

Capital Allocation and the Market for Mutual Funds: Inspecting the Mechanism*

Jules H. van Binsbergen[†] Jeong Ho (John) Kim[‡] Soohun Kim[§]

August 21, 2025

Abstract

We exploit heterogeneity in decreasing returns to scale parameters across mutual funds to analyze the importance of scalability for investors' capital allocation decisions. We find strong evidence that steeper decreasing returns to scale attenuate flow sensitivity to performance. We calibrate a rational model of active fund management and show that a large fraction of cross-sectional variation in assets-under-management is due to investors anticipating the effects of scale on return performance. We conclude that decreasing returns to scale play a key role in achieving equilibrium in the intermediated investment management market.

*We are grateful for comments from our discussants, Laurent Barras, Jaewon Choi, Yi Li, Ľuboš Pástor, and Andrew Patton. We are also grateful for comments from Itay Goldstein, Nikolai Roussanov, Robert Stambaugh, and Youchang Wu, conference participants at the American Finance Association Annual Meeting, the Australasian Finance & Banking Conference, the China International Risk Forum, the Conference on Financial Economics and Accounting, the Econometric Society World Congress, the Finance Down Under Conference, the Fixed Income and Financial Institutions Conference, and the Northern Finance Association Meetings, and seminar participants at Emory University, Federal Reserve Board, Florida State University, Georgia Tech, and University of Toronto.

[†]University of Pennsylvania and NBER. Email: julesv@wharton.upenn.edu

[‡]Florida State University. Email: johnkimjeongho@gmail.com

[§]KAIST. Email: soohun.kimi@gmail.com

1 Introduction

An important determinant of the Net Present Value (NPV) of an investment project is its scalability. Even if the marginal profitability on a project is large at small scale, when the profitability deteriorates quickly with size, agents will choose not to commit much capital to such projects, and, in the presence of fixed costs, may choose to forgo them altogether. Despite the importance of scalability, surprisingly little empirical work has quantitatively evaluated how important its cross-sectional variation is for capital allocation. In this paper, we fill this void by focusing on the mutual fund market, where measuring the scalability of investment strategies has become commonplace.¹ In particular, the literature has argued that decreasing returns to scale (DRS) play a key role in equilibrating the mutual fund market.² Consistent with this argument, we show that scalability is an important driver of investors' capital allocation decisions by exploiting heterogeneity across funds in DRS parameters: steeper DRS attenuate flow sensitivity to performance (FSP). Further, we calibrate a rational model of active fund management and show that 54% of the cross-sectional variation in fund size can plausibly be attributed to heterogeneity in DRS.

Our approach closely follows the insights from Berk and Green (2004). As the percentage fees charged by funds change infrequently, equilibration operates primarily through their size (or Assets Under Management (AUM)). When a fund outperforms, investors rationally learn that the fund is a positive NPV investment at its current size. In turn, flows go to that fund, eroding this positive NPV due to decreasing returns to scale: as the fund grows, its manager

¹See, for example, Chen et al. (2004), Yan (2008), Zhu (2018), and Barras, Gagliardini, and Scaillet (2022), all of which find that an active fund's ability to outperform its benchmark declines as its size increases.

²See Berk and Green (2004), Berk and van Binsbergen (2015), Pástor, Stambaugh, and Taylor (2020), and Barras, Gagliardini, and Scaillet (2022).

finds it increasingly difficult to put the new inflows to good use, leading to a deterioration of the fund’s performance. The inflows will stop when the fund is no longer a positive NPV investment opportunity, and its abnormal return to investors has reverted back to zero.

We inspect this equilibrating mechanism more closely by formally deriving, in the context of the Berk and Green model, the relation between DRS and flow sensitivity to performance: as a fund’s returns decrease in scale more steeply (steeper DRS), the positive net alpha is competed away with a smaller amount of capital inflows, making flows less sensitive to performance (weaker FSP).

To test this theoretical insight, one needs a source of variation in DRS in addition to observing investor reactions to this variation. We demonstrate that there is a substantial amount of heterogeneity in DRS across individual funds, with correspondingly heterogeneous FSP across funds. Our approach can be interpreted as inferring how the subjective size-performance relation, perceived by investors in real time, is incorporated into the flow-performance relation going forward. Consistent with our hypothesis, we find that a steeper DRS parameter predicts a lower FSP.

The main challenge in estimating the effect of DRS on FSP is the estimation error in fund-specific DRS (from fund-by-fund regressions), which is likely to induce attenuation bias in the point estimates of the DRS-FSP relation. Indeed, adjusting these DRS-FSP relation estimates for the errors-in-variable bias under the classical measurement error assumption—the errors are independent to the actual DRS—suggests that they are biased toward zero.

To address this issue, we estimate the DRS-FSP relation by instrumenting for the heterogeneity in DRS with a set of fund characteristics that are plausibly related to the scalability

of investment strategies.³ In particular, by regressing the fund-specific DRS estimates on these characteristics, we obtain fitted values that we use as a more robust way of obtaining cross-sectional variation. Importantly, we show that, while the characteristics-based estimates of the DRS-FSP relation remain statistically significant, they become substantially more negative, comparable in magnitude to those implied by the classical measurement error assumption. This result suggests that the characteristics-based approach is able to alleviate the errors-in-variables problem.

Next, we turn to the economic significance of our estimates. In particular, we assess how equilibrium fund size is affected by the cross-sectional variation in DRS parameters. This exercise does require model assumptions. We calibrate a rational model in the spirit of Berk and Green (2004). After simulating data in which investors know the DRS can vary by fund, we check how much of the simulated size can be explained by counterfactual fund sizes computed under the assumption that investors believe the DRS is the same for all funds. We find that, on average, more than half (54%) of the variance of fund sizes across funds and periods can be related to cross-sectional variation in DRS parameters. Importantly, although we do not target the DRS-FSP relation in our calibration, our model produces DRS-FSP relation estimates that are quantitatively similar to those estimated from the actual data. Thus, the model does a good job of approximating the observed equilibrium in the mutual fund market.

Beyond implications for fund flows, the degree of DRS also has implications for fund size. In the model, fund size is directly proportional to the ratio of perceived skill to per-

³We investigate a number of characteristics that seem relevant a priori (also from the previous literature) for heterogeneity in scalability. For example, we find the degree of DRS is stronger for higher-volatility funds, sole-managed funds, small-cap funds, as well as funds charging higher fees.

ceived scalability: all else equal (holding the alpha earned on the first dollar fixed), the DRS parameter should be lower for larger funds. This prediction is confirmed in our empirical analysis. Moreover, if investors learn about funds as in the model, the (log) fund size should converge to the (log) optimal size—the ratio of true skill to true scalability—as funds grow older. In Appendix D, we provide empirical evidence consistent with this prediction: the estimated optimal size largely explains capital allocation across older funds in the data. The size of older funds remains significantly related to their estimated optimal size, even when we control for an alternative proxy for optimal size that ignores cross-sectional heterogeneity in DRS. Again, investors seem to account for not only the average DRS, but also the heterogeneity of DRS across funds.

Taken together, our results demonstrate that investors *do* account for the adverse effects of fund scale in making their capital allocation decisions.⁴ The previous literature has often deemed mutual fund investors as naive return chasers because fund flows respond to past performance although performance is not persistent,⁵ and because funds show little evidence of outperformance.⁶ In contrast, Berk and Green (2004) argue that they are consistent with a model of how competition between rational investors determines the net alpha in equilibrium. We contribute to this debate by presenting findings that are hard to reconcile with anything other than the existence of rational fund flows.

⁴It is plausible that some investors are better than others at understanding the role of scalability for fund performance. Indeed, Choi and Robertson (2020) find that individual investors believe that actively managed funds do not suffer from diseconomies of scale. What is testable, however, is whether the *marginal* investor understands the role of scalability and accounts for this in making his capital allocation decisions. For example, even if individual investors allocate too much capital to a fund because they ignore the adverse effects of fund scale, sophisticated investors of the fund can allocate their capital away from it to adjust its size on the margin.

⁵See Chevalier and Ellison (1997) and Sirri and Tufano (1998), among others.

⁶See Carhart (1997), Fama and French (2010), and Del Guercio and Reuter (2013), among others.

Closely related to our paper is Barras, Gagliardini, and Scaillet (2022), who also find that both skill and scalability—the degree of fund-level DRS—vary substantially across funds. They find that the majority of funds add value, consistent with rational equilibrium models of active mutual fund management. In contrast, we propose scalability as a key determinant of the flow-performance relation based on such models,⁷ a hypothesis that we test by exploiting the fact that scalability varies substantially across funds. Furthermore, contrary to their analysis, we quantify the importance of cross-sectional variation in DRS for capital allocation decisions.

2 Definitions and Hypothesis

Let R_{it}^n denote the return in excess of the risk-free rate earned by fund i 's investors at time t and let R_{it}^B denote the excess return of the manager's benchmark over the same time interval.

At time t , the investor observes the manager's net return outperformance,

$$\alpha_{it+1} \equiv R_{it}^n - R_{it}^B. \quad (1)$$

We assume throughout that α_{it} can be expressed as follows:

$$\alpha_{it} = a_i - b_i h(q_{it-1}) + \epsilon_{it}, \quad (2)$$

where q_{it-1} denotes the size (i.e., real AUM) of fund i at time $t - 1$, a_i denotes a parameter that captures fund i 's gross alpha on the first dollar net of the percentage fee its manager

⁷Studies that discuss other cross-sectional determinants of the flow-performance relation include Huang, Wei, and Yan (2007) and Chen, Goldstein, and Jiang (2010).

charges, and ϵ_{it} is the noise in observed performance. Here $b_i h(q)$ captures the decreasing returns to scale (DRS) the manager faces, which can vary by fund: $b_i > 0$ is a parameter that captures the cross-sectional variation in DRS technology. For the form of DRS technology, we use the logarithmic specification ($h(q) = \log(q)$) commonly used in empirical studies for simplicity in the rest of the paper and provide a necessary and sufficient condition on the DRS technology for our hypothesis (Proposition 1) in Appendix A.⁸

Now note that α_{it} is an informative signal about a_i : high α_{it} implies good news about a_i and low α_{it} implies bad news about a_i . Thus, at time t , investors use the time- t information set I_t to update their beliefs on a_i implying that the expectation of a_i at time t is:

$$\theta_{it} \equiv E[a_i | I_t]. \quad (3)$$

Let $\bar{\alpha}_{it}(q)$ denote investors' subjective expectation of α_{it+1} when fund i has size q at time t , i.e., fund i 's *net alpha*:

$$\bar{\alpha}_{it}(q) = \theta_{it} - b_i h(q). \quad (4)$$

In equilibrium, the size of the fund q_{it} adjusts to ensure that there are no positive NPV investment opportunities so $\bar{\alpha}_{it}(q_{it}) = 0$ and

$$\frac{\theta_{it}}{b_i} = h(q_{it}) = \log(q_{it}). \quad (5)$$

Following Berk and Green (2004), we assume in the rest of the paper that (i) investors' prior

⁸Another widely used functional form of DRS technology in the size-performance analysis is the linear model ($h(q) = q$). In Appendix B, we demonstrate that the data favor a log DRS technology specification over a linear one.

is that a_i is normally distributed with mean θ_{i0} and variance σ_0^2 , and (ii) ϵ_{it} is normally distributed with mean zero and variance σ_ϵ^2 , but we relax these assumptions in Appendix A.⁹ Then, it is straightforward that the mean of investors' posteriors satisfies the following recursion:

$$\theta_{it} = \theta_{it-1} + \frac{\sigma_0^2}{\sigma_\epsilon^2 + t\sigma_0^2} \alpha_{it}. \quad (6)$$

Next, let the flow of capital into mutual fund i at time t be denoted:

$$F_{it} \equiv \log(q_{it}/q_{it-1}) = \frac{\theta_{it} - \theta_{it-1}}{b_i} = \frac{\sigma_0^2}{\sigma_\epsilon^2 + t\sigma_0^2} \frac{\alpha_{it}}{b_i},$$

where the first equality follows from (5) and the last equality follows from (6). Differentiating this expression with respect to α_{it} ,

$$\frac{\partial F_{it}}{\partial \alpha_{it}} = \frac{\sigma_0^2}{\sigma_\epsilon^2 + t\sigma_0^2} \frac{1}{b_i} > 0, \quad (7)$$

so good (bad) performance results in an inflow (outflow) of funds. This result is one of the important insights from Berk and Green (2004).

Taking the derivative of the flow-performance sensitivity with respect to b_i , we see that steeper DRS lead to a weaker flow sensitivity to performance (FSP):

$$\frac{\partial}{\partial b_i} \left(\frac{\partial F_{it}}{\partial \alpha_{it}} \right) = -\frac{\sigma_0^2}{\sigma_\epsilon^2 + t\sigma_0^2} \frac{1}{b_i^2} < 0. \quad (8)$$

This leads to the following proposition, which will be our main hypothesis that we take to

⁹All we need is that the time- $t-1$ conditional probability density of α_{it} , $f_{t-1}(\alpha_{it})$, satisfies the monotone likelihood ratio property— $f_{t-1}(\alpha_{it} | a_i) / f_{t-1}(\alpha_{it} | a_i^c)$ is increasing (decreasing) in α_{it} if $a_i \geq (\leq) a_i^c$.

the data.

Proposition 1 *Steeper DRS lead to a weaker FSP.*

Intuitively, as a fund’s returns decrease in scale more steeply (steeper DRS), the positive net alpha is competed away with a smaller amount of capital inflows, making flows less sensitive to performance (weaker FSP).¹⁰

Remark 1 Note that the scalability parameter is assumed to be constant for a given fund, but this assumption is *not* essential for Proposition 1. If the scalability parameter is time-varying—with b_{it} denoting the true DRS fund i faces at time t —it is straightforward to show that the derivative (8) becomes

$$\frac{\partial}{\partial b_{it}} \left(\frac{\partial F_{it}}{\partial \alpha_{it}} \right) = -\frac{\sigma_0^2}{\sigma_\epsilon^2 + t\sigma_0^2} \frac{1}{b_{it}^2} < 0, \quad (9)$$

so steeper DRS still lead to a weaker FSP.

Remark 2 Note that the scalability of a fund is implicitly assumed to be known by investors, as in much of the earlier literature, but this assumption is *not* essential for Proposition 1.¹¹

If the scalability parameter obeys an AR(1) process—with $b_{it} = \phi_0 + \phi_1 b_{it-1} + \eta_{it}$ —and investors’ prior is that a_i and b_{it} follow a bivariate normal distribution, it is straightforward

¹⁰Equation (8) further suggests that the relationship between FSP and scalability is nonlinear, specifically that the coefficient of FSP on DRS squared should be positive. Motivated by this, we test for nonlinearity in the DRS-FSP relation and find some evidence of convexity, although the estimated convexity is small relative to the linear term. These results are available upon request.

¹¹For models that relate capital allocation to learning about scalability, see Pástor and Stambaugh (2012) and Kim (2022).

to show that the derivative (8) becomes

$$\frac{\partial}{\partial \hat{b}_{it}} \left(\frac{\partial F_{it}}{\partial \alpha_{it}} \right) = - \frac{\frac{\partial \theta_{it}}{\partial \alpha_{it}} - 2 \log(q_{it}) \frac{\partial \hat{b}_{it}}{\partial \alpha_{it}}}{\hat{b}_{it}^2}, \quad (10)$$

where

$$\begin{aligned} \frac{\partial \theta_{it}}{\partial \alpha_{it}} &= \frac{\text{Var}_{t-1}(a_i) - \text{cov}_{t-1}(a_i, b_{it-1}) \log(q_{it-1})}{\text{Var}_{t-1}(\alpha_{it})} \\ \frac{\partial \hat{b}_{it}}{\partial \alpha_{it}} &= -\phi_1 \frac{\text{Var}_{t-1}(b_{it-1}) \log(q_{it-1}) - \text{cov}_{t-1}(a_i, b_{it-1})}{\text{Var}_{t-1}(\alpha_{it})} \end{aligned}$$

The derivative in equation (10) is negative under two sufficient conditions: (i) the conditional covariance $\text{cov}_{t-1}(a_i, b_{it-1})$ is sufficiently small and (ii) both θ_{it} and θ_{it-1} are sufficiently large. The first condition implies that investors interpret strong performance as a signal of both higher-than-expected skill ($\theta_{it} > \theta_{it-1}$) and better-than-expected scalability ($\hat{b}_{it} \leq E_{t-1}(b_{it})$). If, instead, investors believe a priori that skill and scalability are tightly positively correlated, they may interpret superior performance as a sign of lower skill or worse scalability, resulting in a negative flow-performance relation. The second condition is likely to hold in practice, since funds face fixed operating costs and will optimally exit when they can no longer cover these costs. This behavior naturally bounds θ_{it} and θ_{it-1} from below. In sum, even when the scalability parameter is time-varying and unobservable, steeper DRS (perceived by investors in real time) still lead to a weaker FSP under empirically plausible conditions.

3 Data

Our data come from CRSP and Morningstar. We require that funds appear in both the CRSP and Morningstar databases, which allows us to validate data accuracy across the two. We merge CRSP and Morningstar based on funds' tickers, CUSIPs, and names. We then compare assets and returns across the two sources in an effort to check the accuracy of each match following Berk and van Binsbergen (2015) and Pástor, Stambaugh, and Taylor (2015). We refer the readers to the data appendices of those papers for the details. Our mutual fund data set contains 3,066 actively managed domestic equity-only mutual funds in the U.S. between 1991 and 2014.¹² Finally, we drop any fund observations before the fund's (inflation-adjusted) AUM reaches \$5 million.

We now define the key variables used in our empirical analysis: fund performance, fund size, and fund flows. Summary statistics are in Table 1.

3.1 Fund Performance

We take two approaches to measuring fund performance. First, we use the standard risk-based approach. The recent literature finds that investors use the CAPM in making their capital allocation decisions (Berk and van Binsbergen (2016) and Barber, Huang, and Odean (2016)), so we adopt the CAPM. In this case the risk adjustment R_{it}^{CAPM} is given by:

$$R_{it}^{\text{CAPM}} = \beta_{it}\text{MKT}_t,$$

¹²We start the sample in 1991, the first year in which CRSP provides monthly data on funds' size.

where MKT_t is the realized market excess return and β_{it} is the market beta of fund i . We estimate β_{it} by regressing the fund's excess return to investors onto the market portfolio over the 60 months prior to month t . To produce reliable beta estimates, we require a fund to have at least two years of track record to estimate its betas from the rolling window regressions.

Second, we follow Berk and van Binsbergen (2015) by taking the set of available Vanguard index funds as the alternative investment opportunity set,¹³ so the benchmark of a fund is defined as the closest portfolio in that set to it. Let R_t^j denote the excess return earned by investors in the j 'th Vanguard index fund at time t . Then the benchmark return for fund i is given by:

$$R_{it}^{\text{VG}} = \sum_{j=1}^{n(t)} \beta_i^j R_t^j,$$

where $n(t)$ is the number of Vanguard index funds available at time t and β_i^j is obtained from the appropriate linear projection of fund i onto the set of Vanguard index funds. As pointed out by Berk and van Binsbergen (2015), using Vanguard funds as the benchmark ensures that this alternative investment opportunity set was marketed and tradable at the time. Again, we require a fund to have at least 24 months of data to estimate its projection coefficients (β_i^j) used to calculate the Vanguard benchmark for fund i .

