-5.9669)	(-4.5504)
Family TNA		-0.07474***	-0.08246***		0.00003984***	0.00004055***	0.006093***
		(-8.5246)	(-9.4667)		(10.637)	(10.784)	(6.9889)
Active Funds			-0.04858			0.0003238***	-0.009882
			(-0.2628)			(10.806)	(-0.5593)
Active Families			-0.1035			-0.0005166***	0.02089
			(-0.3192)			(-9.1705)	(0.6837)
LIQv			4.5836			0.001146**	-0.9407***
			(1.5937)			(2.4111)	(-3.4694)
market premium			0.05478**			-0.000004509	-0.007240***
			(2.3813)			(-1.3072)	(-3.2226)
riskfree rate			0.3259**			-0.0001215***	0.002303
			(2.0557)			(-4.3848)	(0.1561)
HML			-0.1386***			0.000001404	0.01558***
			(-3.4842)			(0.2798)	(4.1963)
SMB			0.07548*			-0.00006200***	0.003144
			(1.9488)			(-9.1840)	(0.8870)
RMW			-0.1906***			-0.00003044***	0.01632***
			(-4.1077)			(-4.7584)	(3.9123)
CMA			0.2781***			0.00002520***	-0.02921***
			(5.6762)			(3.8846)	(-6.3504)
Constant	2.5299***	3.9633***	4.4486***	-0.002349***	-0.003116***	-0.002260***	-0.3272***
	(21.359)	(18.959)	(6.3616)	(-31.249)	(-29.281)	(-13.413)	(-5.0094)
Fund FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	352686	352686	352057	352686	352686	352057	352057
Adjusted R^2	0.197	0.198	0.203	0.256	0.256	0.257	0.194

t statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table X Panel Regressions of Log Fund Size Coefficients

This table presents results of a regression of log size and scale coefficients on Fama-French 5 factors, the Pastor-Stambaugh Liquidity factor, and a range of fund-level characteristics.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	b^S (log)	b^S (log)	b^S (log)	b^q (log)	b^q (log)	b^q (log)	a (log)
Fund Age	-0.1450***	-0.1437***	-0.05337***	0.0008583***	0.0003983	-0.002710***	0.02068***
	(-11.554)	(-11.580)	(-4.2278)	(3.3733)	(1.5461)	(-8.1352)	(9.0596)
Fund Turnover	-0.008249***	-0.008020***	-0.008607***	-0.0003238***	-0.0003177***	-0.0002865***	0.002790***
	(-4.0828)	(-3.9740)	(-4.5543)	(-10.255)	(-10.072)	(-9.3236)	(10.062)
Fund Expense	-1.1370***	-1.0938***	-0.7450***	-0.009332**	-0.01860***	-0.04571***	0.3511***
	(-4.3860)	(-4.4063)	(-2.6786)	(-1.9720)	(-3.8828)	(-8.0296)	(9.2232)
N Manager	-0.07316***	-0.06872***	-0.03720***	0.002077***	0.001849***	0.0006906***	-0.001754
	(-8.1044)	(-7.7661)	(-4.2866)	(8.2503)	(7.3733)	(2.7303)	(-1.0260)
Inst. Proportion	-0.2408***	-0.2493***	-0.1606***	0.008598***	0.008094***	0.005140***	-0.02269***
	(-8.4626)	(-8.6848)	(-6.0897)	(11.020)	(10.466)	(6.7367)	(-4.4844)
Fund Size	-0.4938***	-0.4534***	-0.4223***	0.01789***	0.01754***	0.01673***	-0.05452***
	(-17.740)	(-16.230)	(-15.366)	(16.693)	(16.270)	(15.476)	(-7.2791)
Sqr Fund Size	0.02632***	0.02522***	0.02302***	-0.001546***	-0.001519***	-0.001468***	0.007027***
	(13.490)	(12.981)	(12.088)	(-18.860)	(-18.546)	(-17.856)	(12.256)
Family NFund		0.02172	0.1025***		0.002804***	-0.0005510	-0.01276***
		(1.1546)	(6.3632)		(6.4300)	(-1.2083)	(-4.2105)
Family TNA		-0.07080***	-0.07827***		-0.0002437	-0.0001912	0.01294***
		(-8.0638)	(-8.9673)		(-1.0257)	(-0.8021)	(7.7508)
Active Funds			-0.04758			0.02247***	-0.09949***
			(-0.2566)			(8.4404)	(-4.3174)
Active Families			-0.1655			-0.03152***	0.1576***
			(-0.5085)			(-6.0323)	(3.6766)
LIQv			6.8389**			0.2358***	-2.5392***
			(2.3956)			(5.5027)	(-6.7883)
market premium			0.05474**			-0.0001463	-0.006962**
			(2.3774)			(-0.4881)	(-2.3375)
riskfree rate			0.1800			-0.01884***	0.1196***
			(1.1511)			(-8.0521)	(6.1452)
HML			-0.1170***			-0.001989***	0.02498***
			(-2.9606)			(-4.4298)	(5.6635)
SMB			0.06582*			-0.008672***	0.04841***
			(1.7191)			(-14.533)	(10.066)
RMW			-0.1662***			-0.004038***	0.03398***
			(-3.6260)			(-7.7180)	(6.5743)
CMA			0.2434***			0.005710***	-0.05773***
			(5.0029)			(10.754)	(-10.522)
Constant	2.6096***	3.9638***	4.8085***	-0.06346***	-0.06068***	-0.01474	-0.4551***
	(22.198)	(19.017)	(6.8932)	(-17.579)	(-10.577)	(-1.0171)	(-4.1275)
Fund FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	352686	352686	352057	352686	352686	352057	352057
Adjusted R^2	0.198	0.199	0.204	0.153	0.153	0.158	0.176

t statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table XI Persistence of Rolling Window Coefficients

This table presents a markov transition matrix of adjusted coefficient and adjusted benchmark-adjusted return persistence over consecutive 1-year and 3-year periods. Coefficients are estimated monthly fund-by-fund using 2SLS RD regressions over 3-year rolling windows. Coefficients and benchmark-adjusted returns are adjusted for fund, family and market characteristics. Funds are grouped by their mean monthly coefficient into deciles for each period. The probability that the fund transitions from a decile in one period to a decile in the next period is estimated.

	1-year				3-year			
	to-1	to-2	to-9	to-10	to-1	to-2	to-9	to-10
	Residual Benchmark-adjusted 3-year gross return							
from-1	0.546	0.198	0.015	0.011	0.265	0.108	0.040	0.105
from-2	0.237	0.272	0.012	0.013	0.148	0.140	0.060	0.063
from-9	0.011	0.017	0.265	0.210	0.070	0.067	0.139	0.144
from-10	0.014	0.012	0.196	0.505	0.102	0.078	0.180	0.201
	a_r (log)							
from-1	0.324	0.173	0.048	0.082	0.204	0.099	0.082	0.128
from-2	0.162	0.204	0.044	0.045	0.080	0.120	0.112	0.090
from-9	0.065	0.044	0.216	0.191	0.115	0.108	0.067	0.115
from-10	0.110	0.050	0.157	0.375	0.118	0.085	0.126	0.156
	b_r^q (log)							
from-1	0.376	0.155	0.053	0.143	0.179	0.123	0.064	0.149
from-2	0.198	0.188	0.066	0.082	0.132	0.130	0.095	0.103
from-9	0.046	0.063	0.192	0.130	0.070	0.073	0.145	0.091
from-10	0.086	0.057	0.166	0.291	0.130	0.091	0.099	0.208
	b_r^S (log)							
from-1	0.386	0.184	0.040	0.042	0.166	0.104	0.090	0.166
from-2	0.179	0.200	0.039	0.036	0.078	0.111	0.109	0.098
from-9	0.030	0.044	0.212	0.171	0.119	0.085	0.101	0.090
from-10	0.044	0.044	0.181	0.372	0.178	0.120	0.109	0.122
	23745 observations				5209 observations			

Table XII Fund Flows - Size and Scale Coefficients

Panel A presents the results of a regression of 3-month fund flows on lagged log size and scale coefficients. Coefficients are estimated monthly fund-by-fund using 2SLS RD regressions over rolling 3-year windows. In Panel B, the independent variables are the size and scale coefficients adjusted for macro-level and fund-level variables. In columns (5-8), control variables include gross return for the prior month and the prior 12 months and the prior 24 months, prior month log TNA, prior month expense ratio, prior month flow, and log fund age. Fixed effects for Morningstar 3x3 category × month are included, and t-statistics estimated using robust standard errors clustered by fund and month are given in brackets. ***,**, and * denote significance at the 1%, 5% and 10% level, respectively.

