

Scale and Skill in Active Management

Ľuboš Pástor

Robert F. Stambaugh

Lucian A. Taylor *

August 30, 2013

Abstract

We empirically analyze the nature of returns to scale in active mutual fund management. We find strong evidence of decreasing returns at the industry level: As the size of the active mutual fund industry increases, a fund's ability to outperform passive benchmarks declines. In contrast, estimates that avoid econometric biases do not detect decreasing returns at the fund level. We also find that funds born more recently exhibit more skill. This upward trend in skill coincides with industry growth, which precludes the skill improvement from boosting fund performance. Finally, we find that performance deteriorates over a typical fund's lifetime. This result can also be explained by industry growth and industry-level decreasing returns to scale.

*Pástor is at the University of Chicago Booth School of Business. Stambaugh and Taylor are at the Wharton School of the University of Pennsylvania. We are grateful for comments from the workshop participants at the University of Melbourne, University of Queensland, and University of Western Australia. We are also grateful to Yeguang Chi for superb research assistance. This research was funded in part by the Initiative on Global Markets at the University of Chicago Booth School of Business.

1. Introduction

The performance of active mutual funds has been of long-standing interest to financial economists.¹ The extent to which an active fund can outperform its passive benchmark depends not only on the fund’s raw skill in identifying investment opportunities but also on various constraints faced by the fund. One constraint discussed prominently in recent literature is decreasing returns to scale. If scale impacts performance, skill and scale interact: for example, a more skilled large fund can underperform a less skilled small fund. Therefore, to learn about skill, we must understand the effects of scale.

What is the nature of returns to scale in active management? The literature has advanced two hypotheses. The first one is *fund-level* decreasing returns to scale: as the size of an active fund increases, the fund’s ability to outperform its benchmark declines (e.g., Perold and Solomon, 1991, and Berk and Green, 2004). The second hypothesis is *industry-level* decreasing returns to scale: as the size of the active mutual fund industry increases, the ability of any given fund to outperform declines (Pástor and Stambaugh, 2012). Both hypotheses have been motivated by liquidity constraints. At the fund level, a larger fund’s trades have a larger impact on asset prices, eroding the fund’s performance. At the industry level, as more money chases opportunities to outperform, prices move, making such opportunities more elusive. Consistent with such liquidity constraints, there is mounting evidence that trading by mutual funds is capable of exerting meaningful price pressure in equity markets.²

Whether returns to scale operate at the fund level or the industry level, or both or neither, is not clear a priori. At one extreme, if all funds were to follow exactly the same strategy, their performance would more likely depend on their combined size than on their individual sizes. At the other extreme, if they were to follow completely unrelated strategies, the opposite would be true. The relative merits of the two hypotheses must be evaluated empirically. The fund-level hypothesis has been tested in a number of recent studies, with mixed results.³ We provide the first evidence regarding the industry-level hypothesis, to our knowledge. We also reexamine the fund-level hypothesis by using cleaner data and econometric techniques that avoid inherent biases.

¹See, for example, Jensen (1968), Ferson and Schadt (1996), Carhart (1997), Daniel et al. (1997), Wermers (2000), Pástor and Stambaugh (2002), Cohen, Coval, and Pástor (2005), Kacperczyk, Sialm, and Zheng (2005, 2008), Kosowski et al. (2006), Barras, Scaillet, and Wermers (2010), Fama and French (2010), etc.

²For example, Edelen and Warner (2001) find that aggregate flow into equity mutual funds has an aggregate impact on market returns. Wermers (2003), Coval and Stafford (2007), Khan, Kogan, and Serafeim (2012), and Lou (2012) also find significant price impact associated with mutual fund trading. Edelen, Evans, and Kadlec (2007) report that trading costs are a major source of diseconomies of scale for mutual funds.

³See, for example, Chen et al. (2004), Pollet and Wilson (2008), Yan (2008), Ferreira et al. (2013a,b), and Reuter and Zitzewitz (2013). We discuss this evidence in more detail later in the introduction.

One of the challenges in estimating the effect of fund size on performance is the endogeneity of fund size. If size were randomly assigned to funds, one could simply run a panel regression of funds' benchmark-adjusted returns on lagged fund size, and the OLS slope estimate would correctly measure the effect of size on performance. Alas, size is not randomly paired with funds; for example, larger funds are likely to be run by managers with higher skill (e.g., Berk and Green, 2004). Skill is likely to be correlated with both fund size and performance, yet we cannot control for skill as it is unobservable. As a result, the simple OLS estimate of the size-performance relation suffers from an omitted-variable bias.

The omitted-variable bias can be eliminated by including fund fixed effects in the regression model. These fixed effects absorb the cross-sectional variation in performance that is due to differences in skill across funds. Unfortunately, while adding fund fixed effects removes one bias, it introduces another. This second bias results from the positive contemporaneous correlation between changes in fund size and unexpected fund returns. In general, a nonzero correlation between a regressor's innovations and the regression disturbances introduces a finite-sample bias in OLS estimates (Stambaugh, 1999; Hjalmarsson, 2010).

To address the second bias, we develop a *recursive demeaning* procedure that closely builds on the methods of Moon and Phillips (2000) and Hjalmarsson (2010). This procedure runs a panel regression of forward-demeaned returns on forward-demeaned fund size, while instrumenting for the latter quantity by its backward-demeaned counterpart. The resulting estimator eliminates the bias, as proved by Hjalmarsson and confirmed in our simulation analysis. Our simulations also highlight the bias in both OLS estimators, with and without fund fixed effects. In addition to being biased, the OLS estimators heavily overreject the null hypothesis of no returns to scale even when this hypothesis is true.

Our empirical analysis relies on a cross-validated dataset of actively managed U.S. equity mutual funds. We reconcile the key data items in the CRSP and Morningstar databases, building on the work of Berk and Binsbergen (2012). Our dataset covers 3,126 funds from 1979 through 2011, a period during which the mutual fund industry grew dramatically.

We begin our analysis by using panel data to estimate the slope coefficient of fund performance regressed on lagged fund size. OLS regressions both with and without fund fixed effects deliver negative estimates that, while statistically significant, are small in economic magnitude. These OLS procedures are biased in opposite directions, depending on whether or not fixed effects are included. As a result, the estimates from these regressions deliver a mixed message about the presence of economically meaningful fund-level decreasing returns. To avoid the biases in OLS, we apply the recursive demeaning procedure. The estimates

of fund-level returns to scale are again negative but small, and they become statistically insignificant as well. This result is robust to the inclusion of numerous controls. Overall, we do not find reliable evidence of decreasing returns at the fund level.