Our measures of fund performance are then $\hat{\alpha}_{it}^{\text{CAPM}}$ and $\hat{\alpha}_{it}^{\text{VG}}$, the realized return for the fund in month t less $\hat{R}_{it}^{\text{CAPM}}$ and $\hat{R}_{it}^{\text{VG}}$, respectively. The average of $\hat{\alpha}_{it}^{\text{CAPM}}$ is +1.0 basis points per month, while the average $\hat{\alpha}_{it}^{\text{VG}}$ is -1.7 basis points per month.

¹³See Table 1 of that paper for the list of Vanguard Index Funds used to calculate the Vanguard benchmark.

3.2 Fund Size and Flows

We adjust all AUM numbers by inflation by expressing them in January 1, 2000 dollars. Adjusting AUM by inflation reflects the notion that the fund’s real (rather than nominal) size is relevant for capturing decreasing returns to scale in active management—lagged real AUM corresponds to q_{it-1} in the model from Section 2. There is considerable dispersion in real AUM: the inner-quartile range is from \$44 million to \$621 million, while the 99th percentile is orders of magnitude larger at \$16 billion.

Flows are measured in two different ways. First, as in the model, we define fund flow F as the logarithmic change in real AUM—the percentage change in fund size. Alternatively, we calculate flows for fund i in month t as:

$$F_{it} = \frac{AUM_{it} - AUM_{it-1}(1 + R_{it})}{AUM_{it-1}(1 + R_{it})},$$

where AUM_{it} is fund i ’s nominal AUM at the end of month t , and R_{it} is fund i ’s total return in month t .¹⁴ Under this more standard definition of F , flows represent the percentage change in new assets. The flow of fund data contain some implausible outliers, so we winsorize the two flow variables at their 1st and 99th percentiles. Mean monthly changes in fund size and in new assets are 0.8% and 0.5%, respectively.

¹⁴Note that we use $AUM_{it-1}(1 + R_{it})$ in the denominator rather than AUM_{it-1} , which is more common in the literature. Unfortunately, this definition distorts the flow for very large negative returns (see Berk and Green (2004)). Regardless, our findings are unaffected if we were to use AUM_{it-1} in the denominator.

4 Method

Our analysis relies on a theoretical link between DRS and flow sensitivity to performance.

We discuss how we estimate each part in the following sections.

4.1 Fund-Specific Decreasing Returns to Scale (DRS)

Empirically, the net alpha earned by fund i 's investors in month t is given by:

$$\alpha_{it} = a_i - b_i \log(q_{it-1}) + \epsilon_{it}, \quad (11)$$

where a_i is the fund fixed effect, b_i captures the size effect, which can vary by fund, and q_{it-1} is the fund's lagged real AUM.¹⁵ This simple regression model corresponds to the model in Section 2.

We depart from much of the literature by allowing for heterogeneity in the size-performance relation across funds. Indeed, the effect of scale on a fund's performance is unlikely to be constant across funds. For example, a fund's returns should be decreasing in scale more steeply for those that invest in small and illiquid stocks.

We start our analysis by estimating fund-specific b_i parameters. It is well known that the OLS estimators of b_i in (11) are subject to a small-sample bias (Stambaugh (1999)). The small-sample bias arises because changes in fund size tend to be positively correlated with unexpected fund returns. To address this bias, we follow Amihud and Hurvich (2004) and Barras, Gagliardini, and Scaillet (2022) and include a proxy for the size innovation v_{it}^c (see

¹⁵Our results going forward are the same if we were to use gross (instead of net) alpha in equation (11). This robustness is consistent with existing literature showing that fee changes are rare, so they are unlikely to play an important role in equilibrating the mutual fund market.

Appendix C):¹⁶ for each fund i at time t , we define the fund-specific DRS estimate $\hat{b}_{it}$ to be the coefficient of $-\log(q_{i\tau-1})$ in the time-series regression of $\hat{\alpha}_{i\tau}$ on $-\log(q_{i\tau-1})$ and $v_{i\tau}^c$ (including an intercept) using sixty months of data before time t . We require at least 3 years of data to estimate fund-specific DRS of a fund.

Intuitively, the estimate of b_i , $\hat{b}_{it}$, represents investors' perception of the effect of size on performance for fund i at time t based on information prior to time t . Panel A of Figure 1 shows how the cross-sectional distribution of $\hat{b}_{it}$ using the CAPM alpha varies over time. For each month in 1991 through 2014, the figure plots the average as well as the percentiles of the estimated fund-specific b parameters across all funds in that month. The plot shows considerable heterogeneity in DRS across funds:¹⁷ the interquartile range is more than 4 times larger than the estimates' cross-sectional median. We find that, for the average fund, a 1% increase in fund size is associated with a sizeable decrease in performance of about 0.4 basis points per month. This evidence suggests that the subjective size-performance relation, perceived by investors in real time, provides identifying variation in the extent of DRS.

Panel B of Figure 1 shows how the cross-sectional distribution of $\hat{b}_{it}$ varies over time when we estimate fund-specific DRS using $\hat{\alpha}_{it}^{VG}$. Similar to when we use the CAPM alpha to estimate $\hat{b}_{it}$ in Panel A, this plot shows considerable heterogeneity in DRS across funds, although these estimates typically indicate milder DRS.

¹⁶Online Appendix considers alternative proxies for $v_{i\tau}^c$ using approaches from the existing literature, as well as using a novel approach that relies on our model. Our results are robust to these alternative approaches.

¹⁷Some of the heterogeneity in DRS could be attributable to estimation error. See Barras, Gagliardini, and Scaillet (2022) who also find substantial heterogeneity in DRS even after adjusting for the bias arising from estimation error.

4.2 Fund-Specific Flow Sensitivity to Performance (FSP)

We estimate the fund-specific FSPs by estimating the following regression fund-by-fund:

$$F_{it} = c_i + \gamma_i P_{it-1} + v_{it}, \quad (12)$$

where P_{it-1} is annual alpha for the year leading to month $t - 1$, computed by compounding the monthly alphas. This regression is consistent with empirical evidence that investors do not respond immediately.¹⁸ Parameter $\gamma_i > 0$ captures the positive time-series relation between performance and fund flows, which can vary by fund.

For each fund i at time t , we calculate the fund's FSP by estimating (12) using its data over the subsequent 5 years. Let $\widehat{FSP}_{it}$ be the flow-performance sensitivity estimate from that model. We require these coefficient estimates to be obtained from at least 3 years of data. For the average fund, we observe that an increase of 1% in the annual CAPM alpha is associated with a 0.1% increase in monthly flows next month.

Panel A of Figures 2 and 3 display the evolution of the $\widehat{FSP}_{it}$ distributions over time measuring flows as percentage changes in fund size and in new assets, respectively. Both plots manifest considerable heterogeneities in the flow-performance relation across funds. Moreover, these plots show that while the average $\widehat{FSP}_{it}$ do not exhibit any obvious trend, they are certainly time varying. Also noteworthy is the fact that the distributions remain roughly the same over our sample period, conditional on the median.

Panel B of Figures 2 and 3 display the evolution of the $\widehat{FSP}_{it}$ distributions over time

¹⁸For example, Busse and Irvine (2006) show that long-term performance predicts flows better than short-term performance.

when we estimate fund-specific FSP using $\hat{\alpha}_{it}^{\text{VG}}$. Similar to when we use the CAPM alpha to estimate $\widehat{FSP}_{it}$ in Panel A, these plots manifest considerable heterogeneities in the flow-performance relation across funds.

5 Results

5.1 DRS and Flow Sensitivity to Performance

To examine whether fund-specific DRS parameters affect capital allocation decisions, we run panel regressions of fund i 's FSP going forward in month t , $\widehat{FSP}_{it}$, on its DRS estimated as of the previous month-end, $\hat{b}_{it}$. We test the null hypothesis that the slope on $\hat{b}_{it}$ is zero.¹⁹ We report the results based on raw estimates in Table 2.²⁰ In Panel A, we present results using estimates of fund-specific DRS and FSP based on the CAPM alpha, and in Panel B, we present results based on $\hat{\alpha}_{it}^{\text{VG}}$.

We focus on variation in sensitivity coming from the market equilibrating mechanism by including both fund and month fixed effects. Fund fixed effects absorb variation in FSP, e.g., due to cross-sectional differences in investor clientele,²¹ or baseline fund scalability, while month fixed effects soak up variation in FSP due to factors like time-varying investor attention allocation.²² Conceptually, we think the relationship between FSP and scalability in our regression models is driven by within-fund variation in *perceived* scalability over time:

¹⁹While the measurement error in $\widehat{FSP}_{it}$ will not induce bias in the OLS coefficients, the measurement error in $\hat{b}_{it}$ will bias the OLS estimator toward zero. For now, we do not worry, as the errors-in-variables problem will work against us from finding a statistically significant relation that the model predicts.

²⁰Table 2 reports the double-clustered (by fund and time) standard errors.

²¹See Berk and Tonks (2007) for evidence of clientele differences—some investors tend to update faster than others.

²²These fixed effects also subsume any potential time-series variation in FSP due to different stages of development in the US (Ferreira et al. (2012)) and/or across market states (Franzoni and Schmalz (2017)).

as investors update their beliefs by observing a fund’s returns and size, the fund’s perceived scalability fluctuates, leading to fluctuations in the flow sensitivity to the fund’s performance.

In the odd columns, we only include month and fund fixed effects. The results in Panel A are consistent with the main prediction of our model: the estimated coefficients on $\widehat{b}_{it}$ are significantly negative, with t -statistics of -6.3 in column 1 and -3.0 in column 3. These findings are unaffected by including a host of controls in the even columns, where we add proxies for participation costs considered by Huang, Wei, and Yan (2007),²³ as well as performance volatility and fund age.²⁴ The slopes on $\widehat{b}_{it}$ remain negative and highly significant, with t -statistics of -7.1 in column 2 and -3.3 in column 4, and their magnitude increases.

In Panel B, the same conclusions continue to hold when we use estimates of fund-specific DRS and FSP based on $\widehat{\alpha}_{it}^{VG}$. Just like in Panel A, the estimated coefficients on $\widehat{b}_{it}$ are significantly negative, and their magnitude increases when we include a host of controls.

Table 3 repeats this exercise with percentile ranks in each month based on $\widehat{b}_{it}$ and $\widehat{FSP}_{it}$. Now, we do not use month fixed effects: percentile ranks already control for time variation in the flow-performance relation. In each column, the estimated coefficient on $\widehat{b}_{it}$ is significantly negative at the 1% confidence level.

To summarize, we find a strong negative relation between DRS and FSP, consistent with the presence of investors rationally accounting for the adverse effects of fund scale in making their capital allocation decisions. Unfortunately, the coefficient values in Table 2 are likely biased towards zero because of the measurement error in $\widehat{b}_{it}$. In Section 5.1.1, we first gauge

²³Specifically, we use marketing expenses, star family affiliation, family size, and fund size to proxy for the variation in investors’ information costs across funds.

²⁴Huang, Wei, and Yan (2012) find a weaker flow-performance sensitivity for funds with more volatile past performance and longer track records.

the severity of attenuation bias under the classical measurement error assumption. In Section 5.1.2, we then exploit a set of fund characteristics that are plausibly related to the scalability of investment strategies as instruments for heterogeneity in DRS parameters across funds to address the attenuation bias associated with estimating the DRS-FSP relation. Finally, in Section 5.1.3, we propose a way of assessing the economic magnitude of these estimated coefficients by computing counterfactual fund sizes.

5.1.1 DRS-FSP Relation Under the Classical Measurement Error Assumption

To gauge the severity of attenuation bias, we adjust the estimated coefficients on $\hat{b}_{it}$ in Table 2 for the errors-in-variable (EIV) problem, assuming that the errors are of the classical type: they are purely random, mean zero, and uncorrelated with the regressors, including the actual b_i , and with the regression errors. Using the standard errors of $\hat{b}_{it}$ to estimate the variance of measurement error in b_i , we calculate the EIV-adjusted coefficients and their standard errors, reported in the last two row of the panel.

As expected, the simple DRS-FSP relation estimates tend to be too small in magnitude. For example, when the DRS-FSP relation is estimated based on the CAPM controlling for other determinants of the flow-performance relation (column 2 of Panel A), the coefficient becomes substantially more negative with the EIV adjustment: -16.11 , compared to -1.30 without this adjustment. Interestingly, the EIV adjustment suggests that the estimated coefficients on $\hat{b}_{it}$ based on the Vanguard benchmark are even more severely biased toward zero: the EIV adjustment makes the coefficients 26 to 45 times larger in magnitude (see the last row of Panel B). Of course, these results are only true if the errors are indeed of the classical type, but they illustrate that our DRS-FSP relation estimates are likely to

be severely biased against confirming our model prediction. Thus, the fact that we find a strong relation between DRS and FSP despite this counterveiling effect further strengthens the support for the model.

5.1.2 DRS-FSP Relation Using the Characteristic Component of DRS

We explore which fund characteristics are correlated with the observed heterogeneity in scalability. Based on this analysis, we obtain an economically interpretable component of $\hat{b}_i$ related to fund characteristics, using which we re-estimate the DRS-FSP relation. The prior evidence of fund-level DRS depending on fund characteristics suggests that this method is likely to deliver a more accurate measure of b_i , thus mitigating the errors-in-variable problem. Indeed, the characteristic-based approach taken here leads to substantially more negative estimates of the DRS-FSP relation.

Determinants of Fund-Level DRS We investigate a number of characteristics that seem relevant a priori (also from the previous literature) for heterogeneity in scalability. The first characteristic is the number of managers. About 59% of our funds are multi-manager funds. The second characteristic is volatility: the standard deviation of fund alphas over the prior 1 year. The next two characteristics we examine are expense ratios and marketing expenses. The fifth characteristic is the international exposure dummy: for any given fund, it is equal to one if we reject the null hypothesis that the coefficients on three Vanguard international index funds are 0 at the 5% confidence level.²⁵ Although we focus on domestic funds, about 28% of them are significantly exposed to international shocks. The sixth characteristic is average

²⁵Three Vanguard index funds are international: European Stock Index, Pacific Stock Index, and Emerging Markets Stock Index.

annual turnover (from CRSP).²⁶ Median annual turnover is 64%. The last characteristic is log real AUM, which checks for nonlinearity in the DRS technology. In analyzing the dependence of scalability on fund characteristics, we also control for loadings on the market, size, value, and momentum factors to capture fund style and risk.²⁷

The selection of these characteristics to capture heterogeneity in DRS followed two steps. First, we adopt volatility and turnover from Pástor, Stambaugh, and Taylor (2015), which, to the best of our knowledge, is the most recent paper touching on how scalability depends on fund characteristics. They examine three characteristics: volatility, turnover, and a small-cap indicator. High-turnover funds and small-cap funds tend to face greater trading costs and therefore steeper DRS; similarly, high-volatility funds—being effectively larger in terms of their trading—also exhibit steeper DRS. We exclude the small-cap indicator because its effect is subsumed by our controls for funds’ investment styles—specifically, their loadings on the market, size, value, and momentum factors.

Second, given the absence of systematic evidence on the determinants of scalability, we add four further characteristics based on our a priori reasoning that they are likely to influence a fund’s scalability: the number of managers, expense ratios, marketing expenses and the fund’s international exposure. A multi-manager fund may exhibit milder DRS because the division of labor might alleviate the negative performance impact of size, enabling the fund to deploy capital more easily. We hypothesize that funds charging higher expense ratios face steeper DRS, based on the model of Stambaugh (2020), which predicts that such funds deviate more from benchmark weights and consequently incur higher trading costs.

²⁶We winsorize turnover at the 1st and 99th percentiles.

²⁷We estimate these risk exposures by regressing the fund’s return on the four Fama-French-Carhart factors over the prior sixty months.

In contrast, we hypothesize that funds with higher marketing expenses exhibit flatter DRS because funds are likely to undertake marketing efforts to attract flows only when they can manage the performance erosion associated with growth.²⁸ Finally, funds with international exposure may face less severe DRS, as international markets tend to have less competition among active funds,²⁹ and access to such diversification opportunities can help mitigate the performance decline associated with asset growth.

We study how these characteristics affect the impact of a fund’s scale on its performance by running panel regressions of fund i ’s backward-looking DRS estimate in month t , $\hat{b}_{it}$, on the fund’s characteristics as of the previous month-end. Table 4 shows the estimation results.³⁰ Panel A reports the results using estimates of fund-specific DRS based on $\hat{\alpha}_{it}^{\text{CAPM}}$; Panel B reports the results based on outperformance relative to the Vanguard benchmark.

We find significant relations between $\hat{b}$ and three characteristics: the number of managers, volatility, and expense ratios (see the first three columns). The slope on marketing expenses (column 4) is insignificantly negative, while the slope on fund size (column 7) is insignificantly positive. The slope on turnover (column 6) is insignificant as well, but its sign is mixed, depending on how we measure fund performance. Finally, the relation between scalability and international exposure (column 5) is both statistically and economically insignificant.

When all seven fund characteristics are added simultaneously, the estimated slopes on volatility and expense ratios are robust, indicating steeper DRS for higher-volatility funds

²⁸It is well documented that funds’ marketing efforts are positively related to investor flows (e.g., Gallaher, Kaniel, and Starks (2009), Christoffersen, Evans, and Musto (2013), and Roussanov, Ruan, and Wei (2021)).

²⁹Consistent with this argument, Dyakov, Jiang, and Verbeek (2020) find that the active fund industry in the U.S. appears to have exceeded its optimal size, whereas, outside of the U.S., there is still room for growth.

³⁰Standard errors of these regressions are double-clustered by fund and time.

and funds charging higher expense ratios. We continue to find a negative, albeit insignificant, relation between $\hat{b}$ and the number of managers, indicating steeper DRS for sole-manager funds. Marketing expenses now enter with a significantly negative slope, indicating that DRS are less pronounced for funds with higher marketing expenses. The relation between $\hat{b}$ and fund size remains insignificantly positive. Finally, the slopes on turnover and international exposure now flip to negative, albeit still insignificant. Hence, in the final column of Table 4, we focus on the specification that includes the three jointly significant fund characteristics.

Implications for DRS-FSP Relation Using the estimates from Table 4, we now obtain predicted values of $\hat{b}_i$ based on fund characteristics, denoted by $\hat{b}_i^{Char}$. This approach increases the accuracy of the b_i estimate insofar as differences in DRS are well captured by fund characteristics. This assumption seems reasonable since the characteristics-based approach substantially reduces the percentage of negative b_i estimates from 37% to 5% in line with the fact that, theoretically, all funds must face DRS in equilibrium. Figure 4 shows how the cross-sectional distribution of $\hat{b}_{it}^{Char}$ estimated from the specification in column 8 of Table 4 varies over time. The distribution of $\hat{b}_{it}^{Char}$ is clearly tighter than that of $\hat{b}_{it}$, consistent with the characteristics-based approach eliminating the estimation error in DRS. Importantly, the plot continues to reveal clear heterogeneity in scalability across funds, indicating that a significant portion of the variation in $\hat{b}_{it}$ reflects genuine differences rather than estimation error.