	Flow							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
				Panel A - Coefficients				
L.a_{it}	-0.00995***			-0.00666	-0.00254***			-0.00175
	(-6.531)			(-1.420)	(-2.705)			(-0.470)
L.b^q_{it}		0.0129*		-0.0179		-0.000257		-0.00842
		(1.679)		(-0.586)		(-0.0472)		(-0.344)
L.b^S_{it}			0.00324***	0.00265***			0.00109***	0.000927**
			(7.484)	(4.743)			(3.717)	(2.343)
Category x Month Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	No	No	No	Yes	Yes	Yes	Yes
Constant	-0.0194***	-0.0200***	-0.0202***	-0.0199***	0.00833***	0.00873***	0.00844***	0.00827***
	(-32.56)	(-33.48)	(-36.18)	(-33.37)	(2.755)	(2.888)	(2.794)	(2.708)
Observations	337,252	337,252	337,252	337,252	304,898	304,898	304,898	304,898
R-squared	0.344	0.343	0.345	0.345	0.465	0.465	0.465	0.465
				Panel B - Residual Coefficients				
L.$a_{r,it}$	-0.00165			0.0179***	-0.000244			0.00596
	(-1.075)			(3.469)	(-0.260)			(1.601)
L.$b^q_{r,it}$		-0.0218***		0.0897***		-0.00955*		0.0277
		(-2.842)		(2.756)		(-1.743)		(1.132)
L.$b^S_{r,it}$			0.00245***	0.00411***			0.000857***	0.00140***
			(5.533)	(7.042)			(2.893)	(3.501)
Category x Month Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	No	No	No	Yes	Yes	Yes	Yes
Constant	-0.0208***	-0.0208***	-0.0208***	-0.0208***	0.00859***	0.00845***	0.00916***	0.00906***
	(-36.95)	(-36.95)	(-37.18)	(-37.19)	(2.832)	(2.790)	(3.024)	(2.985)
Observations	333,933	333,933	333,933	333,933	304,277	304,277	304,277	304,277
R-squared	0.346	0.346	0.347	0.347	0.465	0.465	0.465	0.465

Table XIII Fund Flows - Size and Scale Components of Returns

Panel A presents the results of a regression of 3-month fund flows on lagged log size and scale coefficients multiplied by their respective independent variables ($b^S_{it} * S_{t-1}$ and $b^q_{it} * q_{it-1}$). Coefficients are estimated monthly fund-by-fund using 2SLS RD regressions over rolling 3-year windows. In Panel B, the size and scale coefficients are adjusted for macro-level and fund-level variables. In columns (5-8), control variables include gross return for the prior month and the prior 12 months and the prior 24 months, prior month log TNA, prior month expense ratio, prior month flow, and log fund age. Fixed effects for Morningstar 3x3 category $\times$ month are included, and t-statistics estimated using robust standard errors clustered by fund and month are given in brackets. ***,**, and * denote significance at the 1%, 5% and 10% level, respectively.