In contrast, we consistently find evidence of decreasing returns to scale at the industry level. Using the same panel regressions, we find a negative relation between industry size and fund performance. When we include both fund size and industry size in the regression, fund size is insignificant in the bias-free specification, whereas industry size is negative and significant. In addition, we find that the negative relation between industry size and fund performance is stronger for funds with higher turnover, higher volatility, as well as small-cap funds. These results seem sensible since funds that are more aggressive in their trading, as well as funds that trade less liquid assets, are likely to face larger total price impact costs when competing in a more crowded industry.

The evidence of industry-level decreasing returns to scale has important implications for our assessment of fund manager skill. We measure skill by the estimated fund fixed effect from our panel regression. This fixed effect is essentially equal to the average benchmark-adjusted gross fund return (i.e., the usual gross alpha) that is further adjusted for any potential fund-level and industry-level returns to scale. We find that the average fund's skill has increased substantially over time, from -5 basis points (bp) per month in 1979 to +13 bp per month in 2011. The improvement in skill is steeper among the better-skilled funds: e.g., the 90th percentile of the cross-sectional distribution of skill grows from 51 bp to 88 bp per month. In short, funds have become more skilled over time.

This improvement in skill has failed to boost fund performance, though, judging by the non-trending average benchmark-adjusted gross fund return. How can we reconcile the upward trend in skill with no trend in performance? Our explanation combines industry-level decreasing returns to scale with the observed steady growth in industry size. We argue that the growing industry size makes it harder for fund managers to outperform despite their improving skill. The active management industry today is bigger and more competitive than it was 30 years ago, so it takes more skill just to keep up with the rest of the pack.

The upward trend in average fund skill is not driven by rising skill within funds, because our measure of a fund's skill is constant over the fund's lifetime. Instead, our evidence suggests that the new funds entering the industry are more skilled, on average, than the existing funds. Consistent with this interpretation, we find that younger funds outperform older funds in a typical month. We sort funds into portfolios based on their age and find that funds aged up to three years outperform those aged more than 10 years by a statistically

significant 0.9% per year, based on gross benchmark-adjusted returns. Funds aged between three and six years also outperform the oldest funds. The young-minus-old portfolio differences are smaller when measured in net returns, suggesting that the younger funds capture a portion of their higher skill by charging higher fees.

The negative age-performance relation holds not only across funds but also within funds. We find that performance deteriorates over a typical fund's lifetime. This result does not seem to be due to the incubation bias (Evans, 2010) because the performance decline continues well beyond the first few years of the fund's existence. Instead, this erosion in fund performance seems to be driven by industry growth during the fund's lifetime. As the fund ages, the industry keeps growing, and the sustained entry of skilled competitors hurts the fund's performance. Consistent with this argument, we find that the negative relation between a fund's age and its performance disappears after we control for industry size.

Taken together, our results are consistent with the following narrative. New funds entering the industry tend to be more skilled than the incumbent funds, perhaps due to better education or greater command of new technology. As a result of their superior skill, the new funds tend to outperform their benchmarks as well as older funds. As these funds grow older, though, their performance suffers as a result of the continued growth in industry size, which is associated with steady arrival of skilled competition.

Our measure of a fund's skill is the gross alpha earned on the first dollar invested in the fund, with no other funds present in the industry. We seek to measure the fund's ability to identify profitable investment opportunities before they are eroded by decreasing returns to scale. In contrast, traditional measures of skill such as alpha or the Sharpe ratio do not separate the effects of scale. Fund size does play a role in the measure of Berk and Binsbergen (2012) but that measure quantifies a different dimension of skill—dollar value added by the fund—whereas we attempt to measure the fund's expected benchmark-adjusted return while taking into account the adverse effects of scale.

While our focus on industry-level returns to scale is novel, others have investigated returns to scale at the fund level. Chen et al. (2004) find a negative relation between fund return and lagged fund size, consistent with fund-level decreasing returns. The negative relation is strongest among small-cap funds, leading the authors to conclude that the adverse scale effects are related to liquidity. Yan (2008) reaches the same conclusion based on more direct measures of liquidity—bid-ask spread and market impact. Yan finds a stronger negative size-performance relation for funds that hold less liquid portfolios, as well as for growth funds and high-turnover funds, which tend to demand immediacy. Further support for liquidity-

related diminishing returns comes from Bris et al. (2007), who analyze mutual funds that have closed to new investment, and from Pollet and Wilson (2008), who examine the response of mutual funds to asset growth.⁴

The prior evidence of fund-level decreasing returns to scale is not pervasive across funds. Ferreira et al. (2013a) analyze the performance of active equity mutual funds in 27 countries. They find diseconomies of scale for U.S. funds but not for non-U.S. funds; in fact, the latter funds seem to exhibit increasing returns to scale. Even in the U.S., the negative size-performance relation seems to obtain only for the subset of funds most affected by illiquidity. For example, Chen et al. (2004) find that this relation is significantly negative only for small-cap funds. Yan (2008) finds the negative relation only among funds with the least liquid holdings, while Bris et al. (2007) find it only among funds with large inflows.

The finding of fund-level diminishing returns is also not universal among prior studies. Early evidence from Grinblatt and Titman (1989) is mixed, depending on how one measures fund returns. More recently, Reuter and Zitzewitz (2013) recognize the endogeneity of fund size and the resulting difficulty in identifying the causal impact of size on performance. To generate exogenous variation in size, these authors exploit a discontinuity in fund flows across Morningstar star ratings. They note that small differences in fund performance can cause discrete changes in Morningstar ratings, which then produce sharp changes in fund size. After applying their regression discontinuity approach to U.S. funds, they find no evidence of fund-level diseconomies of scale. We address the endogeneity of fund size in a different way, namely, by including fund fixed effects to account for heterogeneity in skill. We also show how to obtain unbiased estimates of the size-performance relation in such a setting.

The paper is organized as follows. Section 2. discusses the econometric problems associated with estimating the size-performance relation at the fund level. It also presents a fix—a recursive demeaning procedure—and evaluates its effectiveness in simulations. Section 3. describes our mutual fund dataset. Section 4. presents our empirical results. We first analyze the nature of returns to scale (fund-level vs industry-level), followed by the determinants of the size-performance relation. We then examine the evolution of fund skill as well as the relation between fund performance and fund age. Section 5. concludes.

⁴There is also some evidence of decreasing returns to scale outside the mutual fund industry. For example, Kaplan and Schoar (2005) report such evidence in venture capital and leveraged buyout investing, and Fung et al. (2008) find decreasing returns in the hedge fund industry.