To address the attenuation bias associated with estimating the DRS-FSP relation, we replace $\hat{b}_i$ by $\hat{b}_i^{Char}$ and rerun the regressions in Table 2, with results tabulated in Table 5. Given that we have multiple characteristics which are significantly related to $\hat{b}_i$, we report

the results based on $\widehat{b}_i^{Char}$ estimated using various first-stage specifications from Table 4. We now obtain even stronger evidence that steeper DRS attenuate FSP: not only are the slopes on $\widehat{b}_i^{Char}$ significant throughout, but they are substantially more negative than those on $\widehat{b}_i$. For example, when flows are measured as the change in fund size, the estimated coefficients in the first four columns of Table 5 are more than seven times larger than the corresponding estimate in column 2 of Table 2. Moreover, we find that the DRS-FSP relation estimates are very similar in magnitude across the four alternative first-stage specifications: in any pairwise comparison, the estimate from one specification lies well within one standard error of the estimate from the other.

In summary, when we conduct the analysis using cleaner measures of DRS, the estimated effects of DRS on capital allocation only become stronger. The magnitudes of these DRS-FSP relation estimates are comparable to those implied by the classical measurement error assumption, and they are also robust to the choice of first-stage specification.

Discussion of the Effect Size Here, we discuss the estimated effect size of scalability on FSP. The coefficient estimates from the last four columns of Panel A of Table 5 (ranging from -5.69 to -6.24) indicate that a one standard deviation increase in fund DRS (ranging from 0.0023 to 0.0031)³¹ is associated with a decrease in FSP of 0.0137 – 0.0193 , or 18% – 25% of the median FSP (0.0756) in our sample (Table 1).

To illustrate, consider two funds, A and B, that begin with the same size and produce the same positive net alpha. The coefficient estimates indicate that, if (i) fund A faces one-standard-deviation steeper DRS than fund B and (ii) fund B faces the median FSP, then fund

³¹Note that the standard deviation of $\widehat{b}_{it}^{Char}$ depends on the first-stage specification used to estimate $\widehat{b}_{it}^{Char}$.

A would attract only about 80% of the new money that fund B would receive. This effect is economically meaningful, especially given that managerial compensation primarily depends on fund size (Berk and van Binsbergen (2015)). Notably, this calculation is conservative in that the estimates from the first four columns of Panel A would suggest a much larger effect of DRS on FSP.

We further benchmark scalability’s explanatory power relative to other determinants of FSP. To that end, we also show the coefficient estimates of the controls in Table 5. First, note that scalability is the only determinant that is always statistically significant. Second, fund size and star family affiliation are the only controls that are significant in most specifications. Two other controls are statistically significant in a few specifications: volatility (even columns of Panel A) and marketing expenses (final column of Panel B). The fact that proxies for participation costs considered by Huang, Wei, and Yan (2007) are mostly statistically insignificant is consistent with their argument that the effect of investors’ participation costs depends on the performance level: funds with lower participation costs have a higher (lower) flow sensitivity to medium (high) performance. Our evidence suggests that these opposing effects cancel out, yielding an insignificant unconditional effect on FSP.³²

Using the same four columns of Panel A, we estimate that a one standard deviation increase in log fund size (1.89) reduces FSP by 0.0378, while a one standard deviation increase in star family affiliation (0.496) reduces FSP by 0.0060. In the two specifications where volatility is also significant (columns 6 and 8), a one standard deviation increase in volatility (0.0116) raises FSP by 0.012.³³

³²Given that many of the controls are insignificant, we checked to make sure that our findings are unaffected by excluding controls that are statistically insignificant. All of these results are available upon request.

³³In unreported results, we find that the estimated coefficients on volatility turn negative when we exclude

Thus, in terms of estimated effect size, scalability dominates all other determinants of FSP except fund size. However, this likely understates the economic importance of scalability, as the substantial cross-sectional dispersion in fund size inflates its seeming variability relative to that of scalability, making scalability’s effect appear smaller by comparison.

Since our fixed-effect approach identifies the relationship between FSP and its various determinants from *within-fund* variation, it is more appropriate to compare the coefficients using within-fund—rather than pooled—measures of variability. Consistent with many funds following highly dynamic strategies (Mamaysky, Spiegel, and Zhang (2008)), the within-fund standard deviation of scalability (ranging from 0.0016 to 0.0023) is about 70% of its pooled standard deviation. In contrast, fund size shows much less within-fund variability (0.876), less than half its pooled value, consistent with strong persistence of fund size. Using within-fund measures of variability, we find that one standard deviation increases in DRS and fund size are associated with reductions in FSP of around 15% and 24% of the median FSP, respectively—effects that are much more similar in magnitude than the pooled analysis suggests.

Taken together, these findings highlight scalability as a key determinant of FSP.

5.1.3 Simulated DRS-FSP Relation

Finally, we use our model to ask how much capital is allocated the way it is because of these differences in DRS—we compute counterfactual fund sizes by assuming the investors believe a priori that returns are decreasing in scale at the same (average) rate for all funds.

scalability, consistent with the evidence of Huang, Wei, and Yan (2012) who find that mutual funds with more volatile performance have weaker flow-performance sensitivity. Our evidence suggests that their result is driven by how alpha volatility affects flow-performance relationship *through* its impact on scalability.

Two factors determine the magnitude of capital response to performance in a rational model: (i) the degree of DRS and (ii) the prior and posterior beliefs about fund skill. Thus, for a given value of b in equation (11), the prior uncertainty about a , σ_0 , can be inferred from the flow-performance relation, provided investors update their posteriors as Bayesians.

We simulate fund alphas from equation (11) by drawing the error terms ϵ_{it} from $N(0, \sigma_\epsilon^2)$. Using (5), we compute fund size as:

$$q_{it} = \exp\left(\frac{\theta_{it}}{b_i}\right),$$

where the mean of investors' posteriors θ_{it} satisfies recursion (6).

Following Berk and Green (2004), we set $\sigma = 20\%$ per year, or 5.77% per month. Since investors are assumed to have rational expectations, we draw each fund's skill a_i from $N(\theta_{i0}, \sigma_0^2)$, while we draw b_i from a scaled Beta distribution that approximates the empirical distribution of b_i , i.e., the distribution of $\hat{b}_{it}^{Char}$ using the CAPM alpha.³⁴ Assuming that $\theta_{i0} = \theta_0$ for all funds leads to a considerably more disperse size distribution than in our actual sample: simulated fund sizes tend to be too big (small) for funds whose returns decrease in scale more gradually (steeply). Accordingly, we model the prior mean as a linear function of b_i , $\theta_0(b_i)$, setting the coefficients such that the simulated mean and standard deviation of log fund size match the empirical benchmark values of 5.13 and 1.89, respectively.

Given all other parameters, we set the prior uncertainty (σ_0) so that the average $\hat{\gamma}_i$ across funds in a typical simulated sample matches the average $\widehat{FSP}_{it}$ in our actual sample, where

³⁴Specifically, we calibrate the scaled Beta distribution so that the first four moments of the simulated b_i match those of $\hat{b}_{it}^{Char}$ (estimated using the specification in column 8 of Panel A of Table 4).

$\hat{\gamma}_i$ is each fund's estimated FSP from the following regression using data for just that fund:

$$\log(q_{it}/q_{it-1}) = c_i + \gamma_i \alpha_{it} + v_{it}.$$

Panel A of Table 6 shows the calibrated parameter values used in our simulation analysis. Panel A also reports the moments we target in our calibration. Note that the simulated moments in the model closely match the target moments from the actual data.

To assess the economic magnitude of the DRS-FSP relation estimates from the actual data, we estimate the DRS-FSP relation in our simulated samples. Panel B of Table 6 reports summary statistics across simulations. Note that the DRS-FSP relation estimates in the model correspond to columns 1–2 of Panel A of Table 2 and columns 1–4 of Panel A of Table 5, which use the CAPM for risk adjustment and the log change in fund size as the flow measure. The simulated estimates tend to be larger in magnitude than the corresponding EIV-adjusted estimates in columns 1–2 of Panel A of Table 2 (–16.1 to –9.14) and the estimates based on the characteristic component of DRS in columns 1–4 of Panel A of Table 5 (–11.1 to –9.11). But importantly, the empirical DRS-FSP relation estimates lie comfortably within the 90% confidence interval for simulated estimates, and vice versa. Thus, the magnitude of the empirical DRS-FSP relation estimates is consistent with what the model predicts, suggesting that the calibrated model does a good job of capturing capital allocation patterns in the data.

To quantitatively assess the role of heterogeneity in scalability in capital allocation, we must construct a counterfactual. We construct the counterfactual by assuming investors wrongly believe that every fund exhibits the same, average degree of DRS: $b_i = 0.0047$

for all funds. Then, by updating investors' beliefs about each fund's skill with its history $\{\alpha_{i\tau}, q_{i\tau-1}\}_{\tau=1}^t$ under the counterfactual assumption, we compute what the size of the fund would have been, q_{it}^C , for every i and t . In a given simulated sample, we calculate the R^2 from a regression of $\log(q_{it})$ on $\log(q_{it}^C)$ to check the goodness of fit by the counterfactual.

We report the results from counterfactual simulations in Panel C of Table 6. The counterfactually computed fund sizes explain about 46% of the variation of simulated fund sizes. While counterfactual sizes are positively related to actual sizes, they are considerably larger than actual sizes and their distributions are substantially tighter than those of actual sizes. Thus, the counterfactuals ignoring heterogeneity in DRS are very different than the actual size. In this sense, we can interpret $1 - R^2$ as a lower bound on the role of heterogeneity in scalability on capital allocation: more than half of the variance of fund sizes can be related to cross-sectional variation in DRS parameters, which is economically significant.

To summarize, Table 6 shows that a significant fraction of equilibrium capital allocation can be plausibly explained by investor response to differences in DRS. Moreover, the magnitude of empirical DRS-FSP relation estimates is quantitatively consistent with what our simple model would predict when calibrated to match the empirical fund size distribution.

5.2 DRS and Fund Size

Another immediate implication of our model is that steeper DRS shrink fund size. Recall that, in the model, fund size is directly proportional to the ratio of perceived skill to perceived scalability (see equation (5)). That is, large funds either earn a high alpha on the first dollar and/or implement strategies that are highly scalable. We now investigate the significance of

the latter effect, controlling for the former.³⁵ Table 7 presents the results of this exercise.

To control for the effect of perceived skill, we sort funds into quintiles based on $\hat{a}_{it}^{Char}$, i.e., fund i 's estimated skill in month t . Skill is measured by the average of $\hat{a}_{i\tau} + \hat{b}_{it}^{Char} \log(q_{i\tau-1})$ over the prior sixty months, where $\hat{b}_{it}^{Char}$ is the characteristic component of fund i 's DRS estimate in month t (see Section 5.1.2). Within each $\hat{a}_{it}^{Char}$ quintile, we again sort funds into quintiles based on $\hat{b}_{it}^{Char}$. We conduct this double sort for funds in the same Morningstar category and the same age group.³⁶ This control procedure creates $\hat{b}_{it}^{Char}$ quintiles with comparable distributions of perceived skill, fund style, and fund age, thereby controlling for differences in skill, style, and age. For each $\hat{b}_{it}^{Char}$ quintile, we calculate the average fund size (log real AUM) at the end of month t , whose time-series average is reported in the last column of Table 7. The table also reports the average fund size for each of the 25 $\hat{a}_{it}^{Char} \times \hat{b}_{it}^{Char}$ portfolios.

The last row of Table 7 reports the difference in average fund size between the first and fifth $\hat{b}_{it}^{Char}$ quintiles in each column.³⁷ The differences in fund size in the bottom right entry indicate that the real AUM of funds perceived to face steepest DRS is typically 85% smaller than that of funds perceived to be relatively immune to diseconomies of scale. These differences have robust t -statistics more negative than -14 . Moreover, fund size declines in an almost perfectly monotonic fashion from the lowest $\hat{b}_{it}^{Char}$ quintile to the highest $\hat{b}_{it}^{Char}$ quintile (reading down each column).

³⁵ Given that we control for perceived skill, the process of fund size adjustments that compete away alpha in equation (11) is necessary but not sufficient for fund size to increase with scalability. The pattern further requires that our DRS measure capture variation in scalability that is largely independent of perceived skill. Hence, the empirical confirmation of this pattern would demonstrate that scalability represents an important dimension of fund heterogeneity, distinct from skill.

³⁶ Specifically, we assign funds to three groups based on fund age: $[0, 5]$, $(5, 10]$, and > 10 years.

³⁷ To account for strong persistence in fund size, we calculate standard errors for these differences using 60 Newey-West lags.

Hence, steeper DRS shrink fund size, consistent with the above prediction of our model. This effect is both economically and statistically significant. Also noteworthy is that the real AUM of funds with higher perceived skill is typically larger than that of funds with lower perceived skill (reading from left to right), again consistent with our model.

In Appendix D, we further explore whether fund size converges toward an optimal level over time: if investors learn about funds as in the model, their perception of skill and scalability should become more accurate as funds grow older, implying that fund size should become more closely aligned with “optimal” fund size—the size at which the net alpha is driven to zero—over a typical fund’s lifetime. We find empirical support for this prediction. A detailed discussion of this convergence result is relegated to the appendix, however, as it relies on a stronger assumption that the optimal fund size is time-invariant.

6 Conclusion

One important feature that determines the value of an investment opportunity is its degree of scalability. In this paper, we empirically study the scalability of investment projects in the context of actively managed mutual funds. One common assumption in that literature is that all investment managers face the same degree of DRS while differing in the marginal profitability (gross alpha) on the first dollar invested. In this paper, we show that that assumption does not hold in the data. Heterogeneity in the degree of scalability is a key determinant of investors’ capital allocation decisions. Not only do we find that steeper DRS attenuate flow sensitivity to performance, we also find that differences in DRS across funds are quantitatively important for explaining capital allocation in the market for mutual funds.

This heterogeneity is thus a key driver of the cross-sectional distribution of fund size (AUM).

Appendix

A Necessary–Sufficient Condition for Our Hypothesis

Let R_{it}^n denote the return in excess of the risk-free rate earned by fund i 's investors at time t and let R_{it}^B denote the excess return of the manager's benchmark over the same time interval.

At time t , the investor observes the manager's net return outperformance,

$$\alpha_{it+1} \equiv R_{it}^n - R_{it}^B. \quad (\text{A1})$$

We assume throughout that α_{it} can be expressed as follows:

$$\alpha_{it} = a_i - b_i h(q_{it-1}) + \epsilon_{it}, \quad (\text{A2})$$

where q_{it-1} denotes the size (i.e., real AUM) of fund i at time $t - 1$, a_i denotes a parameter that captures fund i 's gross alpha on the first dollar net of the percentage fee its manager charges, and ϵ_{it} is the noise in observed performance. Here $b_i h(q)$ captures the decreasing returns to scale (DRS) the manager faces, which can vary by fund: $b_i > 0$ is a parameter that captures the cross-sectional variation in DRS technology and $h(q)$ is a strictly increasing function of q , which determines the form of DRS technology common across all funds.

Now note that α_{it} is an informative signal about a_i : high α_{it} implies good news about a_i and low α_{it} implies bad news about a_i . Formally, this is equivalent to assuming that the conditional probability density of α_{it} at time $t - 1$, $f_{t-1}(\alpha_{it})$, satisfies the monotone likelihood

ratio property— $f_{t-1}(\alpha_{it} | a_i) / f_{t-1}(\alpha_{it} | a_i^c)$ is increasing (decreasing) in α_{it} if $a_i \geq (\leq) a_i^c$ —which ensures that

$$\frac{\partial \theta_{it}}{\partial \alpha_{it}} > 0 \quad (\text{A3})$$

(see Milgrom (1981)). Thus, at time t , investors use the time- t information set I_t to update their beliefs on a_i implying that the expectation of a_i at time t is:

$$\theta_{it} \equiv E[a_i | I_t]. \quad (\text{A4})$$

Note that $q_{it}, R_{it}^n, R_{it}^B$ are elements of I_t . Let $\bar{\alpha}_{it}(q)$ denote investors' subjective expectation of α_{it+1} when fund i has size q at time t , i.e., fund i 's *net alpha*:

$$\bar{\alpha}_{it}(q) = \theta_{it} - b_i h(q). \quad (\text{A5})$$

In equilibrium, the size of the fund q_{it} adjusts to ensure that there are no positive NPV investment opportunities so $\bar{\alpha}_{it}(q_{it}) = 0$ and

$$\frac{\theta_{it}}{b_i} = h(q_{it}). \quad (\text{A6})$$

The following lemma shows how q_{it} depends on the information in α_{it} or the parameter b_i .

Lemma 2

$$\frac{\partial q_{it}}{\partial \alpha_{it}} = \frac{1}{b_i h'(q_{it})} \frac{\partial \theta_{it}}{\partial \alpha_{it}} \quad (\text{A7})$$

$$\frac{\partial q_{it}}{\partial b_i} = -\frac{h(q_{it})}{b_i h'(q_{it})} \quad (\text{A8})$$

Proof. First, note that $s_{it} = a_i + \epsilon_{it}$ corresponds to the new information about fund skill contained in α_{it} . Since rescaling the fund's DRS technology—changing the parameter b_i —does not change the signal s_{it} , we can conclude that

$$\frac{\partial \theta_{it}}{\partial b_i} = 0. \quad (\text{A9})$$

Now differentiating (A6) with respect to s_{it} , using the Inverse Function Theorem, and using the fact that these signals are independent of b_i ($\partial b_i / \partial s_{it} = 0$), gives

$$\frac{\partial q_{it}}{\partial s_{it}} = \frac{1}{h'(q_{it})} \frac{\partial (\theta_{it}/b_i)}{\partial s_{it}} = \frac{1}{b_i h'(q_{it})} \frac{\partial \theta_{it}}{\partial s_{it}}.$$

That s_{it} has a unit beta with respect to the realized performance ($\partial s_{it} / \partial \alpha_{it} = 1$) then gives (A7). Similarly, differentiate (A6) with respect to b_i , use the Inverse Function Theorem, and use (A9) to substitute for $\partial \theta_{it} / \partial b_i$ in this expression. This gives (A8). ■

Next, let the flow of capital into mutual fund i at time t be denoted by F_{it} :

$$F_{it} \equiv \log(q_{it}/q_{it-1}).$$

Differentiating this expression with respect to α_{it} ,

$$\frac{\partial F_{it}}{\partial \alpha_{it}} = \frac{1}{q_{it}} \frac{\partial q_{it}}{\partial \alpha_{it}} = \frac{1}{q_{it+1}} \frac{1}{b_i h'(q_{it+1})} \frac{\partial \theta_{it}}{\partial \alpha_{it}} > 0, \quad (\text{A10})$$

where the second equality follows from (A7) and the inequality follows from (A3), so good (bad) performance results in an inflow (outflow) of funds. This result is one of the important insights from Berk and Green (2004).