	Flow							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Panel A - Coefficients x Independent Vars							
L.a_{it}	-0.00995***			0.0383***	-0.00254***			0.0180**
	(-6.531)			(3.292)	(-2.705)			(2.423)
L.$b^q_{it} * q_{it-1}$		0.00432***		0.0450***		0.000557		0.0195**
		(3.463)		(3.691)		(0.649)		(2.473)
L.$b^S_{it} * S_{it-1}$			0.0346***	0.0755***			0.0122***	0.0311***
			(7.298)	(6.236)			(3.844)	(4.114)
Category x Month Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	No	No	No	Yes	Yes	Yes	Yes
Constant	-0.0194***	-0.0198***	-0.0202***	-0.0197***	0.00833***	0.00866***	0.00842***	0.00824***
	(-32.56)	(-33.24)	(-36.11)	(-32.90)	(2.755)	(2.862)	(2.790)	(2.726)
Observations	337,252	337,252	337,252	337,252	304,898	304,898	304,898	304,898
R-squared	0.344	0.343	0.345	0.346	0.465	0.465	0.465	0.465
	Panel B - Residual Coefficients x Independent Vars							
L.$a_{r,it}$	-0.00165			0.0465***	-0.000244			0.0210***
	(-1.075)			(6.427)	(-0.260)			(4.747)
L.$b^q_{r,it} * q_{it-1}$		-0.00153		0.0405***		-0.000725		0.0180***
		(-1.530)		(6.222)		(-1.094)		(4.474)
L.$b^S_{r,it} * S_{it-1}$			0.0255***	0.0734***			0.00985***	0.0313***
			(5.270)	(8.989)			(3.076)	(6.139)
Category x Month Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	No	No	No	Yes	Yes	Yes	Yes
Constant	-0.0208***	-0.0208***	-0.0208***	-0.0210***	0.00859***	0.00853***	0.00919***	0.00838***
	(-36.95)	(-36.98)	(-37.17)	(-37.65)	(2.832)	(2.817)	(3.034)	(2.751)
Observations	333,933	333,933	333,933	333,933	304,277	304,277	304,277	304,277
R-squared	0.346	0.346	0.347	0.348	0.465	0.465	0.465	0.465

Table XIV Fund Flows and Fitted and Residual Components of Returns

Columns (1) and (3) present the results of regressions of 3-month fund flows on lagged benchmark-adjusted return components. In Columns (2) and (4), flows are regressed on fitted and residual log size and scale coefficients multiplied by their respective independent variables ($b^S_{it} * S_{t-1}$ and $b^q_{it} * q_{it-1}$). Coefficients are estimated monthly fund-by-fund using 2SLS RD regressions over rolling 3-year windows. In columns (2) and (4), control variables include gross return for the prior month and the prior 12 months and the prior 24 months, prior month log TNA, prior month expense ratio, prior month flow, and log fund age. Fixed effects for Morningstar 3x3 category × month are included, and t-statistics estimated using robust standard errors clustered by fund and month are given in brackets. ***,**, and * denote significance at the 1%, 5% and 10% level, respectively.

	Flow			
	(1)	(2)	(3)	(4)
L.Ret_adj	0.605***		0.326***	
	(13.85)		(8.311)	
L.a_{it}		0.00878		0.00476
		(1.487)		(1.141)
L.$b^q_{p,it} * q_{it-1}$		0.123***		0.244***
		(6.096)		(11.83)
L.$b^S_{p,it} * S_{it-1}$		0.224***		0.530***
		(5.135)		(12.02)
L.$b^q_{r,it} * q_{it-1}$		0.00806		0.00345
		(1.530)		(0.933)
L.$b^S_{r,it} * S_{it-1}$		0.0407***		0.0127***
		(5.720)		(2.700)
Cat x Month FE	Yes	Yes	Yes	Yes
Controls	No	No	Yes	Yes
Constant	-0.0198***	-0.0108***	0.0113***	-0.0853***
	(-35.30)	(-5.904)	(3.752)	(-13.58)
Observations	337,252	333,933	304,898	304,277
R-squared	0.350	0.354	0.465	0.469

Table XV Flows with Linear, Quadratic and Cubic Return Components

This table gives the results of regressions of 3-month flows on linear, quadratic and cubic vigintile ranks of components of returns, lagged 1 month. There are 3 separate regressions for each variable. Each regression includes gross return for the prior month and the prior 12 months and the prior 24 months, prior month log TNA, prior month expense ratio, prior month flow, and log fund age as controls. Fixed effects for Morningstar 3x3 category × month are included, and t-statistics estimated using robust standard errors clustered by fund and month are given in brackets. ***,**, and * denote significance at the 1%, 5% and 10% level, respectively.