2. Methodology

Estimating the effect of fund size on performance is a challenge because size is determined endogenously. The next subsection explains why the simple regression approach taken in a number of studies delivers biased estimates, and why adding fund fixed effects removes this bias while introducing another one. Section 2.2. presents a recursive-demeaning (RD) estimator that eliminates both biases. Section 2.3. uses simulations to illustrate the bias in OLS estimators, as well as the RD estimator’s ability to avoid the bias.

2.1. Biases in OLS estimators

Let R_{it} denote the benchmark-adjusted return of fund i in period t , and let q_{it-1} denote the fund’s size at the end of period $t - 1$. A simple approach to investigating returns to scale is to use panel data across funds and periods to estimate the regression model

$$R_{it} = a + \beta q_{it-1} + \varepsilon_{it} . \tag{1}$$

If size were random across funds, independent of manager skill, the OLS estimate of β would successfully identify the effect of size on performance. Specifically, a negative estimate of β would indicate decreasing returns to scale. However, independence of fund size and skill is unlikely. Larger funds are likely to be paired with higher-skill managers for two reasons. First, higher-skill managers tend to perform better, *ceteris paribus*, and their superior performance increases fund size, in part by attracting flow.⁵ Second, larger funds tend to collect larger total fees, so they can in principle afford to hire better managers. The positive relation between size and skill in the cross section works against any potential diseconomies of scale, and thus the pooled regression (1) is ill-suited for detecting a negative relation between size and performance even if one truly exists. In econometric terms, the regression (1) suffers from an omitted-variable bias: skill has a positive effect on performance and is positively correlated with fund size, so omitting skill from the regression biases the estimate of β upwards.⁶ Given the potential bias, we prefer not to base inference on the simple regression (1), while recognizing that previous studies have nevertheless done so. For example, Ferreira et al. (2013a,b) estimate a pooled OLS panel regression of fund performance on size, as in equation (1), while Chen et al. (2004) and Yan (2008) estimate

⁵The performance-flow relation is both an empirical observation (e.g., Chevalier and Ellison, 1997, and Sirri and Tufano, 1998) and a theoretical prediction (e.g., Berk and Green, 2004).

⁶This bias has been noted in the literature (e.g., Chen et al. (2004) and Reuter and Zitzewitz (2013)). Applying the omitted-variable bias formula (e.g., Angrist and Pischke, 2009), the bias is equal to the effect of skill on performance times the slope of skill on fund size, both of which are likely to be positive.

the same pooled model using a Fama-MacBeth regression approach. Those studies include control variables, but such controls necessarily omit skill, which is unobservable.

Fortunately, the omitted-variable bias can be eliminated by including a fund fixed effect, denoted by a_i , so that equation (1) is replaced by

$$R_{it} = a_i + \beta q_{it-1} + \varepsilon_{it} . \quad (2)$$

The fund fixed effects soak up any variation in performance due to cross-sectional differences in fund skill, as long as that skill is constant over time. Identification in the fixed-effect (FE) model comes from variation over time within a fund, not from variation across funds.

The simple regression model in equation (2) can be motivated, for example, by the model of Berk and Green (2004). That model assumes fund-level diseconomies of scale, which imply $\beta < 0$ in equation (2). Note that this implication does not contradict Berk and Green’s result that fund size should not predict performance from the real-time perspective of investors. While Berk and Green’s investors perceive no relation between fund size and the fund’s expected return in real time, the true relation—one examined by an econometrician analyzing historical data—is negative. This difference between the objective and subjective size-performance relations stems from the unobservability of fund skill. As investors update their beliefs about skill, their perception of skill fluctuates even though true skill is time-invariant. Changes in perceived skill lead to changes in fund size, which negatively impact the true expected fund return due to diseconomies of scale. For example, when a fund’s perceived skill exceeds its true skill, the fund exceeds its optimal size and its expected future return is lower. Conversely, when perceived skill is below true skill, the fund is smaller and its expected return is higher.⁷

Unfortunately, eliminating the omitted-variable bias associated with equation (1) by including fixed effects as in equation (2) introduces a second bias if the latter specification is estimated with OLS. The omitted-variable bias exists even in large samples, whereas this second bias arises in finite samples through the channel discussed by Stambaugh (1999). To understand the latter bias in the OLS fixed-effects estimator $\hat{\beta}_{FE}$, consider first the OLS estimator $\hat{\beta}_i$, the estimator of β in equation (2) using the data for just a single fund i . As shown by Stambaugh (1999), $\hat{\beta}_i$ in that simple predictive regression is downward biased when

⁷This point has not been fully appreciated in the literature. For example, Elton, Gruber, and Blake (2012) write that “Berk and Green (2004) argue that there is no predictability” (p. 38), and that fund size could predict returns only if Berk and Green’s investors were slow to move capital in response to returns (p. 33). Reuter and Zitzewitz (2013) argue that the Berk-Green model implies that “fund size will be uncorrelated with future returns, thereby frustrating standard approaches to estimate diseconomies of scale” (p. 2). Both studies interpret no predictability as an empirical implication, rather than a statement about investor expectations.

the regression disturbance ε_{it} in equation (2) is positively correlated with the innovation in q_{it} . This positive correlation arises in our setting for two reasons. The first is a mechanical link between ε_{it} and q_{it} : a high fund return in period t corresponds to an increase in the fund’s asset values and thus to a higher fund size at the end of that period. The second is the performance-flow relation—a high return during period t attracts new money into the fund, also contributing to a higher fund size at the end of that period.

To see intuitively why $\hat{\beta}_i$ is negatively biased, suppose $a_i = \beta = 0$ and we have a two-period sample ($t = 1, 2$) with no net flow. Given the positive correlation between ε_{it} and q_{it} , we have $q_{i1} < q_{i0}$ if $\varepsilon_{i1} < 0$, and $q_{i1} > q_{i0}$ if $\varepsilon_{i1} > 0$. Since in either scenario ε_{i2} is zero on average, the higher of the two $q_{i,t-1}$ ’s will tend to precede the lower of the two ε_{it} ’s (which are equal to the R_{it} ’s since $a_i = \beta = 0$). In other words, a fund that outperforms by chance (i.e., $\varepsilon_{i1} > 0$) will grow in size (i.e., $q_{i1} > q_{i0}$), but its future performance is expected to be worse (because $E(\varepsilon_{i2}) = 0$). Conversely, a fund that underperforms by chance will shrink in size, but its future performance is expected to be better. This effect produces a spurious negative relation between changes in fund size and future fund performance. This is a small-sample problem because the tendency for a sample’s highest $q_{i,t-1}$ ’s to precede its lowest R_{it} ’s even when $\beta = 0$ is strongest in small samples. As sample length grows, a given level of $q_{i,t-1}$ eventually gets paired with as many high values as low values of R_{it} .