We are now ready to state the necessary and sufficient condition on the DRS technology for our hypothesis. Taking the derivative of the flow-performance sensitivity with respect to b_i , steeper DRS lead to a weaker flow sensitivity to performance (FSP) if and only if

$$\frac{\partial}{\partial b_i} \left(\frac{\partial F_{it}}{\partial \alpha_{it}} \right) < 0, \quad (\text{A11})$$

We show in Proposition 3 that (A11) is equivalent to

$$\frac{\partial}{\partial q_{it}} \left(\frac{\partial \log(h(q_{it}))}{\partial \log(q_{it})} \right) < 0, \quad (\text{A12})$$

which means that the size elasticity of performance is decreasing in fund size. This assumption is satisfied for many functional forms, including the logarithmic specification ($h(q) = \log(q)$) commonly used in empirical studies. In practice, such “concavity” of DRS technology can arise endogenously from funds changing their investment behavior as they grow: indeed, prior studies find that larger funds trade less and hold more-liquid stocks to mitigate the performance erosion due to diseconomies of scale.³⁸ This leads to the following

³⁸See, e.g., Pollet and Wilson (2008), Pástor, Stambaugh, and Taylor (2020), and Busse et al. (2021).

proposition, which is our hypothesis that we take to the data.

Proposition 3 *Steeper DRS lead to a smaller FSP if and only if condition (A12) holds.*

Proof.

$$\begin{aligned}
\frac{\partial}{\partial b_i} \left(\frac{\partial F_{it}}{\partial \alpha_{it}} \right) &= \frac{\partial}{\partial b_i} \left(\frac{1}{q_{it}} \frac{1}{b_i h'(q_{it})} \right) \frac{\partial \theta_{it}}{\partial \alpha_{it}} \\
&= - \frac{q_{it} h'(q_{it}) + \frac{\partial q_{it}}{\partial b_i} (b_i h'(q_{it}) + q_{it} b_i h''(q_{it}))}{q_{it}^2 (b_i h'(q_{it}))^2} \frac{\partial \theta_{it}}{\partial \alpha_{it}} \\
&= - \frac{q_{it} h'(q_{it}) - h(q_{it}) \left(1 + \frac{q_{it} h''(q_{it})}{h'(q_{it})} \right)}{q_{it}^2 (b_i h'(q_{it}))^2} \frac{\partial \theta_{it}}{\partial \alpha_{it}}, \tag{A13}
\end{aligned}$$

where the first equality is implied by expression (A10) and the fact that $\frac{\partial}{\partial b_i} \left(\frac{\partial \theta_{it}}{\partial \alpha_{it}} \right) = 0$ (since θ_{it} is solely a function of the history of realized signals and is not a function of b_i), and the last equality invokes expression (A8). What (A13) combined with (A3) tells us is that steeper DRS must lead to a smaller flow of funds response to performance if and only if

$$q_{it} h'(q_{it}) - h(q_{it}) \left(1 + \frac{q_{it} h''(q_{it})}{h'(q_{it})} \right) > 0. \tag{A14}$$

Condition (A14) is equivalent to

$$\frac{h'(q_{it})}{(h(q_{it}))^2} \times \left[q_{it} h'(q_{it}) - h(q_{it}) \left(1 + \frac{q_{it} h''(q_{it})}{h'(q_{it})} \right) \right] > \frac{h'(q_{it})}{(h(q_{it}))^2} \times 0 \tag{A15}$$

because $h(q)$ is a strictly increasing function of q , ensuring that $h'(q_{it}) > 0$. Notice that the

left-hand side of (A15) is equal to $-\frac{\partial}{\partial q_{it}} \left(\frac{\partial \log(h(q_{it}))}{\partial \log(q_{it})} \right)$, so (A15) can be rewritten as

$$-\frac{\partial}{\partial q_{it}} \left(\frac{\partial \log(h(q_{it}))}{\partial \log(q_{it})} \right) > 0,$$

which is also equivalent to (A12), completing the proof. ■

B Log vs. Linear Specifications for DRS Technology

The key assumption underpinning the main hypothesis is concavity in the DRS technology. This assumption is satisfied for the log specification ($h(q) = \log(q)$) in the paper, which is widely used in empirical studies.³⁹ In practice, such “concavity” of DRS technology can arise endogenously from funds changing their investment behavior as they grow: indeed, prior studies find that larger funds trade less and hold more-liquid stocks to mitigate the performance erosion due to diseconomies of scale.⁴⁰ But this assumption is not satisfied for the linear model ($h(q) = q$), another commonly used functional form of DRS technology in the size-performance analysis.⁴¹ To the best of our knowledge, none of these studies has tested whether the log specification is a more (or less) suitable functional form of decreasing returns to scale than the linear model. We demonstrate that the data support the logarithmic specification for DRS technology over a linear one in this appendix.

Recall that, for each fund i at time t , we estimate the regression $\alpha_{it} = a_i - b_i \log(q_{it-1}) + \epsilon_{it}$ using sixty months of the fund’s data before time t , which we use to compute the 1-month ahead forecast error based on the log specification $\hat{\epsilon}_{it} = \alpha_{it} - [\hat{a}_{it} - \hat{b}_{it} \log(q_{it-1})]$. There is

³⁹See, e.g., Chen et al. (2004), Yan (2008), Ferreira et al. (2013), and Zhu (2018).

⁴⁰See, e.g., Pollet and Wilson (2008), Pástor, Stambaugh, and Taylor (2020), and Busse et al. (2021).

⁴¹See, e.g., Pástor, Stambaugh, and Taylor (2015) and Barras, Gagliardini, and Scaillet (2022).

an analogous calculation of forecast errors for the linear model that we compare with $\hat{\epsilon}_{it}$: for each fund i at time t , we estimate the regression $\alpha_{it}^{\text{lin}} = a_i^{\text{lin}} - b_i^{\text{lin}} q_{it-1} + \epsilon_{it}^{\text{lin}}$ using sixty months of the fund's data before time t , which we use to compute the 1-month ahead forecast error based on the linear model $\hat{\epsilon}_{it}^{\text{lin}} = \alpha_{it}^{\text{lin}} - [\hat{a}_{it}^{\text{lin}} - \hat{b}_{it}^{\text{lin}} q_{it-1}]$.

A commonly used measure to quantify the forecasting performance of statistical models is the root mean square of forecast errors (RMSFE) in an out-of-sample exercise, which is given by

$$\text{RMSFE} = \sqrt{\frac{\sum_{it} e_{it}^2}{\# \text{Obs}}},$$

where e_{it} denotes the forecast error using the log specification or the linear model. We find that, measuring return outperformance relative to the CAPM (Vanguard benchmark), the log specification, with an RMSFE of 0.02216 (0.01619), outperforms the linear model, which has a higher RMSFE of 0.02283 (0.01659). Another measurements of forecasting performance is the mean absolute forecast errors

$$\text{MAFE} = \frac{\sum_{it} |e_{it}|}{\# \text{Obs}}.$$

Again, we find that, measuring return outperformance relative to the CAPM (Vanguard benchmark), the log specification, with an MAFE of 0.01598 (0.01150), outperforms the linear model, which has a higher MAFE of 0.01663 (0.01189). In short, the log specification is a better description of DRS technology than the linear model.

The same conclusions continue to hold if we were to assume away heterogeneity in the size-performance relation across funds following much of the literature: when we estimate

panel regressions $\alpha_{it} = a_i - b_i \log(q_{it-1}) + \epsilon_{it}$ and $\alpha_{it}^{\text{lin}} = a_i^{\text{lin}} - b_i^{\text{lin}} q_{it-1} + \epsilon_{it}^{\text{lin}}$ measuring return outperformance relative to the CAPM (Vanguard benchmark),⁴² their root mean square errors are 0.02239 (0.01603) and 0.02251 (0.01607), respectively, so the log specification has a smaller RMSE and still outperforms the linear model.

C Estimation Procedure for Fund-Specific DRS

This appendix describes the details of how we estimate fund-specific b_i parameters in the time-series regression $\alpha_{it} = a_i - b_i \log(q_{it-1}) + \epsilon_{it}$. It is well known that the OLS estimators of the coefficients b_i are subject to a small-sample bias. The small sample-bias arises because of the flow-performance relation, which induces a positive correlation between the regression disturbance ϵ_{it} and the innovation in $\log(q_{it})$: if $\log(q_{it})$ obeys an AR(1) process,

$$\log(q_{it}) = \chi_i + \rho_i \log(q_{it-1}) + v_{it}, \quad (\text{C1})$$

Stambaugh (1999) shows that $\widehat{b}_i^{OLS}$ is upward biased, and proposes a first-order bias-corrected estimator of b_i . Amihud and Hurvich (2004, hereafter “AH”) improve upon this estimator by noting that adding a proxy v_{it}^c for the innovations in the autoregressive model can reduce the small-sample bias. The proxy v_{it}^c takes the form, $v_{it}^c = \log(q_{it}) - (\widehat{\chi}_i^c + \widehat{\rho}_i^c \log(q_{it-1}))$, where $\widehat{\chi}_i^c$ and $\widehat{\rho}_i^c$ are any estimators of χ_i and ρ_i constructed based on size data. We adopt this estimation procedure, except we use a different estimator of ρ_i than AH.⁴³ Specifically, for each month t ,

⁴²We estimate these panel regressions by applying the recursive demeaning procedure of Zhu (2018).

⁴³This choice is guided by a horse race among various ρ_i estimators for recovering the b_i coefficients using simulations from our calibrated model in the Online Appendix.

1. $\hat{\rho}^{OLS\ FE}$ (the coefficient of $\log(q_{i\tau-1})$ in a panel regression of $\log(q_{i\tau})$ on $\log(q_{i\tau-1})$ with fund fixed effects based on sixty months of size data for all funds prior to month t) is used to construct the panel median-unbiased estimator (PMUE) $\hat{\rho}_t^c$ of the size's persistence ρ_i as $\hat{\rho}_t^c = m^{-1}(\hat{\rho}^{OLS\ FE})$, where $m(\rho)$ is the unique median of $\hat{\rho}^{OLS\ FE}$ when the true $\rho \in (-1, 1)$ is homogeneous across funds.⁴⁴
2. the proxy for size innovations is $v_{i\tau}^c = \log(q_{i\tau}) - (\hat{\chi}_{it}^c + \hat{\rho}_t^c \log(q_{i\tau-1}))$, where $\hat{\chi}_{it}^c$ is chosen to ensure that $\{v_{i\tau}^c\}$ has zero mean for each fund i ; and
3. $\hat{b}_{it}$ is the coefficient of $-\log(q_{i\tau-1})$ in the time-series regression of $\alpha_{i\tau}$ on $-\log(q_{i\tau-1})$ and $v_{i\tau}^c$, with intercept, using data from months $\tau = t - 60, \dots, t - 1$ for each fund i .

D DRS and Optimal Fund Size

If investors learn about funds as in the model, their perception of optimal size should converge to the true optimal size over a typical fund's lifetime. This logic predicts that the (log) fund size should converge to the (log) optimal size—the ratio of true skill to true scalability—as funds grow older.⁴⁵ In this appendix, we test this prediction and find empirical support for it.

As noted earlier, we find that 36% of the fund-by-fund b_i estimates (i.e., $\hat{b}_{it}$) are negative. This has not been an issue so far because our analysis has relied on the *relative* steepness of the DRS technology. However, it is a problem for estimating the fund's optimal size, which requires $b_i > 0$ since, theoretically, all funds must face DRS in equilibrium. A simple way

⁴⁴Key to this estimator is the fact that the distribution of $\hat{\rho}^{OLS\ FE}$ depends only on ρ when the homogeneous dynamic panel model is correct (see Phillips and Sul (2003)).

⁴⁵Note that the optimal fund size here is the size at which the net alpha is driven to zero.

to deal with this econometric defect is to “shrink” the b_i estimates toward their prior mean, specifically, the homogeneous fund-level DRS parameter b estimated using the recursive demeaning (RD) procedure of Zhu (2018).⁴⁶ The RD estimate $\widehat{b}^{RD2}$, which is statistically significant, indicates that an 1% increase in fund size results in a decrease in a fund’s CAPM alpha of 0.47 bp per month.⁴⁷ All of the “shrinkage” b_i estimates, denoted by $\widehat{b}_{it}^{Shr}$, are positive.⁴⁸ Then, the corresponding estimator of skill, denoted by $\widehat{a}_{it}^{Shr}$, is equal to the average of $\widehat{\alpha}_{i\tau} + \widehat{b}_{it}^{Shr} \log(q_{i\tau-1})$ over the prior sixty months. We measure fund i ’s (log) optimal size, $\log(q_i^*)$, by the average of the ratios $\widehat{a}_{it}^{Shr}/\widehat{b}_{it}^{Shr}$ over its lifetime.

This measure of optimal size can be different than what investors come to think of as optimal (ex-post) if they ignore individual heterogeneity in DRS—they use the homogeneous fund-level DRS parameter estimate, $\widehat{b}^{RD2}$. In this case, the corresponding estimator of skill, denoted by $\widehat{a}_{it}^{RD2}$, is equal to the average of $\widehat{\alpha}_{i\tau} + \widehat{b}^{RD2} \log(q_{i\tau-1})$ over the prior sixty months, and the alternative measure of fund i ’s optimal size, $\log(q_i^{*RD2})$, is given by the average of the ratios $\widehat{a}_{it}^{RD2}/\widehat{b}^{RD2}$ over its lifetime.

To test the above prediction, we examine how the relation between our measure of optimal size and fund size depends on fund age. Specifically, we assign funds to three groups based on fund age: $[0, 5]$, $(5, 10]$, and > 10 years. In each age-based sample, we run panel regressions of fund i ’s log real AUM in month t on its estimated log optimal size, $\log(\widehat{q}_i^*)$. We report

⁴⁶Pástor, Stambaugh, and Taylor (2015) develop an RD procedure to analyze returns to scale. They find coefficients indicative of fund-level DRS, albeit statistically insignificant. Zhu (2018) improves upon the empirical strategy in PST and establishes strong evidence of fund-level diseconomies of scale.

⁴⁷When we calculate outperformance relative to the Vanguard benchmark, the RD estimate $\widehat{b}^{RD2}$, which is again statistically significant, indicates that an 1% increase in size results in a decrease in expected fund performance of 0.15 bp per month.

⁴⁸Formally, the shrinkage estimator of b_i is $\widehat{b}_{it}^{Shr} = w_{it}\widehat{b}_{it} + (1 - w_{it})\widehat{b}^{RD2}$, where $w_{it} = \frac{1/\sigma_{b_{it}}^2}{1/\sigma_{b_{it}}^2 + 1/\kappa^2}$, $\sigma_{b_{it}}$ is the standard error of $\widehat{b}_{it}$, and κ is a constant controlling the amount of shrinkage toward $\widehat{b}^{RD2}$. We set this constant to ensure that the resulting $\widehat{b}_{it}^{Shr}$ values are positive.

the results in the first three columns of Table 8.⁴⁹

The slopes on $\log(q_i^*)$ are consistently significantly positive, and their magnitude increases as we move from the samples of young funds to middle-aged funds, and, then, to old funds. In addition, the R^2 of the regressions increases monotonically as we move from the samples of young funds to old funds. In short, as funds get older, the estimated optimal size plays an increasingly important role in capital allocation.

In columns 4 through 6 of Table 8, we run multiple regressions of $\log(q_{it})$ on both $\log(q_i^*)$ and $\log(q_i^{*RD2})$ in all three age-sorted samples. While the slopes on $\log(q_i^*)$ become much smaller, they continue to be significantly positive in the samples of middle-aged and old funds, and their magnitude increases monotonically as we move from the samples of young funds to old funds. In contrast, although the slopes on $\log(q_i^*)$ are significantly positive, their magnitude decreases monotonically as we move from the samples of young funds to old funds. Taken together, these results suggest that investors allocate capital to older funds using the sophisticated measure of optimal size, but allocate capital to young funds using the simple measure of optimal size. Consistent with this interpretation, the R^2 improvement from adding $\log(q_i^{*RD2})$ decreases as we move from the samples of young funds to middle-aged funds, and, then, to old funds (for which the R^2 remains about the same).

Our results offer the following narrative. Investors want to account for DRS heterogeneity, but they need to learn about fund-specific values. Given that such fund-specific information is not yet available for young funds, investors use the sample-wide b instead, and they only use fund-specific b_i in making their capital allocation decisions when funds grow old enough

⁴⁹Table 8 reports the double-clustered (by fund and time) standard errors.

such that the remaining Bayesian uncertainty on these values is relatively modest. Thus, it seems that investors are learning about not only skill but also scalability. Future research can explore the capital allocation implications of learning about fund heterogeneity in DRS.

E Variable Definitions

Variable Names	Variable Definitions
Net return (R_{it})	Return received by investors (in units of fraction per month)
Net alpha ($\hat{\alpha}_{it}^{\text{CAPM}}$)	Net return minus the return on benchmark portfolio, constructed using the CAPM (in units of fraction per month)
Net alpha ($\hat{\alpha}_{it}^{\text{VG}}$)	Net return minus the return on benchmark portfolio, constructed using a set of Vanguard index funds (in units of fraction per month)
Total AUM (AUM_{it})	Nominal assets under management at the end of month t
Fund size (q_{it-1})	Total AUM at the end of the previous month, adjusted for inflation (in 2000 \$millions)
Log fund size ($\log(q_{it-1})$)	Natural logarithm of lagged real AUM
Flow, v.1 ($F_{it}^{\text{v}.1}$)	$\log(q_{it}/q_{it-1})$
Flow, v.2 ($F_{it}^{\text{v}.2}$)	$\frac{AUM_{it} - AUM_{it-1}(1+R_{it})}{AUM_{it-1}(1+R_{it})}$
Fund-specific DRS ($\hat{b}_{it}$)	Bias-corrected estimator of b_i in $\hat{\alpha}_{it} = a_i - b_i \log(q_{it-1}) + \epsilon_{it}$ using sixty months of data before month t
Fund-specific FSP ($\widehat{FSP}_{it}$)	OLS estimator of γ_i in $F_{it} = c_i - \gamma_i \left(\prod_{s=t-12}^{t-1} (1 + \hat{\alpha}_{is}) - 1 \right) + v_{it}$ using sixty months of data after month t

Variable Names	Variable Definitions
Expense ratio	Annual expense ratio as of the previous month-end (in units of fraction per year)
Marketing expenses	Expense ratio plus 1/7th of the up-front load fees (in units of fraction per year)
Star family affiliation	Dummy variable that is equal to one if a fund is affiliated with a star family—parent company of a fund with a five-star rating based on the 3-year Morningstar rating—but is not a star itself, and zero otherwise
Log family size	Natural logarithm of the sum of lagged real AUM across all funds within the same fund family
Diverse offerings	Dummy variable that is one if the number of different fund categories offered by a fund family is larger than the median number for all fund families in the same month, and zero otherwise
Fund age	Number of years since a fund's first offer date (from CRSP or, if missing, from Morningstar)
Log fund age	Natural logarithm of fund age
Volatility ($SD(\hat{\alpha}_{it})$)	Standard deviation of a fund's monthly net alpha estimates over the prior twelve months
Number of managers	Number of managers in a fund as of the previous month-end
Intl exposure	Indicator equal to one if a fund has significant international risk exposure, and zero otherwise (based on rejecting the joint null that its loadings on all three Vanguard international index funds are 0 at the 5% level)
Turnover	Average annual turnover (in units of fraction per year)
$\hat{\beta}_{it}^{mkt}$	Estimated market beta from the regression of a fund's return on the four Fama-French-Carhart factors over the prior sixty months
$\hat{\beta}_{it}^{smb}$	Estimated size loading from the regression of a fund's return on the four Fama-French-Carhart factors over the prior sixty months
$\hat{\beta}_{it}^{hml}$	Estimated value loading from the regression of a fund's return on the four Fama-French-Carhart factors over the prior sixty months
$\hat{\beta}_{it}^{mkt}$	Estimated momentum loading from the regression of a fund's return on the four Fama-French-Carhart factors over the prior sixty months