		Flow	
VARIABLE	(1)	(2)	(3)
Ret_adj	0.000898***	0.000267	0.00348***
	(4.263)	(0.803)	(3.399)
Ret_adj^2		3.01e-05***	-0.000322***
		(2.609)	(-3.495)
Ret_adj^3			1.12e-05***
			(3.827)
$b^q q$	0.000125	0.00132***	-0.000897
	(1.649)	(4.060)	(-1.343)
$b^q q^2$		-5.67e-05***	0.000203***
		(-3.745)	(2.645)
$b^q q^3$			-8.27e-06***
			(-3.364)
$b^S S$	0.000639***	-0.00112***	-0.00107
	(6.120)	(-3.208)	(-1.181)
$b^S S^2$		8.39e-05***	7.87e-05
		(5.106)	(0.826)
$b^S S^3$			1.66e-07
			(0.0556)
a	-0.000402***	-0.000391	0.000819
	(-4.509)	(-1.015)	(0.894)
a^2		-5.29e-07	-0.000141
		(-0.0293)	(-1.504)
a^3			4.43e-06
			(1.616)

		Flow	
VARIABLE	(1)	(2)	(3)
$b^q_p q$	0.00516***	0.00531***	0.0127***
	(18.78)	(10.29)	(11.03)
$b^q_p q^2$		-6.23e-06	-0.000823***
		(-0.333)	(-7.017)
$b^q_p q^3$			2.61e-05***
			(6.859)
$b^q_r q$	-0.000127*	0.00116***	0.000983
	(-1.701)	(3.552)	(1.436)
$b^q_r q^2$		-6.22e-05***	-4.21e-05
		(-4.040)	(-0.556)
$b^q_r q^3$			-6.37e-07
			(-0.268)
$b^S_p S$	0.00415***	0.00348***	0.0104***
	(16.04)	(7.026)	(8.858)
$b^S_p S^2$		3.00e-05	-0.000740***
		(1.608)	(-6.189)
$b^S_p S^3$			2.46e-05***
			(6.471)
$b^S_r S$	0.000463***	-0.00123***	-0.00163*
	(4.348)	(-3.526)	(-1.774)
$b^S_r S^2$		8.11e-05***	0.000127
		(4.987)	(1.316)
$b^S_r S^3$			-1.46e-06
			(-0.484)

Table A1 Variable Definitions

Name	Description
Gross return	Monthly gross fund return
Benchmark-adjusted gross return	Monthly gross fund return adjusted by the fund's Morningstar 3x3 category benchmark
TNA	Monthly fund size (total net assets), in millions of USD, winsorized at the 1% and 99% levels, adjusted for inflation by the ratio of the value all US stocks that month to the value of all US stocks in December 2011
LTNA	TNA of a fund, lagged one month
LlogTNA	Log of TNA of a fund, lagged one month
Institutional proportion AUM	Proportion of a fund's TNA that consists of institutional investors' funds, estimated in December each year
Fund size	Synonym for TNA
Sqr Fund size	Fund size squared
Fund age	Fund age in years
Fund Turnover	Monthly fund turnover (annual turnover divided by 12), winsorized at the 1% and 99% levels
Fund Expense ratio	Monthly fund expense ratio (annual expense ratio divided by 12), winsorized at the 1% and 99% levels
Active Funds	The number of actively managed funds in the sample, estimated annually
Active Families	The number of fund management firms that have actively managed funds in the Morningstar 3x3 categories, estimated annually
N Manager	The number of fund managers running a fund, estimated annually
Family TNA	Monthly sum of the TNA of all the fund family's actively managed funds in the 3x3 Morningstar categories, in millions of USD
Family Nfund	Monthly count of all the fund family's actively managed funds in the 3x3 Morningstar categories
Industry size	Monthly ratio of the sum of the TNA (not adjusted for inflation) of all funds in the sample divided by the value of all publicly traded stocks on US stock markets that month
Industry scale	Synonym for industry size
LIndSizeRatio	Industry size, lagged one month
a	Benchmark-adjusted gross return adjusted for fund size and industry size ratio effects, estimated over the full sample period. Winsorized at the 1% and 99% levels
b^q	The coefficient on lagged TNA estimated over the full sample period. Winsorized at the 1% and 99% levels
b^S	The coefficient on lagged Industry size, estimated over the full sample period. Winsorized at the 1% and 99% levels
a_i	Benchmark-adjusted gross return adjusted for fund size and industry size effects, estimated for each fund i over the full sample period. Winsorized at the 1% and 99% levels
b_i^q	The coefficient on lagged TNA estimated for each fund$i$ over the full sample period. Winsorized at the 1% and 99% levels.
b_i^S	The coefficient on lagged Industry size, estimated for each fund i over the full sample period. Winsorized at the 1% and 99% levels
a_{it}	Benchmark-adjusted gross return adjusted for fund size and industry size effects, estimated for each fund i each month t over rolling windows. Winsorized at the 1% and 99% levels
b_{it}^q	The coefficient on lagged TNA estimated for each fund i each month t over rolling windows. Winsorized at the 1% and 99% levels
b_{it}^S	The coefficient on lagged Industry size, winsorized at the 1% and 99% levels, estimated for each fund i each month t over rolling windows. Winsorized at the 1% and 99% levels
Rf	Monthly risk-free rate
MktRf	Monthly excess return of the market over the risk-free rate
SMB	Monthly Fama-French 5-Factor size factor
HML	Monthly Fama-French 5-Factor value factor
CMA	Monthly Fama-French 5-Factor investment factor
RMW	Monthly Fama-French 5-Factor profitability factor
LIQv	Monthly Pastor-Stambaugh liquidity factor