Now consider the OLS estimator $\hat{\beta}_{FE}$. It is straightforward to show that $\hat{\beta}_{FE} = \sum_{i=1}^N w_i \hat{\beta}_i$, where $\sum_{i=1}^N w_i = 1$ and the w_i ’s are positive.⁸ Thus, the negative bias in $\hat{\beta}_{FE}$ is essentially just the weighted average of the negative biases in each of the $\hat{\beta}_i$ ’s. As a result of this negative bias, the OLS fixed-effects estimator can “detect” decreasing returns to scale even when there are none.

2.2. Recursive demeaning

Fortunately, there is an estimator that allows fund fixed effects while avoiding the finite-sample bias. To understand this estimator, it is useful to begin with an alternative explanation of the source of the bias in the OLS FE estimator. This explanation as well as our implementation of the estimator that avoids the bias largely follow Hjalmarsson (2010).

The OLS estimator of β in equation (2) is equivalent to the OLS estimator for the demeaned model $\tilde{R}_{it} = \beta \tilde{q}_{it-1} + \tilde{\varepsilon}_{it}$, where \tilde{R}_{it} , \tilde{q}_{it-1} , and $\tilde{\varepsilon}_{it}$ are equal to R_{it} , q_{it-1} , and ε_{it} minus

⁸See, for example, Juhl and Lugovskyy (2010). Specifically, $w_i = T_i \hat{\sigma}_{q_i}^2 / \sum_{j=1}^N T_j \hat{\sigma}_{q_j}^2$, where T_i is the number of observations for fund i and $\hat{\sigma}_{q_i}^2$ is the sample variance of q_{it} .

their full-sample time-series means at the fund level. That is, $\hat{\beta}_{FE} = (\sum_{t,i} \tilde{q}_{i,t-1}^2)^{-1} (\sum_{t,i} \tilde{q}_{i,t-1} \tilde{R}_{it})$, and thus

$$\hat{\beta}_{FE} - \beta = \left(\sum_{t,i} \tilde{q}_{i,t-1}^2 \right)^{-1} \left(\sum_{t,i} \tilde{q}_{i,t-1} \tilde{\varepsilon}_{it} \right). \quad (3)$$

The bias in $\hat{\beta}_{FE}$ arises because, even though $q_{i,t-1}$ and ε_{it} have zero correlation, $\tilde{q}_{i,t-1}$ and $\tilde{\varepsilon}_{it}$ do not, as a result of which the second factor in equation (3) has nonzero expectation. Because a fund's full-sample time-series mean is subtracted when computing the demeaned series, the value of $\tilde{q}_{i,t-1}$ depends on observations after period $t-1$. In particular, a high value of q_{it} increases the time-series mean, which decreases $\tilde{q}_{i,t-1}$. Therefore, $\tilde{q}_{i,t-1}$ is negatively correlated with the innovation in q_{it} , which in turn is positively correlated with ε_{it} . Recall that the latter correlation is the source of the bias. The effect of that correlation in the context of equation (3) is a negative correlation between $\tilde{q}_{i,t-1}$ and $\tilde{\varepsilon}_{it}$, which produces a negative expectation for the second factor, resulting in the negative bias in $\hat{\beta}_{FE}$.

If $q_{i,t-1}$ were instead backward-demeaned by a mean computed using only fund i 's observations prior to period $t-1$, rather than the fund's full-sample mean, then that demeaned value of $q_{i,t-1}$ would be uncorrelated with ε_{it} . Such backward demeaning, applied recursively through time, forms the basis for the instrumental variable estimator we employ to eliminate the bias. While demeaning in a recursive fashion adds noise compared to demeaning with a fund's less noisy full-sample mean, applying such an approach in a panel setting nevertheless yields reliable inferences by aggregating information across a large cross section of funds.

In applying the recursive demeaning (RD) estimator, we expand the FE model to include a vector of regressors, x_{it-1} , that potentially include lagged size, q_{it-1} :

$$R_{it} = a_i + \beta' x_{it-1} + \varepsilon_{it} . \quad (4)$$

Following the notation of Moon and Phillips (2000), we define the recursively backward-demeaned regressors, \underline{x}_{it-1} , for $t = 2, \dots, T_i$, as

$$\underline{x}_{it-1} = x_{it-1} - \frac{1}{t-1} \sum_{s=1}^{t-1} x_{is-1} . \quad (5)$$

Similarly, recursively forward-demeaned variables are

$$\bar{x}_{it-1} = x_{it-1} - \frac{1}{T_i - t + 1} \sum_{s=t}^{T_i} x_{is-1} \quad (6)$$

$$\bar{R}_{it} = R_{it} - \frac{1}{T_i - t + 1} \sum_{s=t}^{T_i} R_{is} . \quad (7)$$

Substituting these definitions into equation (4) makes the fixed effects a_i drop out:

$$\bar{R}_{it} = \beta' \bar{x}_{it-1} + \bar{\varepsilon}_{it} , \quad (8)$$

where $\bar{\varepsilon}_{it}$ is defined in a manner analogous to \bar{R}_{it} .

We estimate regression (8) by using the instrumental variables (IV) approach. When x_{it-1} includes fund size (q_{it-1}), we instrument for \bar{q}_{it-1} by using \underline{q}_{it-1} . We treat the elements of x_{it-1} other than fund size, such as industry size or fund turnover, as exogenous regressors, since their innovations are not plausibly correlated with the fund’s benchmark-adjusted return. When x only includes fund size, the IV estimator of regression (8) is simply

$$\hat{\beta}_{RD} = \left(\sum_{i=1}^n \sum_{t=2}^{T_i} \bar{q}_{it-1} \underline{q}'_{it-1} \right)^{-1} \left(\sum_{i=1}^n \sum_{t=2}^{T_i} \bar{R}_{it} \underline{q}'_{it-1} \right). \quad (9)$$

This estimator is the same as Hjalmarsson’s (2010), except that we backward-demean our instrument. (This backward-demeaning is necessary in our setting because, unlike the regressor in Hjalmarsson’s setting, our regressor, fund size, does not have zero mean.) Since the estimator in equation (9) is an IV estimator, we can implement it via two-stage least squares. We first regress \bar{q}_{it-1} on \underline{q}_{it-1} , and then we regress \bar{R}_{it} on the fitted values from the first-stage regression. Neither regression includes an intercept.