Variable Names	Variable Definitions
Characteristic component of DRS ($\widehat{b}_{it}^{Char}$)	Fitted value of $\widehat{b}_{it}$ from a panel regression on fund characteristics; see Table 4 for the various specifications of the regression
Perceived skill ($\widehat{a}_{it}^{Char}$)	Average of $\widehat{\alpha}_{i\tau} + \widehat{b}_{it}^{Char} \log(q_{i\tau-1})$ over the prior sixty months, where $\widehat{b}_{it}^{Char}$ is the characteristic component of DRS estimated using the specification in column 8 of Table 4
Log optimal size, v.1 ($\log(q_i^*)$)	Average of the ratios $\widehat{a}_{it}^{Shr} / \widehat{b}_{it}^{Shr}$ over a fund's lifetime, where $\widehat{b}_{it}^{Shr}$ is the shrinkage estimator of b_i (see footnote 47) and $\widehat{a}_{it}^{Shr}$ is equal to the average of $\widehat{\alpha}_{i\tau} + \widehat{b}_{it}^{Shr} \log(q_{i\tau-1})$ over the prior sixty months
Log optimal size, v.2 ($\log(q_i^{*RD2})$)	Average of the ratios $\widehat{a}_{it}^{RD2} / \widehat{b}_{it}^{RD2}$ over a fund's lifetime, where $\widehat{b}_{it}^{RD2}$ is the recursive demeaning (RD) estimator of the common fund-level DRS parameter b (see Zhu (2018)) and $\widehat{a}_{it}^{RD2}$ is equal to the average of $\widehat{\alpha}_{i\tau} + \widehat{b}_{it}^{RD2} \log(q_{i\tau-1})$ over the prior sixty months

References

- AMIHUD, Y., AND C. M. HURVICH (2004): “Predictive Regressions: A Reduced-Bias Estimation Method,” *Journal of Financial and Quantitative Analysis*, 39(4), 1897–1923.
- ANDREWS, D. W. K. (1993): “Exactly Median-Unbiased Estimation of First Order Autoregressive/Unit Root Models,” *Econometrica*, 61(1), 139–165.
- BARBER, B. M., X. HUANG, AND T. ODEAN (2016): “Which Factors Matter to Investors? Evidence from Mutual Fund Flows,” *Review of Financial Studies*, 29(10), 2600–2642.
- BARRAS, L., P. GAGLIARDINI, AND O. SCAILLET (2021): “Skill, Scale, and Value Creation in the Mutual Fund Industry,” *Journal of Finance*, Forthcoming.
- BERK, J. B., AND R. C. GREEN (2004): “Mutual Fund Flows and Performance in Rational Markets,” *Journal of Political Economy*, 112(6), 1269–1295.
- BERK, J. B., AND I. TONKS (2007): “Return Persistence and Fund Flows in the Worst Performing Mutual Funds,” Working Paper, 13042, NBER.
- BERK, J. B., AND J. H. VAN BINSBERGEN (2015): “Measuring Skill in the Mutual Fund Industry,” *Journal of Financial Economics*, 118(1), 1–20.
- BERK, J. B., AND J. H. VAN BINSBERGEN (2016): “Assessing Asset Pricing Models using Revealed Preference,” *Journal of Financial Economics*, 119(1), 1–23.
- BUSSE, J. A., AND P. J. IRVINE (2006): “Bayesian Alphas and Mutual Fund Persistence,” *Journal of Finance*, 61(5), 2251–2288.

- BUSSE, J. A., T. CHORDIA, L. JIANG, AND Y. TANG (2021): “Transaction Costs, Portfolio Characteristics, and Mutual Fund Performance,” *Management Science*, 67(2), 1227–1248.
- CARHART, M. M. (1997): “On Persistence in Mutual Fund Performance,” *Journal of Finance*, 52(1), 57–82.
- CHEN, J., H. HONG, M. HUANG, AND J. D. KUBIK (2004): “Does Fund Size Erode Mutual Fund Performance? The Role of Liquidity and Organization,” *American Economic Review*, 94(5), 1276–1302.
- CHEN, Q., I. GOLDSTEIN, AND W. JIANG (2010): “Payoff Complementarities and Financial Fragility: Evidence from Mutual Fund Outflows,” *Journal of Financial Economics*, 97(2), 239–262.
- CHEVALIER, J., AND G. ELLISON (1997): “Risk Taking by Mutual Funds as a Response to Incentives,” *Journal of Political Economy*, 105(6), 1167–1200.
- CHOI, J. J., AND A. Z. ROBERTSON (2020): “What Matters to Individual Investors? Evidence from the Horse’s Mouth,” *Journal of Finance*, 75(4), 1965–2020.
- CHRISTOFFERSEN, S. E. K. , R. EVANS, AND D. K. MUSTO (2013): “What Do Consumers’ Fund Flows Maximize? Evidence from Their Brokers’ Incentives,” *Journal of Finance*, 68(1), 201–235.
- DEL GUERCIO, D., AND J. REUTER (2013): “Mutual Fund Performance and the Incentive to Generate Alpha,” *Journal of Finance*, 69(4), 1673–1704.
- DYAKOV, T., H. JIANG, AND M. VERBEEK (2020): “Trade Less and Exit Overcrowded Markets: Lessons from International Mutual Funds,” *Review of Finance*, 24(3), 677–731.

- FAMA, E. F., AND K. R. FRENCH (2010): “Luck versus Skill in the Cross-Section of Mutual Fund Returns,” *Journal of Finance*, 65(5), 1915–1947.
- FERREIRA, M. A., A. KESWANI, A. F. MIGUEL, AND S. B. RAMOS (2012): “The Flow-Performance Relationship Around the World,” *Journal of Banking & Finance*, 36(6), 1759–1780.
- FERREIRA, M. A., A. KESWANI, A. F. MIGUEL, AND S. B. RAMOS (2013): “The Determinants of Mutual Fund Performance: A Cross-Country Study,” *Review of Finance*, 17(2), 483–525.
- FRANZONI, F., AND M. C. SCHMALZ (2017): “Fund Flows and Market States,” *Review of Financial Studies*, 30(8), 2621–2673.
- GALLAHER, S. T., R. KANIEL, AND L. T. STARKS (2009): “Advertising and Mutual Funds: From Families to Individual Funds,” Working Paper, University of Texas at Austin.
- HUANG, J., K. D. WEI, AND H. YAN (2007): “Participation Costs and the Sensitivity of Fund Flows to Past Performance,” *Journal of Finance*, 62(3), 1273–1311.
- HUANG, J., K. D. WEI, AND H. YAN (2012): “Investor Learning and Mutual Fund Flows,” Working Paper, University of Texas at Austin.
- KIM, J. H. (2022): “Investor Learning and the Aggregate Allocation of Capital to Active Management,” Working Paper, Florida State University.
- MAMAYSKY, H., M. SPIEGEL, AND H. ZHANG (2008): “Estimating the Dynamics of Mutual Fund Alphas and Betas,” *Review of Financial Studies*, 21(1), 233–264.
- MILGROM, P. R. (1981): “Good News and Bad News: Representation Theorems and Applications,” *Bell Journal of Economics*, 12(2), 380–391.

- PÁSTOR, L., AND R. F. STAMBAUGH (2012): “On the Size of the Active Management Industry,” *Journal of Political Economy*, 120(4), 740–781.
- PÁSTOR, L., R. F. STAMBAUGH, AND L. A. TAYLOR (2015): “Scale and Skill in Active Management,” *Journal of Financial Economics*, 116(1), 23–45.
- PÁSTOR, L., R. F. STAMBAUGH, AND L. A. TAYLOR (2020): “Fund Tradeoffs,” *Journal of Financial Economics*, 138(3), 614–634.
- PHILLIPS, P. C. B., AND D. SUL (2003): “Dynamic panel estimation and homogeneity testing under cross section dependence,” *The Econometrics Journal*, 6(1), 217–259.
- POLLET, J. M., AND M. WILSON (2008): “How Does Size Affect Mutual Fund Behavior?” *Journal of Finance*, 63(6), 2941–2969.
- ROUSSANOV, N., H. RUAN, AND Y. WEI (2021): “Marketing Mutual Funds,” *Review of Financial Studies*, 34(6), 3045–3094.
- SIRRI, E. R., AND P. TUFANO (1998): “Costly Search and Mutual Fund Flows,” *Journal of Finance*, 53(5), 1589–1622.
- STAMBAUGH, R. F. (1999): “Predictive Regressions,” *Journal of Financial Economics*, 54(3), 375–421.
- STAMBAUGH, R. F. (2020): “Skill and Profit in Active Management,” Working Paper, 26027, NBER.
- YAN, X. (2008): “Liquidity, Investment Style, and the Relation between Fund Size and Fund Performance,” *Journal of Financial and Quantitative Analysis*, 43(3), 741–767.

ZHU, M. (2018): “Informative Fund Size, Managerial Skill, and Investor Rationality,” *Journal of Financial Economics*, 130(1), 114–134.

Table 1: Summary Statistics

This table reports summary statistics for our sample of active equity mutual funds from 1991–2014. The unit of observation is the fund/month. The characteristic component of DRS ($\hat{b}_{it}^{Char}$) is estimated using the specification in column 8 of Table 4. Detailed definitions of all variables are provided in Appendix E.

	# of obs.	Mean	Stdev.	Min	10%	25%	50%	75%	90%	Max
Net return (R_{it})	394, 351	0.0075	0.0498	-0.1501	-0.0572	-0.0196	0.0122	0.0384	0.0645	0.1307
Net alpha ($\hat{\alpha}_{it}^{CAPM}$)	331, 450	0.0001	0.0225	-0.2355	-0.0237	-0.0105	-0.0002	0.0103	0.0244	0.3447
Net alpha ($\hat{\alpha}_{it}^{VG}$)	389, 721	-0.0002	0.0162	-0.2183	-0.0176	-0.0082	-0.0002	0.0078	0.0173	0.2202
Total AUM (AUM_{it})	394, 335	1196.6	4936.0	0.0072	15.607	49.9	186.3	714.13	2189.8	202306
Fund size (q_{it-1})	391, 615	1031.2	4146.8	0.02298	13.816	44.004	163.52	621.31	1908.0	163443
Log fund size ($\log(q_{it-1})$)	391, 615	5.1260	1.8870	-3.7731	2.6258	3.7843	5.0969	6.4318	7.5538	12.004
Flow, v.1 ($F_{it}^{v.1}$)	391, 611	0.0078	0.0735	-0.2281	-0.0761	-0.0288	0.0096	0.0447	0.0850	0.2711
Flow, v.2 ($F_{it}^{v.2}$)	391, 611	0.0049	0.0504	-0.1506	-0.0307	-0.0134	-0.0020	0.0136	0.0473	0.2705
Fund-specific DRS ($\hat{b}_{it}$)										
where $\hat{\alpha}_{it} = \hat{\alpha}_{it}^{CAPM}$	252, 433	0.0038	0.0153	-0.2178	-0.0112	-0.0027	0.0029	0.0100	0.0202	0.1372
where $\hat{\alpha}_{it} = \hat{\alpha}_{it}^{VG}$	300, 963	0.0023	0.0105	-0.1353	-0.0077	-0.0021	0.0017	0.0064	0.0131	0.1081
Fund-specific FSP ($\widehat{FSP}_{it}$)										
where $\hat{\alpha}_{it} = \hat{\alpha}_{it}^{CAPM}, F_{it} = F_{it}^{v.1}$	266, 376	0.0560	0.2758	-5.3838	-0.2276	-0.0846	0.0432	0.1907	0.3652	4.9262
where $\hat{\alpha}_{it} = \hat{\alpha}_{it}^{CAPM}, F_{it} = F_{it}^{v.2}$	266, 376	0.1045	0.1910	-10.950	-0.0495	0.0140	0.0756	0.1694	0.3035	5.8525
where $\hat{\alpha}_{it} = \hat{\alpha}_{it}^{VG}, F_{it} = F_{it}^{v.1}$	293, 895	0.0993	0.3938	-5.4942	-0.3236	-0.0988	0.0940	0.2973	0.5186	10.923
where $\hat{\alpha}_{it} = \hat{\alpha}_{it}^{VG}, F_{it} = F_{it}^{v.2}$	293, 895	0.1487	0.2898	-12.939	-0.0828	0.0171	0.1094	0.2499	0.4451	6.7011
Expense ratio	391, 631	0.0125	0.0043	0.0025	0.0077	0.0097	0.0120	0.0148	0.0182	0.0268
Marketing expenses	227, 378	0.0186	0.0060	0	0.0101	0.0151	0.0197	0.0218	0.0241	0.1705
Star family affiliation	374, 245	0.4371	0.4960	0	0	0	0	1	1	1
Log family size	393, 707	7.8592	2.5329	-2.6017	4.1335	6.1700	8.3542	9.6621	10.765	13.061
Diverse offerings	389, 135	0.4535	0.4978	0	0	0	0	1	1	1
Fund age	393, 909	12.916	13.038	0.6219	2.1753	4.4795	8.9507	15.970	28.036	67.841
Log fund age	393, 909	2.2849	0.8350	0.4836	1.1554	1.7010	2.2976	2.8314	3.3685	4.2318
Volatility ($SD(\hat{\alpha}_{it}^{CAPM})$)	303, 147	0.0188	0.0116	0.0005	0.0075	0.0105	0.0158	0.0239	0.0341	0.1270
Volatility ($SD(\hat{\alpha}_{it}^{VG})$)	360, 609	0.0138	0.0080	0.0007	0.0061	0.0084	0.0119	0.0172	0.0240	0.1140

	# of obs.	Mean	Stdev.	Min	Percentiles						Max
					10%	25%	50%	75%	90%		
Number of managers	382, 531	2.4144	2.1567	1	1	1	2	3	5	70	
Intl exposure	389, 721	0.2775	0.4477	0	0	0	0	1	1	1	
Turnover	375, 293	0.8366	0.7048	0.03	0.18	0.35	0.6421	1.1	1.71	3.84	
$\hat{\beta}_{it}^{mkt}$	331, 450	0.9670	0.1429	-0.0017	0.7969	0.8910	0.9707	1.0480	1.1302	1.9565	
$\hat{\beta}_{it}^{smb}$	331, 450	0.2157	0.3376	-0.8299	-0.1587	-0.0613	0.1314	0.4764	0.7178	1.6968	
$\hat{\beta}_{it}^{hml}$	331, 450	-0.0046	0.3100	-1.7424	-0.3875	-0.2214	-0.0020	0.2020	0.3905	1.7911	
$\hat{\beta}_{it}^{mkt}$	331, 450	0.0146	0.1374	-0.9930	-0.1362	-0.0598	0.0062	0.0774	0.1784	1.6558	
Characteristic component of DRS ($\hat{b}_{it}^{Char}$)											
where $\hat{\alpha}_{it} = \hat{\alpha}_{it}^{CAPM}$	154, 998	0.0047	0.0040	-0.0263	0.0005	0.0022	0.0042	0.0067	0.0095	0.0343	
where $\hat{\alpha}_{it} = \hat{\alpha}_{it}^{VG}$	180, 913	0.0029	0.0025	-0.0169	0.0002	0.0013	0.0027	0.0042	0.0058	0.0215	
Perceived skill ($\hat{a}_{it}^{Char}$)											
where $\hat{\alpha}_{it} = \hat{\alpha}_{it}^{CAPM}$	154, 998	0.0275	0.0243	-0.1005	0.0027	0.0114	0.0233	0.0392	0.0575	0.2158	
where $\hat{\alpha}_{it} = \hat{\alpha}_{it}^{VG}$	180, 913	0.0159	0.0151	-0.0599	0.0002	0.0059	0.0135	0.0234	0.0348	0.1261	
Log optimal size, v.1 ($\log(q_i^*)$)											
where $\hat{\alpha}_{it} = \hat{\alpha}_{it}^{CAPM}$	360, 749	5.5457	1.9439	-5.1121	3.0933	4.3132	5.6363	6.9069	7.9090	12.196	
where $\hat{\alpha}_{it} = \hat{\alpha}_{it}^{VG}$	376, 934	5.1418	2.6834	-24.326	2.0104	3.7603	5.3465	6.7776	8.1377	24.557	
Log optimal size, v.2 ($\log(q_i^{*RD2})$)											
where $\hat{\alpha}_{it} = \hat{\alpha}_{it}^{CAPM}$	382, 959	5.2923	1.9330	-10.918	2.7794	4.1123	5.4244	6.6087	7.5722	13.438	
where $\hat{\alpha}_{it} = \hat{\alpha}_{it}^{VG}$	384, 132	5.0197	2.4776	-11.373	2.0298	3.5850	5.1547	6.5689	8.0210	25.249	

Table 2: Relation Between DRS and Flow Sensitivity to Performance (FSP)

The table presents results from running panel regressions of fund i 's FSP going forward in month t , $\widehat{FSP}_{it}$, on its DRS estimated as of the previous month-end, $\widehat{b}_{it}$. Panel A shows the results when fund-specific DRS is estimated using the CAPM, and Panel B shows the results using Vanguard index funds as benchmark portfolios. In the odd columns, we only include month and fund fixed effects; in the even columns, we add proxies for participation costs, as well as performance volatility and fund age. Standard errors, two-way clustered by fund and by month, are in parentheses. We also report the EIV-adjusted coefficients and their standard errors in the last two row of each panel.

Panel A: Using Estimates Based on the CAPM				
	$\widehat{FSP}_{it}$			
$\widehat{b}_{it}$	−0.937 (0.148)	−1.299 (0.182)	−0.262 (0.087)	−0.386 (0.117)
Fund FE & Month FE	Yes	Yes	Yes	Yes
Controls	No	Yes	No	Yes
Observations	182, 675	114, 623	182, 675	114, 623
Flow as % Change in	Fund Size	Fund Size	New Assets	New Assets
EIV Adj. Coefficient	−9.143	−16.112	−2.560	−4.788
SE of the EIV Adj. Coefficient	0.375	0.752	0.224	0.365
Panel B: Using Estimates Based on the Vanguard BM				
	$\widehat{FSP}_{it}$			
$\widehat{b}_{it}$	−1.551 (0.288)	−1.714 (0.328)	−0.370 (0.199)	−0.549 (0.182)
Fund FE & Month FE	Yes	Yes	Yes	Yes
Controls	No	Yes	No	Yes
Observations	221, 743	136, 365	221, 743	136, 365
Flow as % Change in	Fund Size	Fund Size	New Assets	New Assets
EIV Adj. Coefficient	−40.984	−77.384	−9.772	−24.781
SE of the EIV Adj. Coefficient	2.786	9.906	1.323	3.658

Table 3: Relation Between Percentile Ranks of DRS and FSP

This table is the same as Table 2 but, instead of using $\hat{b}_{it}$ and $\widehat{FSP}_{it}$, uses their percentile ranks in each month. Standard errors, two-way clustered by fund and by month, are in parentheses.