Table A2 Fund Growth predicted by Rolling Window Coefficients

This table gives the results of panel regressions of fund size growth on lagged fund size and industry scale coefficients estimated using 2SLS RD regressions over rolling 3-year windows and adjusted for fund, family and market characteristics. The dependent variable is the ratio of the log fund size at the end of each year to the full sample mean log fund size. The lag periods are 1, 3 and 5 years. ***,**, and * denote significance at the 1%, 5% and 10% level, respectively.

	(1) logTNAratio	(2) logTNAratio	(3) logTNAratio	(4) logTNAratio
		1-year lag		
L.$r\_b_{it}^{S}$ (log)	0.00193***			0.00888***
	(5.341)			(11.88)
L.r_a_{it} (log)		0.0103***		0.0747***
		(6.430)		(11.12)
L.$r\_b_{it}^{q}$ (log)			-0.0718***	0.373***
			(-6.397)	(8.455)
Constant	1.001***	1.000***	1.000***	1.000***
	(1,505)	(1,459)	(1,455)	(1,451)
Observations	30,351	30,351	30,351	30,351
R-squared	0.001	0.001	0.001	0.006
		3-year lag		
L3.$r\_b_{it}^{S}$ (log)	0.00402***			0.00734***
	(10.64)			(10.28)
L3.r_a_{it} (log)		-0.000106		0.0362***
		(-0.0667)		(5.676)
L3.$r\_b_{it}^{q}$ (log)			-0.0593***	0.169***
			(-5.420)	(3.941)
Constant	0.999***	0.999***	0.998***	0.998***
	(1,397)	(1,351)	(1,336)	(1,335)
Observations	24,971	24,971	24,971	24,971
R-squared	0.005	0.000	0.001	0.006
		5-year lag		
L5.$r\_b_{it}^{S}$ (log)	0.00377***			0.00405***
	(9.149)			(5.309)
L5.r_a_{it} (log)		-0.00577***		0.00335
		(-3.371)		(0.479)
L5.$r\_b_{it}^{q}$ (log)			-0.0268**	0.00539
			(-2.238)	(0.113)
Constant	0.995***	0.996***	0.995***	0.995***
	(1,245)	(1,204)	(1,192)	(1,192)
Observations	20,351	20,351	20,351	20,351
R-squared	0.004	0.000	0.000	0.004

Table A3 Coefficient Correlation Matrix

This table gives the correlations between fund size and industry scale coefficients estimated using 2SLS RD regressions over rolling 3-year windows. Panels A and B gives the correlations for the level and log coefficients respectively. Panel C gives the correlation for log coefficients adjusted for fund, family and market characteristics.

		Panel A	
	a(level)	b^S (level)	b^q (level)
a(level)	1.000	-0.876	-0.257
b^S (level)	-0.876	1.000	-0.054
b^q (level)	-0.257	-0.054	1.000
		Panel B	
	a(log)	b^S (log)	b^q (log)
a(log)	1.000	-0.430	-0.825
b^S (log)	-0.430	1.000	-0.069
b^q (log)	-0.825	-0.069	1.000
		Panel C	
	r_a(log)	r_b^S (log)	r_b^q (log)
r_a(log)	1.000	-0.429	-0.830
r_b^S (log)	-0.429	1.000	-0.067
r_b^q (log)	-0.830	-0.067	1.000