To be a valid instrument for \bar{q}_{it-1} , \underline{q}_{it-1} must satisfy the relevance and exclusion conditions (e.g., Roberts and Whited, 2012). The relevance condition requires that \bar{q}_{it-1} and \underline{q}_{it-1} be significantly related in the first-stage regression. Since \bar{q}_{it-1} and \underline{q}_{it-1} are both derived from q_{it-1} (see equations (5) and (6)), they indeed tend to be closely related.⁹ The exclusion condition requires that $E[\bar{\varepsilon}_{it} | \underline{q}_{it-1}] = 0$, meaning the instrument is unrelated to the innovation in the dependent variable. This condition is likely to hold as well, since the backward-looking information in \underline{q}_{it-1} is unlikely to be helpful in predicting the forward-looking return information in $\bar{\varepsilon}_{it}$. In contrast, $E[\tilde{\varepsilon}_{it} | \tilde{q}_{it-1}] \neq 0$ in the OLS FE estimator, as discussed above. This distinction is the reason why $\hat{\beta}_{RD}$ eliminates the bias in $\hat{\beta}_{FE}$.

⁹For most funds in our data, \bar{q}_{it-1} and \underline{q}_{it-1} are positively related in the first-stage regression. Some funds, however, exhibit a negative relation when we fit this regression through the origin, due to trends in their size. We exclude a small number of these trending funds—less than 2% of observations in Table 3, for example—to prevent them from weakening the first-stage relation. Specifically, we run two regressions for each fund: we regress \bar{q}_{it-1} on \underline{q}_{it-1} , both with and without an intercept. We exclude funds that have both a negative slope in the first regression and an intercept in the second regression whose absolute value is above a threshold. We choose this threshold in each model to exclude as few funds as possible while delivering a positive first-stage relation as well as a first-stage Angrist-Pischke (2009) F -statistic above 10. Stock, Wright, and Yogo (2002) show that the bias from weak instruments is small when the F -statistic is above 10. We apply this procedure to all variables that depend on q_{it-1} when we implement RD.

2.3. Simulation exercise

We use simulations to illustrate the bias in the OLS estimators, with and without fixed effects, as well as the unbiased nature of the RD estimator. After simulating data in which we know the true relation between returns and fund size, we check whether the estimators are able to recover the true relation. To gauge the estimators' size and power, we simulate data both with and without decreasing returns to scale.

The first step is to simulate panel data on funds' returns and size. Simulations include the two features that make the OLS and OLS FE estimators biased: differences in ability across funds, and a contemporaneous correlation between fund size and returns. We simulate benchmark-adjusted fund returns from equation (2). We simulate fund size as follows:

$$\frac{q_{it}}{q_{it-1}} - 1 = c + \gamma R_{it} + v_{it} . \quad (10)$$

Parameter $\gamma > 0$ captures the positive time-series correlation between returns and fund size, which induces a bias in the OLS FE estimator. Equations (2) and (10) imply that higher-ability funds tend to grow larger due to their higher average returns. The resulting positive cross-sectional correlation between skill and size leads to a bias in the OLS estimator.

To obtain some guidance regarding the parameter values, we run the regression (10) on our data, which we describe later in Section 3. We choose $c = 0.0039$ and $\text{Std}(v) = 0.0566$, which are the OLS estimates of these parameters. The point estimate of γ is 0.92; we consider three different values, $\gamma = 0.8, 0.9$, and 1.0. We consider four plausible values of β : 0, -1×10^{-5} , -3×10^{-5} , and -10×10^{-5} . These values produce a wide dispersion in the simulated outcomes. The value of $\beta = -1 \times 10^{-5}$ implies that a \$100 million increase in fund size decreases expected returns by 0.1% per month. We set $\text{Std}(\varepsilon) = 0.0225$, which is the estimate obtained from (2) by using the OLS FE estimator. We simulate a_i , ε_{it} , and v_{it} as independent draws from normal distributions. We draw each fund's skill a_i from a normal distribution with mean 0.2% per month and standard deviation 0.5% per month; these values are close to those we estimate later in the paper. We set funds' starting size to \$250 million, roughly our sample median. We construct 10,000 samples of simulated panel data for 300 funds over 100 months.¹⁰ In each sample, we estimate $\hat{\beta}_{OLS}$, $\hat{\beta}_{FE}$, and $\hat{\beta}_{RD}$.

Table 1 shows the estimation results. Panels A and B show the means and medians of the β estimates across simulated samples. As expected, the simple OLS estimates tend

¹⁰We simulate uncorrelated benchmark-adjusted fund returns, whereas there is some cross-sectional dependence in our actual data, as noted in Section 3. Therefore, we simulate data on fewer funds than in our actual sample, so that the simulated and actual data exhibit similar amounts of independent variation.

to be too high, while the OLS FE estimates tend to be too low. For example, even when the simulated data exhibit no returns to scale (i.e., the true $\beta = 0$), simple OLS estimates indicate increasing returns to scale, while the OLS FE estimates indicate decreasing returns to scale. Bias is typically more severe for simple OLS than for OLS FE. Bias in the OLS FE estimates is typically larger when the contemporaneous relation between returns and size (γ) is stronger, as expected. The RD estimator produces essentially no bias. For instance, when $\beta = 0$, both the mean and median RD estimates round to 0.00 for all three values of γ . For $\beta \neq 0$, the mean and median RD estimates are also very close to the true values.

Panel C of Table 1 shows the fraction of simulations in which we reject the null hypothesis, $\beta = 0$, at the 5% confidence level. Both the OLS and OLS FE estimators almost always produce false positives, rejecting the null in 98 to 100% of simulations when the null is actually true. In contrast, the RD estimator has approximately the right size, rejecting a true null 6% of the time in the 5% test. The RD estimator also possesses nontrivial power to reject the null when the null is false. For example, when $\beta = -3 \times 10^{-5}$, RD rejects the null of $\beta = 0$ about 20% of the time. The OLS estimators reject the same null almost 100% of the time, but they do so regardless of whether the null is true or not.

To summarize, both OLS estimators are biased and much too eager to reject the null of no returns to scale even when the null is true. In contrast, the RD estimator has virtually no bias, nontrivial power, and approximately the right size.

3. Data

The data come from CRSP and Morningstar. The sample contains 3,126 actively managed domestic equity-only mutual funds from the United States between 1979 and 2011. A 34-page Data Appendix on the authors' websites supplements the information below.