Panel A: Using Estimates Based on the CAPM				
	Pctl. rank based on $\widehat{FSP}_{it}$			
Pctl. rank based on $\hat{b}_{it}$	-0.1083 (0.0109)	-0.1251 (0.0129)	-0.0659 (0.0095)	-0.0753 (0.0117)
Fund FE	Yes	Yes	Yes	Yes
Controls	No	Yes	No	Yes
Observations	182, 675	114, 623	182, 675	114, 623
Flow as % Change in	Fund Size	Fund Size	New Assets	New Assets
Panel B: Using Estimates Based on the Vanguard BM				
	Pctl. rank based on $\widehat{FSP}_{it}$			
Pctl. rank based on $\hat{b}_{it}$	-0.0918 (0.0095)	-0.0951 (0.0117)	-0.0560 (0.0086)	-0.0593 (0.0108)
Fund FE	Yes	Yes	Yes	Yes
Controls	No	Yes	No	Yes
Observations	221, 743	136, 365	221, 743	136, 365
Flow as % Change in	Fund Size	Fund Size	New Assets	New Assets

Table 4: Determinants of Fund-Level DRS

The table presents results from running panel regressions of fund i 's backward-looking DRS estimate in month t , $\hat{b}_{it}$, on the fund's characteristics as of the previous month-end. Panel A shows the results when fund-specific DRS is estimated using the CAPM, and Panel B shows the results using Vanguard index funds as benchmark portfolios. The first characteristic is the number of managers. The second characteristic is volatility: the standard deviation of fund alphas over the prior 1 year. The next two characteristics we examine are expense ratio and marketing expenses. The fifth characteristic is the international exposure dummy: for any given fund, it is equal to one if we reject the null hypothesis that the coefficients on three Vanguard international index funds are 0 at the 5% confidence level. The sixth characteristic is average annual turnover (from CRSP). The last characteristic is log real AUM. We also control for loadings on the market, size, value, and momentum factors to capture fund style and risk. Standard errors, two-way clustered by fund and by month, are in parentheses.

Panel A: Using Estimates Based on the CAPM		Dependent Variable: $\hat{b}_{it}$					
		Yes	Yes	Yes	Yes	Yes	Yes
Number of managers	-0.000254 (0.000070)						
Volatility ($SD(\hat{\alpha}_{it}^{CAPM})$)	0.1906 (0.0226)						0.1828 (0.0262)
Expense ratio	0.1622 (0.0457)						0.2128 (0.0689)
Marketing expenses				-0.0338 (0.0358)			-0.1323 (0.0440)
Intl exposure				0.000057 (0.000351)			
Turnover					0.000458 (0.000287)		
Log fund size						0.000041 (0.000109)	
Style & Risk Controls		Yes	Yes	Yes	Yes	Yes	Yes
Observations	248,076	252,361	252,427	154,998	252,433	247,256	252,413
							152,007

Panel B: Using Estimates Based on the Vanguard BM

	Dependent Variable: $\hat{b}_{it}$					
Number of managers	-0.000078 (0.000037)					
Volatility ($(\hat{\alpha}_{it}^{VG})$)	0.0512 (0.0185)					0.0703 (0.0217)
Expense ratio		0.0528 (0.0308)				0.0780 (0.0461)
Marketing expenses			-0.0344 (0.0278)			-0.0818 (0.0340)
Intl exposure				0.000006 (0.000236)		
Turnover					-0.000037 (0.000186)	
Log fund size						0.000045 (0.000069)
Style & Risk Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	294, 922	300, 883	300, 957	180, 913	294, 666	300, 941
						176, 756

Table 5: Relation Between the Characteristic Component of DRS and FSP

This table is similar to Table 2 but replaces $\hat{b}_{it}$ by its characteristic component, $\hat{b}_{it}^{Char}$, based on various first-stage specifications from Table 4. Standard errors, two-way clustered by fund and by month, are in parentheses.

Panel A: Using Estimates Based on the CAPM									
	$\widehat{FSP}_{it}$								
$\widehat{b}_{it}^{Char}$	-9.923 (1.927)	-9.111 (1.977)	-9.419 (1.888)	-11.143 (1.962)	-6.144 (1.237)	-6.177 (1.290)	-5.688 (1.218)	-6.236 (1.243)	
Log fund size	-0.004 (0.008)	-0.005 (0.008)	-0.006 (0.008)	-0.006 (0.008)	-0.020 (0.006)	-0.021 (0.006)	-0.021 (0.006)	-0.022 (0.006)	
Marketing expenses	-0.306 (0.979)	-0.547 (1.017)	-0.096 (0.986)	-1.485 (1.004)	-0.363 (0.599)	-0.469 (0.613)	-0.200 (0.600)	-1.000 (0.613)	
Star family affiliation	-0.006 (0.005)	-0.007 (0.005)	-0.007 (0.005)	-0.007 (0.005)	-0.012 (0.003)	-0.012 (0.003)	-0.012 (0.003)	-0.012 (0.003)	
Log family size	0.015 (0.011)	0.016 (0.011)	0.016 (0.011)	0.015 (0.011)	0.006 (0.007)	0.007 (0.007)	0.007 (0.007)	0.006 (0.007)	
Diverse offerings	0.001 (0.015)	-0.001 (0.014)	-0.000 (0.014)	-0.000 (0.014)	-0.002 (0.009)	-0.003 (0.009)	-0.003 (0.009)	-0.003 (0.009)	
Log fund age	-0.015 (0.030)	-0.015 (0.030)	-0.018 (0.030)	-0.018 (0.030)	-0.013 (0.017)	-0.014 (0.017)	-0.016 (0.017)	-0.016 (0.017)	
Volatility ($SD(\widehat{\alpha}_{it}^{CAPM}))$	0.559 (0.409)	2.179 (0.560)	0.522 (0.405)	2.549 (0.558)	-0.141 (0.245)	1.038 (0.334)	-0.099 (0.239)	1.027 (0.322)	
Fund FE & Month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	112,071	114,623	114,623	114,623	112,071	114,623	114,623	114,623	
Flow as % Change in	Fund Size	Fund Size	Fund Size	Fund Size	New Assets	New Assets	New Assets	New Assets	
$\widehat{b}_{it}^{Char}$ based on Table 4 A, Col.	1	2	3	8	1	2	3	8	
$SD(\widehat{b}_{it}^{Char})$	0.0023	0.0030	0.0024	0.0031	0.0023	0.0030	0.0024	0.0031	

Panel B: Using Estimates Based on the Vanguard BM

	$\widehat{FSP}_{it}$							
$\widehat{b}_{it}^{Char}$	-8.305 (3.492)	-7.409 (3.463)	-7.592 (3.398)	-7.413 (3.759)	-7.248 (2.133)	-7.199 (2.103)	-6.772 (2.059)	-7.206 (2.184)
Log fund size	-0.031 (0.009)	-0.033 (0.009)	-0.033 (0.009)	-0.033 (0.009)	-0.046 (0.007)	-0.048 (0.007)	-0.048 (0.007)	-0.048 (0.007)
Marketing expenses	-1.500 (1.327)	-1.612 (1.348)	-1.494 (1.339)	-2.007 (1.350)	-1.097 (0.785)	-1.107 (0.782)	-0.998 (0.775)	-1.491 (0.789)
Star family affiliation	-0.016 (0.007)	-0.017 (0.007)	-0.017 (0.007)	-0.017 (0.007)	-0.016 (0.004)	-0.017 (0.004)	-0.017 (0.004)	-0.017 (0.004)
Log family size	0.002 (0.013)	0.007 (0.013)	0.007 (0.013)	0.007 (0.013)	0.007 (0.008)	0.010 (0.008)	0.010 (0.008)	0.010 (0.008)
Diverse offerings	-0.004 (0.019)	-0.008 (0.019)	-0.008 (0.019)	-0.008 (0.019)	-0.001 (0.012)	-0.004 (0.012)	-0.003 (0.012)	-0.003 (0.012)
Log fund age	-0.020 (0.032)	-0.017 (0.032)	-0.018 (0.032)	-0.018 (0.032)	-0.010 (0.018)	-0.011 (0.017)	-0.011 (0.017)	-0.011 (0.017)
Volatility ($SD(\widehat{\alpha}_{it}^{VG})$)	-0.586 (0.720)	-0.148 (0.763)	-0.513 (0.708)	0.000 (0.802)	-0.058 (0.487)	0.336 (0.508)	-0.024 (0.471)	0.480 (0.523)
Fund FE & Month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	132, 802	136, 365	136, 365	136, 365	132, 802	136, 365	136, 365	136, 365
Flow as % Change in	Fund Size	Fund Size	Fund Size	Fund Size	New Assets	New Assets	New Assets	New Assets
$\widehat{b}_{it}^{Char}$ based on Table 4 B, Col.	1	2	3	8	1	2	3	8
$SD(\widehat{b}_{it}^{Char})$	0.0016	0.0016	0.0016	0.0017	0.0016	0.0016	0.0016	0.0017

Table 6: Simulation Exercise Quantifying the Role of Heterogeneity in DRS

Panel A shows the calibrated parameter values used in our simulation analysis. It also reports the moments we target in our calibration. We simulate 10,000 samples of fund alphas (α) and size (q) for 300 funds over 100 months. We first draw each fund's scalability b_i from $Beta(A, B) \times C$, and then draw its skill a_i from $N(\theta_0(b_i), \sigma_0)$, where $\theta_0(b_i) = \theta_{0a} + \theta_{0b}b_i$. Simulated alphas follow $\alpha_{it} = a_i - b_i \log(q_{it-1}) + \epsilon_{it}$, and size follows $\log(q_{it}/q_{it-1}) = (\alpha_{it}/b_i) / (0.0577^2/\sigma_0^2 + t)$. In each sample, we estimate the DRS-FSP relation, where $\hat{\gamma}_i$ is each fund's estimated FSP from $\log(q_{it}/q_{it-1}) = c_i + \gamma_i\alpha_{it} + v_{it}$. Panel B shows its summary statistics across the simulated samples. To quantify the role of heterogeneity in DRS for capital allocation, we construct a counterfactual by assuming investors wrongly believe that every fund exhibits the same, average degree of DRS: $b_i = 0.0047$ for all funds. Then, under the counterfactual assumption, we compute what the size of the fund would have been, q_{it}^C , for every i and t . In each sample, we calculate the R^2 from a regression of $\log(q_{it})$ on $\log(q_{it}^C)$. Panel C reports the results from the counterfactual simulations.

Panel A: Calibrated Parameters								
Variable	Value	Target	Emp. Value	Sim. Value				
$x \sim Beta(A, B), b_i = Cx$								
First shape parameter (A)	1.15	Mean of b_i	0.0047	0.0047				
Second shape parameter (B)	9.93	Std dev of b_i	0.0040	0.0040				
Scale parameter (C)	0.045	Skewness of b_i	1.13	1.38				
		Kurtosis of b_i	6.73	5.23				
$\theta_0(b_i) = \theta_{0a} + \theta_{0b}b_i$								
Prior mean for CRS funds (θ_{0a})	0.06%	Mean of $\log(q_{it})$	5.13	5.13				
Prior mean slope on DRS (θ_{0b})	4.59	Std dev of $\log(q_{it})$	1.89	1.87				
Prior uncertainty (σ_0)	0.05%	Mean of $\hat{\gamma}_i$	0.56	0.56				
Panel B: Simulated DRS-FSP Relation								
$\hat{\gamma}_i = k + \lambda b_i + u_i$								
	Percentiles							
Mean	2.5%	5%	25%	50%	75%	95%	97.5%	
$\hat{\lambda}$	-20.4	-62.7	-40.2	-17.6	-12.9	-10.1	-7.73	-7.13
Panel C: Capital Allocation Explained by Counterfactual								
$\log(q_{it}) = \pi_0 + \pi_1 \log(q_{it}^C) + e_{it}$								
	Percentiles							
Mean	2.5%	5%	25%	50%	75%	95%	97.5%	
$\hat{\pi}_0$	-681	-1118	-1106	-947	-683	-433	-220	-177
$\hat{\pi}_1$	145	38.5	47.7	92.9	146	202	236	238
R^2	0.463	0.123	0.155	0.298	0.464	0.639	0.745	0.753

Table 7: Relation Between DRS and Fund Size

We first sort funds into quintiles based on $\hat{a}_{it}^{Char}$, i.e., fund's estimated skill. Within each $\hat{a}_{it}^{Char}$ quintile, we again sort funds into quintiles based on $\hat{b}_{it}^{Char}$, i.e., the characteristic component of fund's DRS estimate. We conduct this double sort for funds in the same Morningstar category and the same age group. Panel A shows the results using estimates based on the CAPM, and Panel B shows the results based on Vanguard index funds as benchmark portfolios. The table reports the average fund size (log real AUM) for each of the 25 $\hat{a}_{it}^{Char} \times \hat{b}_{it}^{Char}$ portfolios. The column labeled "Average" reports the average fund size at the end of month t for each $\hat{b}_{it}^{Char}$ quintile, controlling for differences in perceived skill, style, and age. The row labeled "High-low" reports the difference in average fund size between the first and fifth $\hat{b}_{it}^{Char}$ quintiles in each column. We calculate standard errors for these differences using 60 Newey-West lags.

Panel A: Using Estimates Based on the CAPM						
Group	$\hat{a}_{it}^{Char}$ Quintiles					Average
	1 Low	2	3	4	5	
1 Low $\hat{b}_{it}^{Char}$	5.578	6.463	6.880	7.259	7.501	6.712
2	5.697	6.160	6.387	6.767	7.195	6.410
3	5.326	5.394	5.701	6.083	6.689	5.813
4	4.860	4.846	5.026	5.576	6.565	5.340
5 High $\hat{b}_{it}^{Char}$	3.911	3.973	4.247	4.899	6.461	4.658
High-low	-1.667 (0.173)	-2.490 (0.127)	-2.634 (0.143)	-2.361 (0.165)	-1.040 (0.206)	-2.054 (0.137)

Panel B: Using Estimates Based on the Vanguard BM						
Group	$\hat{a}_{it}^{Char}$ Quintiles					Average
	1 Low	2	3	4	5	
1 Low $\hat{b}_{it}^{Char}$	5.229	6.029	6.511	6.924	7.338	6.364
2	5.428	5.871	6.271	6.659	7.014	6.213
3	5.097	5.190	5.580	5.968	6.548	5.651
4	4.681	4.647	5.031	5.541	6.496	5.239
5 High $\hat{b}_{it}^{Char}$	3.824	3.888	4.261	4.963	6.349	4.612
High-low	-1.405 (0.071)	-2.140 (0.097)	-2.250 (0.117)	-1.961 (0.099)	-0.989 (0.171)	-1.752 (0.083)

Table 8: Relation Between Optimal Size and Fund Size

The table presents results from running panel regressions of fund i 's log real AUM in month t on the fund's estimated (log) optimal size, $\log(q_i^*)$, across three age-sorted samples. The optimal fund size is defined as the size at which the net alpha is driven to zero. The last three columns control for an alternative measure of optimal size, $\log(q_i^{*RD2})$, that ignores fund heterogeneity in DRS. Panel A shows the results using estimates based on the CAPM, and Panel B shows the results based on Vanguard index funds as benchmark portfolios. Standard errors, two-way clustered by fund and by month, are in parentheses. We also report the standard deviations of the regressors for each sub-sample in the last two rows of each panel.

Panel A: Using Estimates Based on the CAPM						
Dependent Variable: Log Real AUM						
$\log(q_i^*)$	0.367 (0.016)	0.674 (0.015)	0.905 (0.012)	-0.052 (0.021)	0.271 (0.023)	0.597 (0.037)
$\log(q_i^{*RD2})$				0.585 (0.025)	0.517 (0.023)	0.332 (0.038)
R^2	0.24	0.63	0.78	0.39	0.71	0.79
Observations	84,310	100,048	176,375	84,310	100,048	176,375
Fund ages	[0, 5] yr.	(5, 10] yr.	> 10 yr.	[0, 5] yr.	(5, 10] yr.	> 10 yr.
Subsample SD($\log(q_i^*)$)	1.952	1.917	1.844	1.952	1.917	1.844
Subsample SD($\log(q_i^{*RD2})$)				2.021	1.760	1.791

Panel B: Using Estimates Based on the Vanguard BM						
Dependent Variable: Log Real AUM						
$\log(\hat{q}_i^*)$	0.203 (0.011)	0.416 (0.028)	0.659 (0.024)	0.031 (0.011)	0.181 (0.025)	0.361 (0.038)
$\log(\hat{q}_i^{*RD2})$				0.300 (0.014)	0.354 (0.024)	0.340 (0.038)
R^2	0.17	0.44	0.65	0.31	0.54	0.68
Observations	98,088	102,325	176,505	98,088	102,325	176,505
Fund ages	[0, 5] yr.	(5, 10] yr.	> 10 yr.	[0, 5] yr.	(5, 10] yr.	> 10 yr.
Subsample SD($\log(\hat{q}_i^*)$)	3.089	2.604	2.315	3.089	2.604	2.315
Subsample SD($\log(\hat{q}_i^{*RD2})$)				2.689	2.309	2.235

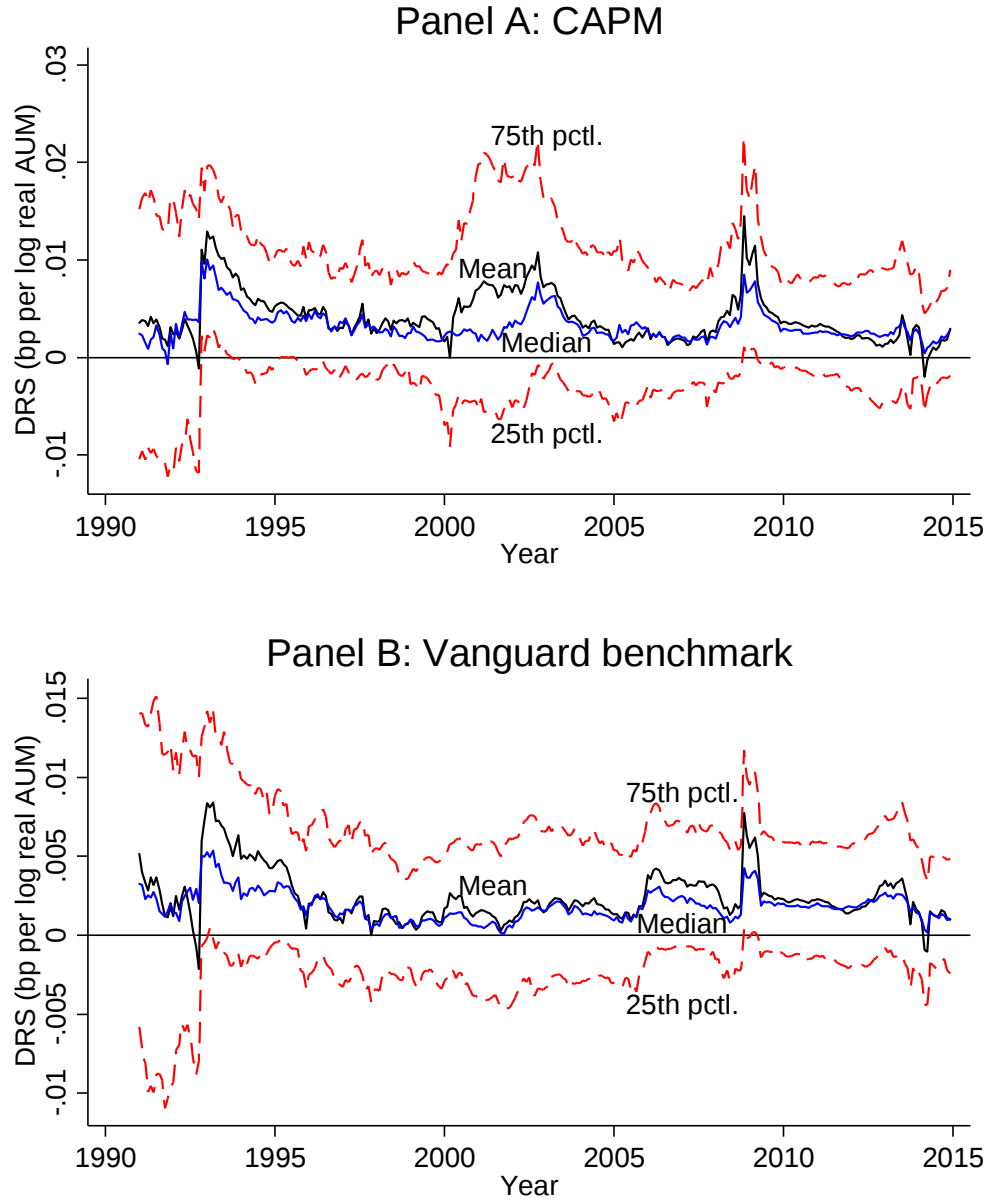


Figure 1: **Cross-Sectional Distribution of Fund-Specific DRS Estimates Over Time.** The figure displays how the cross-sectional distribution of $\hat{b}_{it}$ —fund i 's DRS estimated using sixty months of data before month t —varies over time. Panel A shows the plot when fund-specific DRS is estimated using the CAPM, and Panel B shows the plot using Vanguard index funds as benchmark portfolios.

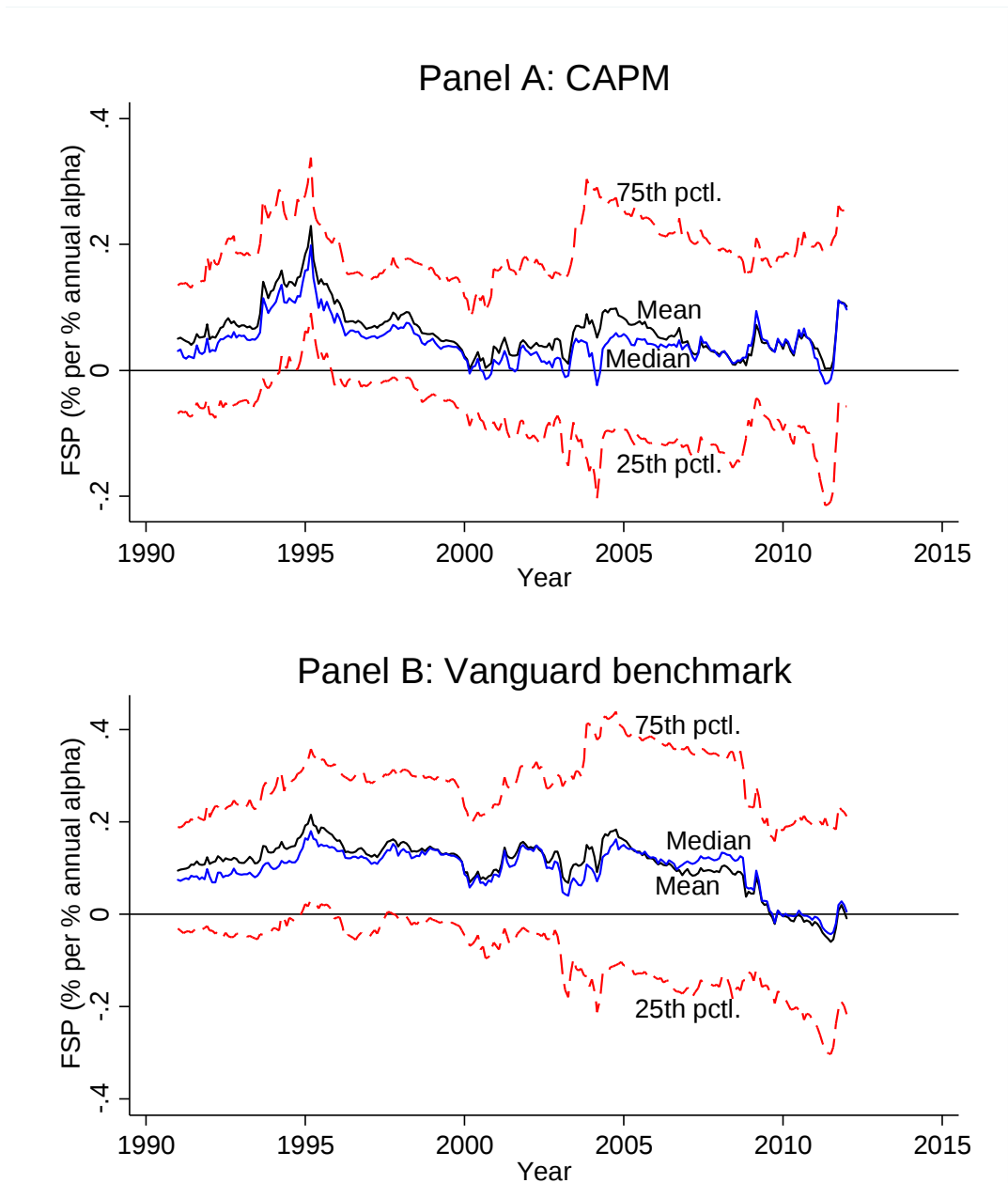


Figure 2: **Cross-Sectional Distribution of Fund-Specific FSP Estimates Over Time When Flow Is Measured As Percentage Change in Fund Size.** The figure displays the distributions of $\widehat{FSP}_{it}$ —the fund’s FSP estimated using its data over the subsequent 5 years—over time. Panel A shows the plot when fund-specific FSP is estimated using the CAPM, and Panel B shows the plot using Vanguard index funds as benchmark portfolios.

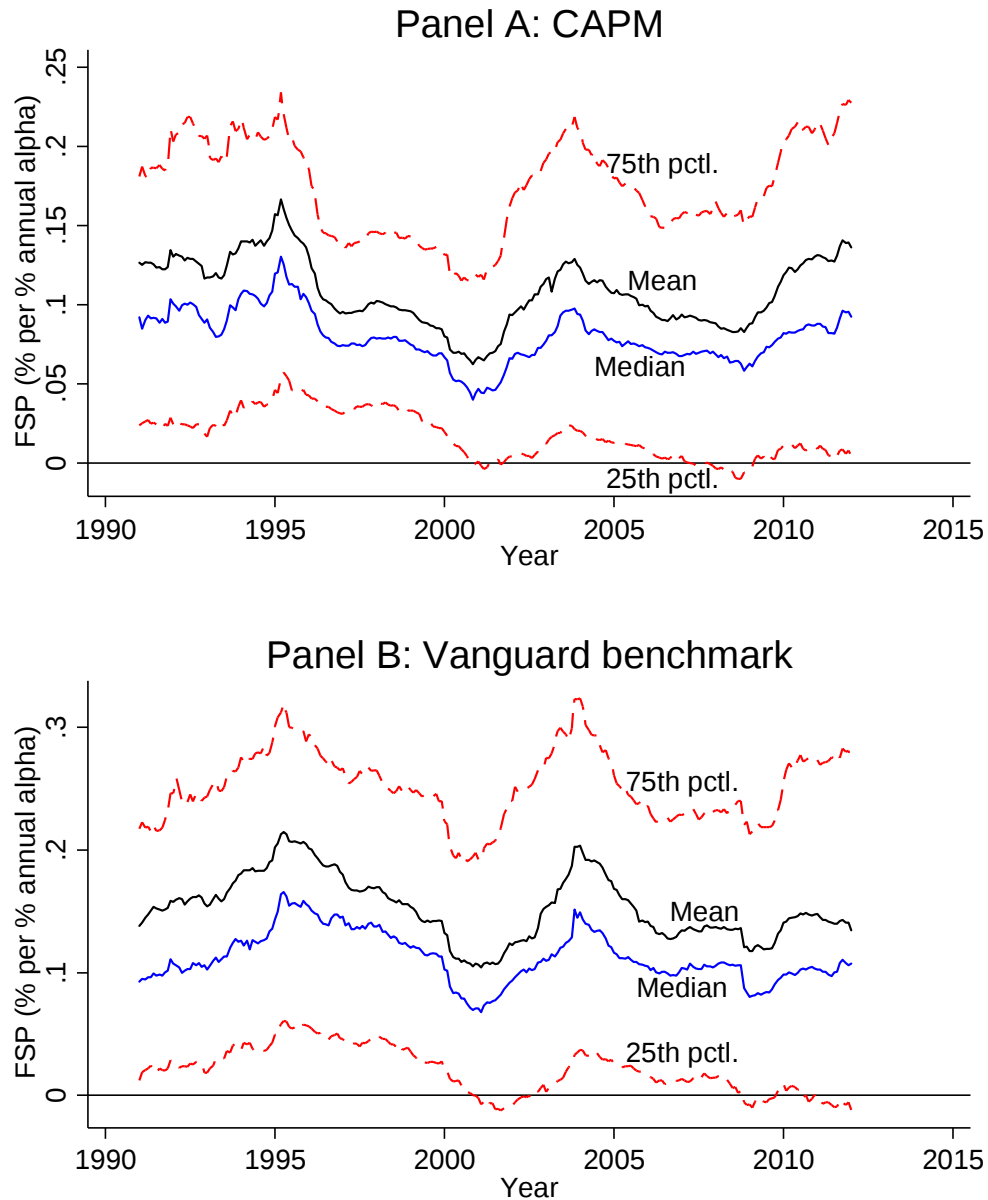


Figure 3: **Cross-Sectional Distribution of Fund-Specific FSP Estimates Over Time When Flow Is Measured As Percentage Change in New Assets.** The figure displays the distributions of $\widehat{FSP}_{it}$ —the fund’s FSP estimated using its data over the subsequent 5 years—over time. Panel A shows the plot when fund-specific FSP is estimated using the CAPM, and Panel B shows the plot using Vanguard index funds as benchmark portfolios.

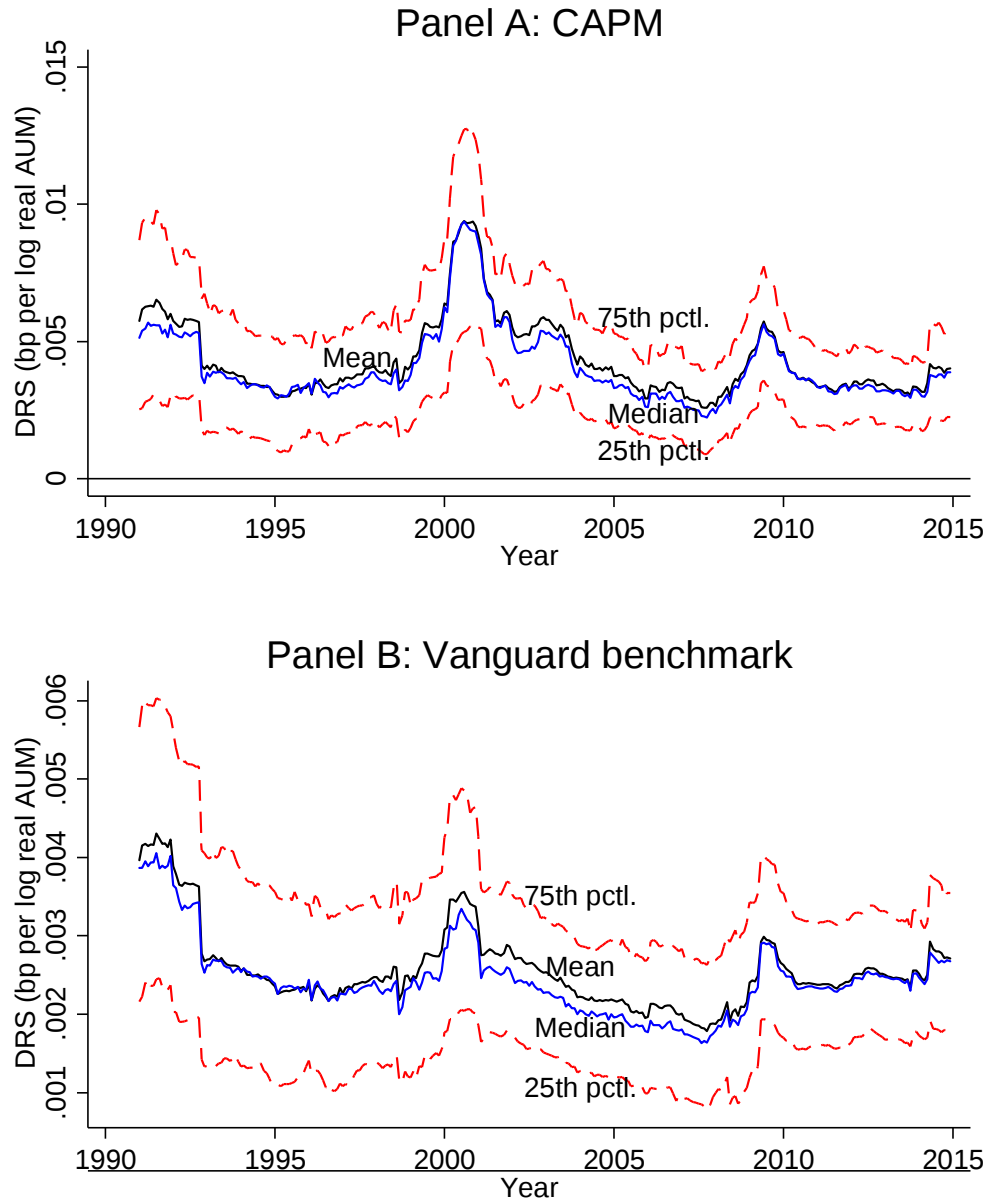


Figure 4: **Cross-Sectional Distribution of the Characteristic Component of DRS Estimates Over Time.** The figure displays how the cross-sectional distribution of $\hat{b}_{it}^{Char}$ —fund i 's DRS estimate in month t explained by contemporaneous fund characteristics—varies over time. Panel A shows the plot when fund-specific DRS is estimated using the CAPM, and Panel B shows the plot using Vanguard index funds as benchmark portfolios.

Online Appendix

to

“Capital Allocation and the Market for Mutual Funds: Inspecting the Mechanism”

OA Alternative Approaches to DRS Estimation

This appendix considers alternative approaches to estimating fund-specific b_i parameters in the time-series regression $\alpha_{it} = a_i - b_i \log(q_{it-1}) + \epsilon_{it}$. It is well known that the OLS estimators of the coefficients b_i are subject to a small-sample bias. The small sample-bias arises because of the flow-performance relation, which induces a positive correlation between the regression disturbance ϵ_{it} and the innovation in $\log(q_{it})$: if $\log(q_{it})$ obeys an AR(1) process,

$$\log(q_{it}) = \chi_i + \rho_i \log(q_{it-1}) + v_{it}, \tag{OA1}$$

Stambaugh (1999) shows that $\widehat{b}_i^{OLS}$ is upward biased, and proposes a first-order bias-corrected estimator of b_i . Amihud and Hurvich (2004, hereafter “AH”) improve upon this estimator by noting that adding a proxy v_{it}^c for the innovations in the autoregressive model can reduce the small-sample bias. The proxy v_{it}^c takes the form, $v_{it}^c = \log(q_{it}) - (\widehat{\chi}_i^c + \widehat{\rho}_i^c \log(q_{it-1}))$, where $\widehat{\chi}_i^c$ and $\widehat{\rho}_i^c$ are any estimators of χ_i and ρ_i constructed based on size data. Specifically, if we let $\widehat{b}_i^c$ be the coefficient of $\log(q_{it-1})$ in a time-series regression of α_{it} on $-\log(q_{it-1})$

and v_{it}^c , AH show that the bias of this estimator is given by:

$$E \left[\widehat{b}_i^c - b_i \right] = -\phi E \left[\widehat{\rho}_i^c - \rho_i \right], \quad (\text{OA2})$$

where $\phi = \sigma_{ev}/\sigma_v^2$. We adopt this estimation procedure, except we use different estimators of ρ_i than AH, guided by a horse race among various ρ_i estimators using simulations as further explained in the next subsection.

OA.1 Reduced-Bias Estimators of the AR(1) Parameter

AH's bias expression (OA2) says that the smaller the bias in the estimator $\widehat{\rho}_i^c$, the smaller the bias in the estimator $\widehat{b}_i^c$. AH suggest using a second-order bias-corrected estimator,

$$\widehat{\rho}_i^{c,AH} = \widehat{\rho}_i^{OLS} + (1 + 3\widehat{\rho}_i^{OLS})/T_i + 3(1 + 3\widehat{\rho}_i^{OLS})/T_i^2,$$

where $\widehat{\rho}_i^{OLS}$ is the coefficient of $\log(q_{it-1})$ in a time-series regression of $\log(q_{it})$ on $\log(q_{it-1})$ and T_i is the number of months. Barras, Gagliardini, and Scaillet (2022) use a variant of this estimator as given by $\widehat{\rho}_i^{c,BGS} = \min(\widehat{\rho}_i^{c,AH}, 0.999)$. Andrews (1993) proposes a median-unbiased estimator (MUE) of the AR parameter ρ_i ,

$$\widehat{\rho}_i^{c,MUE} = m_{ts}^{-1}(\widehat{\rho}_i^{OLS}),$$

where $m_{ts}(\rho_i)$ is the unique median of $\hat{\rho}_i^{OLS}$ when $\rho_i \in (-1, 1)$ is the true parameter.¹ Phillips and Sul (2003) extend this estimator to the fixed-effects setting under homogeneous ρ across funds by using $\hat{\rho}^{OLS\ FE}$ (the coefficient of $\log(q_{it-1})$ in a panel regression of $\log(q_{it})$ on $\log(q_{it-1})$ with fund fixed effects) to infer the true ρ . Similar to the MUE, the panel median-unbiased estimator (PMUE) is $\hat{\rho}^{c, PMUE} = m_{pnl}^{-1}(\hat{\rho}^{OLS\ FE})$, where $m_{pnl}(\rho)$ is the unique median of $\hat{\rho}^{OLS\ FE}$ when the true $\rho \in (-1, 1)$ is homogeneous across funds.²

We now assess the effectiveness of these various ρ_i estimators for recovering the b_i coefficients, which motivates us to propose an improved estimator in the next subsection: we simulate alphas and sizes from our calibrated model to generate 10,000 samples of panel data for 300 funds over 100 months. In each sample, we estimate $\hat{b}_i^{c, AH}, \hat{b}_i^{c, BGS}, \hat{b}_i^{c, MUE}, \hat{b}_i^{c, PMUE}$, which denote the bias-corrected b_i estimators corresponding to $\hat{\rho}_i^{c, AH}, \hat{\rho}_i^{c, BGS}, \hat{\rho}_i^{c, MUE}, \hat{\rho}_i^{c, PMUE}$, respectively; we then compute the positive bias in each $\hat{b}_i^c$ as the average difference between $\hat{b}_i^c$ and b_i across funds. Another way to assess the effectiveness of these b_i estimators is to run a cross-sectional regression of $\hat{b}_i^c$ on b_i in each sample: $\hat{b}_i^c = \psi_0 + \psi_1 b_i + \varepsilon_i$. The results of these exercises are in Table OA1.