We require that funds appear in both CRSP and Morningstar, which offers several benefits. First, it allows us to check data accuracy by comparing the two databases, as detailed below. Second, Morningstar assigns each fund a category (e.g., large growth, Japan stock, muni California intermediate), which helps us classify funds. Finally, Morningstar designates a benchmark portfolio to each fund and provides benchmark returns. Since Morningstar chooses benchmarks based on funds' holdings rather than their reported objective, the Morningstar benchmark does not suffer from the cherry-picking bias of Sensoy (2009). We start the sample in 1979, the first year in which Morningstar provides benchmark returns. We merge CRSP and Morningstar using funds' tickers, CUSIPs, and names. We check the

accuracy of each match by comparing assets and returns across the two databases.

We use keywords in the Morningstar Category variable to exclude bond funds, money market funds, international funds, funds of funds, industry funds, real estate funds, target retirement funds, and other non-equity funds. We also exclude funds identified by CRSP or Morningstar as index funds, as well as funds whose name contains “index.” We exclude fund/month observations with expense ratios below 0.1% per year, since it is extremely unlikely that any actively managed funds would charge such low fees. Finally, we exclude fund/month observations with lagged fund size below \$15 million in 2011 dollars. A \$15 million minimum is also used by Elton, Gruber, and Blake (2001), Chen et al (2004), Yan (2008), and others.

Berk and Binsbergen (2012, hereafter “BB”) carry out a major data project to address problems with the CRSP mutual fund data. We apply many of BB’s data-cleaning steps, stopping short of steps that require manual searches of data from Bloomberg or the SEC. To be conservative, we require that CRSP and Morningstar agree closely on the two key variables in our analysis, returns and fund size. First, we follow BB in reconciling return data between CRSP and Morningstar. Returns differ across the two databases by at least 10 bp per month in 3.1% of observations. By applying BB’s algorithm we reduce the discrepancy rate to 0.6%. We set the remaining return discrepancies to missing. Similarly, total assets under management (AUM) differ between CRSP and Morningstar in 7.3% of observations, even allowing for rounding errors.¹¹ The average of these discrepancies is \$12.3 million. AUM differs by at least \$100,000 and 5% across databases in 1.0% percent of observations; we set these AUM values to missing, otherwise we use CRSP’s value.

We depart from BB’s sample construction somewhat, since we use different Morningstar data. BB purchase every monthly data update from Morningstar starting in January 1995, whereas we use Morningstar’s most recent historical file, which includes data back to 1924. While BB use the union of CRSP and Morningstar, we use the intersection, which allows us to cross-check all observations’ accuracy across the two sources. Besides being significantly less expensive, our Morningstar data include useful additional variables such as CUSIP (which we use to merge CRSP and Morningstar), Category (which we use to categorize funds and assign benchmarks), and FundID (which we use to aggregate share classes).¹²

¹¹BB report a discrepancy rate of 16%. One potential reason for their higher rate is that BB use monthly data updates from Morningstar, whereas we use Morningstar’s single historical database. It is possible that Morningstar corrected errors from the monthly updates when compiling them into the historical database.

¹²Many mutual funds offer multiple share classes, which represent claims on the same underlying assets but have different fee structures. Different share classes of the same fund have the same Morningstar FundID. We aggregate all share classes of the same fund. Specifically, we compute a fund’s AUM by summing AUM across the fund’s share classes, and we compute the fund’s returns, expense ratios, and turnover by asset-weighting

We now define the variables used in our analysis. Summary statistics are in Table 2.

Our measure of fund performance is $GrossR$, the fund’s monthly benchmark-adjusted gross return. We use gross rather than net returns because our goal is to measure a manager’s ability to outperform a benchmark, not the value delivered to clients after fees. $GrossR$ equals the fund’s net return plus its monthly expense ratio minus the return on the benchmark index portfolio designated by Morningstar.¹³ We take expense ratios from CRSP because Morningstar is ambiguous about their timing. The average of $GrossR$ is +5 bp per month, whereas the average benchmark-adjusted net return is −5 bp per month.

The average pairwise correlation in $GrossR$ between funds belonging to the same Morningstar Category is 0.15. To account for these cross-sectional correlations in our subsequent regressions, we cluster standard errors by Morningstar Category \times month. The average correlation between funds from different categories is only 0.04; therefore, we do not cluster by month to avoid adding noise to standard errors. In our RD specifications we also cluster by fund since recursive demeaning can potentially induce serial correlation within funds.

$FundSize$ corresponds to q_{it-1} in the previous section. $FundSize$ equals the fund’s AUM at the end of the previous month, inflated to December 2011 dollars by using the ratio of the total market value of all CRSP stocks in December 2011 to its value at the end of the previous month. The advantage of this inflator is that it makes $FundSize$ capture the size of the fund relative to the universe of stocks that the fund can buy, a reasonable way to measure the limitations on a fund due to its size. There is considerable dispersion in $FundSize$: the inner-quartile range is \$84 million to \$921 million.

The top panel of Figure 1 shows the number of funds in our sample over time. The number of funds with non-missing returns increases from 145 in 1979 to 1,574 in 2011. Comparing the black and blue lines, we see that we lose some observations because of missing expense ratios or benchmark returns. The sawtooth pattern in the red line shows that many funds report AUM only quarterly or yearly before March 1993, which we denote with a vertical dashed line. The middle panel shows another change around March 1993: CRSP and Morningstar report similar expense ratios starting in 1993, whereas they often disagree before then. We also see large jumps in expense ratios in both databases before 1993. Overall, the data appear to be more reliable starting in March 1993. For this reason, we use the period from

across share classes. We take the fund’s age to be the maximum age across the fund’s share classes.

¹³This index-based benchmark adjustment is likely to adjust for fund style and risk more precisely than the commonly-used loadings on Fama-French and Carhart factors. Cremers, Petajisto, and Zitzewitz (2013) argue that the popular Fama-French and Carhart adjustments produce biased assessments of fund performance. The same authors recommend using index-based benchmarks, and find that using such benchmarks better explains the cross-section of mutual fund returns. We follow this advice.

March 1993 to December 2011 as our main sample. We also report results from the extended sample that begins in January 1979. Since there are fewer funds and more missing values before 1993, extending the sample back to 1979 increases its size by only 11%.

IndustrySize is the sum of AUM across all funds in our sample, divided by the total market value of all stocks (i.e., the sum of *FundSize* across all sample funds, up to a constant). It is the fraction of total stock market capitalization that the sample’s mutual funds own at that time. When computing *IndustrySize*, we fill in missing values of *FundSize* by taking the fund’s most recent reported size and updating it by using interim realized total fund returns.¹⁴ The red line in the top panel of Figure 1 shows that without this adjustment, we would obtain a downward-biased, sawtooth pattern in *IndustrySize* before March 1993. The bottom panel of Figure 1 plots *IndustrySize* over time. It starts at 2.4% in January 1979, peaks at 18.6% in July 2008, and finishes at 16.8% in December 2011.