The first takeaway is that the reduced-bias ρ_i estimators from the existing literature *do* reduce the biases in the corresponding b_i estimators, compared to the OLS estimator of b_i whose bias is equal to 3.91. Further, these b_i estimators *do* approximate the b_i coefficients in the sense that every $\hat{b}_i^c$ is positively related to the actual b_i . In particular, $\hat{b}_i^{c, PMUE}$ performs the best, which is why we use the PMUE to estimate fund-specific b_i parameters in our main

¹Key to this estimator is the fact that the distribution of $\hat{\rho}_i^{OLS}$ depends only on ρ_i when (OA1) is correct. In particular, it does not depend on χ_i or σ_v^2 in (OA1).

²As with the MUE, key to this estimator is the fact that the distribution of $\hat{\rho}^{OLS\ FE}$ depends only on ρ when the homogeneous dynamic panel model is correct.

analysis.

Unfortunately, none of the reduced-bias ρ_i estimators from the existing literature completely eliminates the bias in the corresponding b_i estimator if data were generated in a Berk and Green equilibrium, which motivates us to propose our own estimator of b_i .

OA.2 Novel Approach to Estimating the AR(1) Parameter

We now propose a novel estimator of b_i motivated by AH’s bias expression (OA2): set ρ_i to minimize the average size of the bias of the resulting b_i estimates across simulated samples. This process results in the value of $\rho_i = 0.999906$, which we denote by $\hat{\rho}_i^{c,BKK}$, with the corresponding bias-corrected b_i estimator, denoted by $\hat{b}_i^{c,BKK}$. We use the same simulation exercise as before: we use our calibrated model to generate 10,000 samples of panel data for 300 funds over 100 months, in each of which we estimate $\hat{b}_i^{c,BKK}$. We then assess the effectiveness of this bias-corrected b_i estimator by (i) computing its bias (the average difference between $\hat{b}_i^{c,BKK}$ and b_i across funds) and (ii) regressing it on b_i ($\hat{b}_i^{c,BKK} = \psi_0 + \psi_1 b_i + \varepsilon_i$); see the last column of Table OA1.

Fortunately, $\hat{b}_i^{c,BKK}$ *does* eliminate the bias completely, so we report our main empirical results (Paper’s Tables 2 through 5) when we use the model-implied ρ_i estimator— $\hat{\rho}_i^{c,BKK} = 0.999906$ —to estimate fund-specific DRS in Tables OA2 through OA5. This alternative approach leads to very similar results.

OA.3 Further Discussion

We end this appendix by explaining how reduced-bias ρ_i estimators from the existing literature cannot effectively eliminate the bias in the corresponding b_i estimators if data were generated in a Berk and Green equilibrium, and yet our empirical results are robust to the choice of the bias-corrected estimator $\widehat{b}_i^c$.

Consistent with AH’s bias expression (OA2), Table OA1 shows that the effectiveness of $\widehat{b}_i^c$ monotonically worsens in the underlying ρ_i estimator’s distance from $\widehat{\rho}_i^{c,BKK}$ —the “true” ρ_i in the simulated data.³ But every $\widehat{\rho}_i^c$ from the existing literature is actually very close to $\widehat{\rho}_i^{c,BKK}$ (see Panel C), which implies that the poor effectiveness of these b_i estimators must involve very high values of $\phi = \sigma_{ev}/\sigma_v^2$. Such high values of ϕ arise in our model for two reasons. The first is the fact that ϵ_{it} is perfectly correlated with v_{it} : a high fund return in period t is the sole reason for a larger fund size that period in a Berk and Green equilibrium. Therefore, σ_{ev} is very high in our model. The second is the fact that $\sigma_v^2 = (1 - \rho)^2 \frac{\sigma_\epsilon^2}{b_i^2}$ in our model:⁴ a high persistence in fund size coincides with tighter beliefs about skill in a Berk and Green equilibrium and thus a weaker flow-performance relationship—smaller changes in fund size in response to unexpected fund returns. Therefore, σ_v^2 is very small in our model.

In reality, a substantial fraction of flows remain unexplained by performance in the data (Berk and van Binsbergen (2016)).⁵ Consistent with this observation, the average of the ϕ estimates in our actual data is 0.115, which implies that the bias of 0.0058 in $\widehat{\rho}_i^{c,PMUE}$ results in a bias of 0.000667 in $\widehat{b}_i^{c,PMUE}$, which is about 14% of 0.0047 (the mean of b_i). So, our

³The PMUE of ρ_i in our actual data is 0.9988.

⁴Technically, ρ is a function of the fund’s age in our model, but the variation in ρ due to fund age is negligible.

⁵Similarly, Barber, Huang, and Odean (2016) find that less than 18% of flows are explained by performance.

empirical results are robust to the choice of the bias-corrected estimator because small biases in $\hat{\rho}_i^c$ *do* lead to small biases in the corresponding $\hat{b}_i^c$ in the actual data. Importantly, this observation suggests that the estimation error in fund-specific DRS are mostly of the classical type (i.e., mean zero and independent to the actual DRS parameter) in the actual data, so correcting the DRS-FSP relation estimates under the classical measurement error assumption should be able to recover the true relation. This reassures us that the approaches taken in our paper are delivering estimates of the DRS-FSP relation that avoid the errors-in-variable bias.

Table OA1: Horse Race Among the Various Estimators of DRS Using Simulations

This table reports the results of a horse race among the various estimators of the AR parameter ρ_i governing the size process $\log(q_{it}) = \chi_i + \rho_i \log(q_{it-1}) + v_{it}$ for the sake of recovering the b_i coefficients using our model. We simulate equilibrium alphas and sizes from the calibrated model to generate 10,000 samples of panel data for 300 funds over 100 months. In each sample, we estimate $\hat{b}_i^c$ as the coefficient of $\log(q_{it-1})$ in a time-series regression of α_{it} on $-\log(q_{it-1})$ and v_{it}^c , where v_{it}^c is a proxy for the size innovation. The proxy v_{it}^c takes the form, $v_{it}^c = \log(q_{it}) - (\hat{\chi}_i^c + \hat{\rho}_i^c \log(q_{it-1}))$, where $\hat{\chi}_i^c$ and $\hat{\rho}_i^c$ are any estimators of χ_i and ρ_i constructed based on size data. To obtain the proxy v_{it}^c , we consider four popular reduced-bias estimators of ρ_i from the existing literature, details of which are in Section OA.1. Importantly, we propose yet another bias-corrected estimator of b_i motivated by AH's bias expression (OA2): set ρ_i to minimize the magnitude of the mean bias of the b_i estimates across simulated samples. This process results in the value of $\rho_i = 0.999906$, which we denote by $\hat{\rho}_i^{c,*}$, with the corresponding bias-corrected b_i estimator, denoted by $\hat{b}_i^{c,*}$. In each sample, we then compute the non-negative biases in each of the $\hat{b}_i^c$'s as the average difference between $\hat{b}_i^c$ and b_i across funds. Panel A shows the median biases of the b_i estimates across simulated samples. Another way to assess the performance of these bias-corrected estimators of b_i is to regress each of the $\hat{b}_i^c$'s on b_i in each sample: $\hat{b}_i^c = \psi_0 + \psi_1 b_i + \varepsilon_i$. Panel B shows the medians of the ψ_0 and ψ_1 estimates across simulated samples. Panel C reports the medians of the ρ_i estimates (averaged across funds in each sample) across simulated samples.

Panel A: Bias in estimators of b_i corresponding to					
	reduced-bias estimators of ρ_i from the existing literature				our approach to estimating ρ_i
Bias in	$\hat{b}_i^{c,AH}$	$\hat{b}_i^{c,BGS}$	$\hat{b}_i^{c,MUE}$	$\hat{b}_i^{c,PMUE}$	$\hat{b}_i^{c,*}$
Median	0.934	1.717	1.802	0.438	-0.000
Panel B: Relation between b_i and its estimators corresponding to					
	reduced-bias estimators of ρ_i from the existing literature				our approach to estimating ρ_i
	$\hat{b}_i^{c,AH}$	$\hat{b}_i^{c,BGS}$	$\hat{b}_i^{c,MUE}$	$\hat{b}_i^{c,PMUE}$	$\hat{b}_i^{c,*}$
Median $\hat{\psi}_0$	0.033	0.043	0.044	-0.000	-0.000
Median $\hat{\psi}_1$	192.4	357.3	374.9	93.9	1.002
Panel C: Mean estimated ρ_i across funds corresponding to					
	reduced-bias estimators of ρ_i from the existing literature				our approach to estimating ρ_i
	$\hat{\rho}_i^{c,AH}$	$\hat{\rho}_i^{c,BGS}$	$\hat{\rho}_i^{c,MUE}$	$\hat{\rho}_i^{c,PMUE}$	$\hat{\rho}_i^{c,*}$
Mean	0.9874	0.9769	0.9758	0.9941	0.999906

Table OA2: Relation Between DRS and Flow Sensitivity to Performance (FSP)

This table is the same as Paper's Table 2 but uses the model-implied estimator of the size's persistence ρ_i ($\rho_i^{c,BKK} = 0.999906$) to estimate fund-specific b_i parameters.

Panel A: Using Estimates Based on the CAPM				
	$\widehat{FSP}_{it}$			
$\widehat{b}_{it}$	−0.943 (0.148)	−1.308 (0.182)	−0.263 (0.087)	−0.386 (0.117)
Fund FE & Month FE	Yes	Yes	Yes	Yes
Controls	No	Yes	No	Yes
Observations	182, 675	114, 623	182, 675	114, 623
Flow as % Change in	Fund Size	Fund Size	New Assets	New Assets
EIV Adj. Coefficient	−9.815	−17.99	−2.739	−5.315
SE of the EIV Adj. Coefficient	0.409	0.892	0.240	0.415

Panel B: Using Estimates Based on the Vanguard BM				
	$\widehat{FSP}_{it}$			
$\widehat{b}_{it}$	−1.551 (0.288)	−1.709 (0.327)	−0.384 (0.197)	−0.559 (0.181)
Fund FE & Month FE	Yes	Yes	Yes	Yes
Controls	No	Yes	No	Yes
Observations	221, 743	136, 365	221, 743	136, 365
Flow as % Change in	Fund Size	Fund Size	New Assets	New Assets
EIV Adj. Coefficient	−50.64	−135.3	−12.53	−44.28
SE of the EIV Adj. Coefficient	4.002	29.42	1.723	10.14

Table OA3: Relation Between Percentile Ranks of DRS and FSP

This table is the same as Paper's Table 3 but uses the model-implied estimator of the size's persistence ρ_i ($\rho_i^{c,BKK} = 0.999906$) to estimate fund-specific b_i parameters.

Panel A: Using Estimates Based on the CAPM				
	Pctl. rank based on $\widehat{FSP}_{it}$			
Pctl. rank based on $\widehat{b}_{it}$	−0.1084 (0.0109)	−0.1250 (0.0128)	−0.0657 (0.0094)	−0.0750 (0.0117)
Fund FE	Yes	Yes	Yes	Yes
Controls	No	Yes	No	Yes
Observations	182, 675	114, 623	182, 675	114, 623
Flow as % Change in	Fund Size	Fund Size	New Assets	New Assets
Panel B: Using Estimates Based on the Vanguard BM				
	Pctl. rank based on $\widehat{FSP}_{it}$			
Pctl. rank based on $\widehat{b}_{it}$	−0.0914 (0.0095)	−0.0943 (0.0117)	−0.0563 (0.0085)	−0.0597 (0.0108)
Fund FE	Yes	Yes	Yes	Yes
Controls	No	Yes	No	Yes
Observations	221, 743	136, 365	221, 743	136, 365
Flow as % Change in	Fund Size	Fund Size	New Assets	New Assets

Table OA4: Determinants of Fund-Level DRS

This table is the same as Paper's Table 4 but uses the model-implied estimator of the size's persistence ρ_i ($\rho_i^{c,BKK} = 0.999906$) to estimate fund-specific b_i parameters.

Panel A: Using Estimates Based on the CAPM		Dependent Variable: $\hat{b}_{it}$			
		Yes	Yes	Yes	Yes
Number of managers	-0.000251 (0.000070)				
Volatility ($SD(\hat{\alpha}_{it}^{CAPM})$)	0.1881 (0.0224)				0.1789 (0.0253)
Expense ratio			0.1660 (0.0458)		0.1966 (0.0681)
Marketing expenses				-0.0250 (0.0359)	-0.1306 (0.0432)
Intl exposure				0.000054 (0.000350)	
Turnover				0.000452 (0.000288)	
Log fund size					0.00004 (0.000109)
Style & Risk Controls		Yes	Yes	Yes	Yes
Observations		248, 076	252, 361	252, 427	247, 256
				252, 433	252, 413
					154, 998

Panel B: Using Estimates Based on the Vanguard BM

	Dependent Variable: $\hat{b}_{it}$					
Number of managers	-0.000076 (0.000036)					
Volatility ($SD(\hat{\alpha}_{it}^{VG})$)	0.0473 (0.0182)					0.0644 (0.0211)
Expense ratio		0.0538 (0.0307)				0.0801 (0.0454)
Marketing expenses			-0.0300 (0.0278)			-0.0723 (0.0340)
Intl exposure				0.000001 (0.000235)		
Turnover					-0.000040 (0.000185)	
Log fund size						0.000043 (0.000069)
Style & Risk Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	294, 922	300, 883	300, 957	300, 963	294, 666	300, 941
						180, 913

Table OA5: Relation Between the Characteristic Component of DRS and FSP

This table is the same as Paper's Table 5 but uses the model-implied estimator of the size's persistence ρ_i ($\rho_i^{S,BKK} = 0.999906$) to estimate fund-specific b_i parameters.

Panel A: Using Estimates Based on the CAPM									
	$\widehat{FSP}_{it}$								
$\widehat{b}_{it}^{Char}$	-10.422 (1.970)	-9.725 (2.032)	-9.979 (1.936)	-11.952 (2.030)	-6.323 (1.262)	-6.413 (1.321)	-5.875 (1.246)	-6.466 (1.275)	
Log fund size	-0.004 (0.008)	-0.005 (0.008)	-0.006 (0.008)	-0.006 (0.008)	-0.020 (0.006)	-0.021 (0.006)	-0.022 (0.006)	-0.022 (0.006)	
Marketing expenses	-0.299 (0.979)	-0.542 (1.017)	-0.055 (0.984)	-1.432 (0.999)	-0.360 (0.598)	-0.467 (0.612)	-0.184 (0.599)	-0.955 (0.611)	
Star family affiliation	-0.006 (0.005)	-0.007 (0.005)	-0.007 (0.005)	-0.007 (0.005)	-0.012 (0.003)	-0.012 (0.003)	-0.012 (0.003)	-0.012 (0.003)	
Log family size	0.015 (0.011)	0.016 (0.011)	0.016 (0.011)	0.015 (0.011)	0.006 (0.007)	0.007 (0.007)	0.007 (0.007)	0.006 (0.007)	
Diverse offerings	0.001 (0.015)	-0.001 (0.014)	-0.000 (0.014)	-0.000 (0.014)	-0.002 (0.009)	-0.003 (0.009)	-0.003 (0.009)	-0.003 (0.009)	
Log fund age	-0.015 (0.030)	-0.015 (0.030)	-0.018 (0.030)	-0.018 (0.030)	-0.013 (0.017)	-0.014 (0.017)	-0.016 (0.017)	-0.016 (0.017)	
Volatility ($SD(\widehat{\alpha}_t^{CAPM})$)	0.566 (0.408)	2.277 (0.564)	0.530 (0.405)	2.644 (0.560)	-0.139 (0.245)	1.069 (0.335)	-0.097 (0.239)	1.037 (0.321)	
Fund FE & Month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	
Observations	112,071	114,623	114,623	114,623	112,071	114,623	114,623	114,623	
Flow as % Change in	Fund Size	Fund Size	Fund Size	Fund Size	New Assets	New Assets	New Assets	New Assets	
$\widehat{b}_{it}^{Char}$ based on Table OA4 A, Col.	1	2	3	8	1	2	3	8	
$SD(\widehat{b}_{it}^{Char})$	0.0023	0.0029	0.0023	0.0030	0.0023	0.0029	0.0023	0.0030	

Panel B: Using Estimates Based on the Vanguard BM

	$\widehat{FSP}_{it}$							
$\widehat{b}_{it}^{Char}$	-8.382 (3.537)	-7.484 (3.515)	-7.671 (3.452)	-7.476 (3.834)	-7.324 (2.157)	-7.268 (2.130)	-6.854 (2.087)	-7.249 (2.220)
Log fund size	-0.031 (0.009)	-0.033 (0.009)	-0.033 (0.009)	-0.033 (0.009)	-0.046 (0.007)	-0.048 (0.007)	-0.048 (0.007)	-0.048 (0.007)
Marketing expenses	-1.500 (1.327)	-1.611 (1.348)	-1.490 (1.339)	-1.969 (1.346)	-1.096 (0.785)	-1.106 (0.782)	-0.994 (0.774)	-1.453 (0.786)
Star family affiliation	-0.016 (0.007)	-0.017 (0.007)	-0.017 (0.007)	-0.017 (0.007)	-0.016 (0.004)	-0.017 (0.004)	-0.017 (0.004)	-0.017 (0.004)
Log family size	0.002 (0.013)	0.007 (0.013)	0.007 (0.013)	0.007 (0.013)	0.007 (0.008)	0.010 (0.008)	0.010 (0.008)	0.010 (0.008)
Diverse offerings	-0.004 (0.019)	-0.008 (0.019)	-0.008 (0.019)	-0.008 (0.019)	-0.001 (0.012)	-0.004 (0.012)	-0.003 (0.012)	-0.003 (0.012)
Log fund age	-0.020 (0.032)	-0.017 (0.032)	-0.018 (0.032)	-0.018 (0.032)	-0.010 (0.018)	-0.011 (0.017)	-0.011 (0.017)	-0.011 (0.017)
Volatility ($SD(\widehat{\alpha}_{it}^{VG}))$	-0.587 (0.720)	-0.174 (0.758)	-0.514 (0.708)	-0.058 (0.789)	-0.059 (0.487)	0.311 (0.504)	-0.025 (0.471)	0.423 (0.515)
Fund FE & Month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	132, 802	136, 365	136, 365	136, 365	132, 802	136, 365	136, 365	136, 365
Flow as % Change in	Fund Size	Fund Size	Fund Size	Fund Size	New Assets	New Assets	New Assets	New Assets
$\widehat{b}_{it}^{Char}$ based on Table OA4 B, Col.	1	2	3	8	1	2	3	8
$SD(\widehat{b}_{it}^{Char})$	0.0016	0.0016	0.0016	0.0016	0.0016	0.0016	0.0016	0.0016