The variables defined above—*GrossR*, *FundSize*, and *IndustrySize*—are the main variables used in our empirical analysis of returns to scale. The remaining variables from Table 2 are defined later, in Section 4., as soon as they are first introduced.

4. Empirical results

4.1. Fund-level returns to scale?

To investigate whether there are returns to scale at the fund level, we run panel regressions of fund i ’s benchmark-adjusted gross return in month t , $GrossR(i, t)$, on the fund’s size at the end of the previous month, $FundSize(i, t - 1)$. We consider three regression approaches: plain OLS, OLS with fund fixed effects (OLS FE), and recursive demeaning (RD). All three approaches are discussed in detail in Section 2.: simple OLS corresponds to equation (1), OLS FE to equation (2), and RD to equation (9). We report the results in the first three columns of Table 3. Panel A reports the results from our main sample (1993–2011), whereas Panel B focuses on the extended sample (1979–2011).

In the pooled OLS specification, the estimated coefficients on *FundSize* are negative,

¹⁴We assume no flows in or out of the fund since its last reported AUM. For example, if the fund’s size was \$100 a month ago and the fund then experiences a 10% total return, we impute the current size of \$110. To avoid imputing an AUM for a dead fund, we impute only if the fund reports a return during the given month. We do not look more than 12 months back for a non-missing AUM. Imputing fund size introduces measurement error in *IndustrySize*, but such error would be worse if we were to simply set the missing fund sizes to zero. Note that we only fill in missing values of fund size when computing *IndustrySize*; we do not do so when we use *FundSize* on its own, so there should be no measurement error in *FundSize*.

with t -statistics in the neighborhood of -2, but the coefficient values are economically trivial in both the main and extended samples. Consider a \$100 million increase in fund size, which is substantial as it represents almost a 40% increase in the size of the median fund in our sample (Table 2). The coefficient estimates indicate that such an increase in size is associated with a decrease in expected fund performance of only 0.00014% per month, or 0.17 bp per year. Recall however that the plain OLS estimator is biased upward if skill and size are cross-sectionally correlated, so economic significance of the OLS estimate is likely to be understated. Chen et al. (2004) make a similar observation when obtaining significantly negative estimates under this specification.

In contrast, recall that the bias in the OLS FE estimator is negative, opposite the bias in the plain OLS estimator. Despite this downward bias, the negative OLS FE coefficients in Table 3 are again economically small, indicating that a \$100 million increase in fund size lowers expected return less than 0.0017% per month, or about two bp per year. Therefore, although the OLS FE results produce greater statistical significance, with t -statistics of about -9, economic significance remains small despite its likely overstatement due to the negative bias. We thus see mixed evidence of fund-level decreasing returns coming from the two OLS procedures, each of which is biased but in different directions.

To avoid these biases, we apply the bias-free RD procedure from Section 2.2. We find that the estimated effect of fund size on performance is no longer statistically significant, with t -statistics of about -0.6 (see column 3 of Table 3). The effect is not economically significant either. The estimate from Panel A indicates that a \$100 million increase in fund size depresses performance by 0.0022% per month, or about 2.5 bp per year. In Panel B, the same increase in fund size depresses performance by only 1.3 bp per year.

In sum, we do not find consistent evidence of decreasing returns to scale at the fund level. The biased OLS procedures indicate a negative relation between a fund's size and its performance, but the unbiased RD procedure detects no significant relation. All three procedures produce economically small estimates of the effect of fund size on performance. We show later that these findings are unaffected by including numerous controls such as industry size, sector size, family size, fund age, and fund turnover.

4.2. Industry-level returns to scale?

To explore potential returns to scale at the industry level, we run panel regressions of $GrossR(i, t)$ on $IndustrySize(t - 1)$. We consider the same panel regression approaches

as before: OLS, OLS FE, and RD. The results are in columns 4 through 6 of Table 3.

In the plain OLS specification, the estimated coefficient on *IndustrySize* is negative and marginally significant, with t -statistics of -1.9 in both panels. This evidence is suggestive of decreasing returns to scale at the industry level. However, since the plain OLS specification does not allow for differences in skill across funds, we cannot treat this evidence as conclusive.

To allow for differences in skill, we add fund fixed effects (see column 5 of Table 3). The evidence of decreasing returns to scale then becomes stronger: the estimated coefficients on *IndustrySize* roughly double and the t -statistics drop to -3.6 in Panel A and -4.3 in Panel B.¹⁵ The effect is not only statistically but also economically significant. For example, consider a one percentage point increase in *IndustrySize*, which is not too large.¹⁶ In our main sample, this increase in *IndustrySize* is associated with a sizable decrease in fund performance: 0.0326% per month, or almost 40 bp per year. In the extended sample, the magnitude of the same effect is smaller but still substantial, about 20 bp per year.

The RD estimates of the relation between *GrossR* on *IndustrySize* are shown in column 6 of Table 3. The point estimates are virtually identical to those in column 5 and even though the t -statistics are smaller, the relation remains statistically significant. As noted earlier in Section 2.2., *IndustrySize* can instrument for itself in the RD procedure because it is not plagued by the bias-inducing correlation between the error term and the regressor in equation (2). In particular, there is no reason to believe that innovations in *IndustrySize* are correlated with the benchmark-adjusted returns of any given fund.¹⁷ Therefore, the RD procedure in this case is simply the OLS regression of forward-demeaned *GrossR* on forward-demeaned *IndustrySize*. Since there is no need for a backward-demeaned instrument, forward-demeaning is unnecessary as well. We report the results from the RD procedure only for comparison with the other approaches. The relation between *GrossR* and *IndustrySize* is better captured by the OLS FE results in column 5.

In columns 7 through 9 of Table 3, we run a horserace by regressing $GrossR(i, t)$ on

¹⁵To calculate standard errors, we cluster by sector \times month to allow for potential correlation of benchmark-adjusted fund returns across funds, as explained in Section 3. We do not cluster by fund in this OLS FE specification because there is very little serial correlation within funds: the first ten residual autocorrelations are all smaller than 0.05 in absolute value. If we were to add clustering by fund to address the serial correlation in the residuals, the t -statistics on *IndustrySize* would change from -3.60 to -3.58 in Panel A and from -4.34 to -4.24 in Panel B.

¹⁶Recall that *IndustrySize* is the total AUM of active mutual funds divided by the stock market capitalization. According to Table 2, changing *IndustrySize* by 1% represents movement of only about one fifth of the interquartile range, one seventh of the median, and one sixteenth of the 98-percentile range.

¹⁷Indeed, the R-squared from a panel regression of fraction changes in *IndustrySize* on benchmark-adjusted fund returns is only 0.006. The R-squared is almost 20 times larger, 0.110, if we replace *IndustrySize* with *FundSize* in the regression. We winsorize the regressor at the 1st and 99th percentiles.

both $FundSize(i, t - 1)$ and $IndustrySize(t - 1)$ under all three approaches. We find that $FundSize$ enters significantly in the first two approaches, but its significance disappears in the bias-free RD approach. In contrast, the coefficient on $IndustrySize$ remains negative and significant, and its magnitude is similar to column 5 where $FundSize$ is excluded.

The negative relation between fund performance and industry size emerges not only from the panel regressions in Table 3 but also from simple fund-by-fund regressions. For each fund i , we run the time-series regression of $GrossR(i, t)$ on $IndustrySize(t - 1)$. In our main sample, we find that 62% of the funds' OLS slope estimates are negative, and 9% (4%) are negative and significantly different from zero at the 5% (1%) two-sided confidence level. In the extended sample, the results are very similar: 61% of the estimates are negative, and 10% (4%) are significantly negative at the 5% (1%) confidence level.

To summarize, we find a strong negative relation between fund performance and industry size. This relation, which is both economically and statistically significant, is consistent with the presence of decreasing returns to scale at the industry level.

Table 3 presents results from three different methods, only one of which, RD, removes the bias inherent in estimating fund-level returns to scale. Unless noted otherwise, from now on we report only the bias-free results based on RD for any panel regression that involves $FundSize$ (Tables 5, 6, 10, and 11). When the regression includes no variable involving $FundSize$, so that the bias is not an issue, we report the OLS FE results.

4.3. A closer look at industry size

Recall from Figure 1 that $IndustrySize$ trends upward for most of the sample period. This trend is nonmonotonic—for example, $IndustrySize$ decreases in the late 1990s as well as from 2009 to 2011—but it is clearly present. One might wonder whether $IndustrySize$ is simply capturing a time trend. To address this issue, we define a time trend variable as the number of months elapsed since January 1979. When we run an OLS FE panel regression of $GrossR$ on the linear time trend, we indeed find a significantly negative relation, as shown in column 2 of Table 4. To separate time from $IndustrySize$, we include both variables on the right-hand side of the OLS FE regression. We find that $IndustrySize$ retains its significantly negative slope coefficient in the main sample, and the coefficient's estimated value becomes substantially more negative: -0.0852 , compared to -0.0326 when the time trend is excluded (compare columns 1 and 3 of Panel A of Table 4). In contrast, the sign of the estimated coefficient on the time trend flips from negative to positive. Fund performance therefore

seems negatively related to *IndustrySize* instead of being a simple function of time.

Industry size fluctuates over time as a result of changes in the number of active mutual funds as well as changes in the average fund size. Which of the two components drives the negative relation between industry size and fund performance? To answer this question, we define two new variables: *Number of Funds*, which is a count of the sample funds operating in the given month, and *Average Fund Size*, which is the average AUM across all sample funds in that month, inflated to current dollars. We perform this inflation by dividing the average AUM by the total stock market capitalization in the same month and then multiplying by the total stock market capitalization at the end of 2011. Note that *IndustrySize* equals *Number of Funds* times *Average Fund Size* divided by a constant, namely, the total stock market capitalization at the end of 2011.

Table 4 shows that both components of *IndustrySize* contribute to the negative size-performance relation. When the new variables are included individually in our OLS FE panel regression, *Average Fund Size* exhibits a significantly negative relation with *GrossR* whereas *Number of Funds* is insignificant. When the two variables are included together, though, both of them enter with significantly negative coefficients. Interestingly, both variables lose their statistical significance when *IndustrySize* is also included in the regression (see the last two columns of Table 4). This result suggests that *IndustrySize* does a good job of capturing the joint effect of its two components on fund performance.

4.4. Determinants of the size-performance relation

In this subsection, we take a closer look at the size-performance relation by analyzing its dependence on fund characteristics. We examine three characteristics that have some a priori relevance for the size-performance relation: a small-cap indicator, volatility, and turnover.

The first characteristic, $1(SmlCap)$, is a dummy variable that is equal to one if the fund is classified by Morningstar as a small-cap fund (i.e., a fund trading small-capitalization stocks) and zero otherwise. About 19% of our funds are small-cap funds. The second characteristic, *Turnover*, is the fund's turnover, expressed as a fraction per year. We obtain turnover data from CRSP if available, otherwise from Morningstar. To remove some implausible outliers, we winsorize turnover at its 1st and 99th percentiles. Median turnover is 65% per year. The third characteristic, $Std(AbnRet)$, is the standard deviation of a fund's abnormal returns, expressed as a fraction per month. Abnormal returns are the residuals from the regression of the fund's excess gross returns on excess benchmark returns.

Why might these characteristics affect the size-performance relation? The effect of scale on a fund’s performance is likely to depend on the liquidity of the fund’s assets. Lower liquidity implies a larger price impact for a trade of a given size. Therefore, lower liquidity is likely to make a fund’s returns decrease in scale more steeply. This relation can in principle hold both at the fund level and at the industry level. It can hold at the fund level because a larger fund trades larger amounts, leading to a larger price impact. It can also hold at the industry level because in a more crowded industry, there are likely to be more active funds chasing the same investment opportunities and pushing prices in the same direction.¹⁸ This logic suggests that if there are decreasing returns to scale at either the fund or industry level, they should be decreasing more steeply for both small-cap funds and high-turnover funds, both of which are likely to face larger total price impact costs.

Higher-volatility funds might also exhibit steeper decreasing returns to scale. The reason is that funds with more volatile benchmark-adjusted returns are effectively larger in terms of their trading. Note that a fund’s portfolio can be thought of as a combination of a (potentially levered) benchmark investment and a zero-cost long-short “active” portfolio whose return is uncorrelated with the benchmark return. Since benchmark exposure can be managed cheaply, the cost of managing the fund depends largely on the size of the active portfolio. Given this portfolio’s zero-cost nature, a reasonable measure of its size is its dollar volatility, which is the product of the active portfolio’s volatility and fund size. For example, a “closet indexing” fund looks small by this metric, whereas an equal-sized fund that takes big active bets looms larger. Funds with more volatile active portfolios are likely to face larger trading costs and, consequently, steeper decreasing returns to scale.

To examine these hypotheses, we run panel regressions analogous to those in Table 3, except that we add the interactions of both *FundSize* and *IndustrySize* with $1(SmlCap)$, *Turnover*, and *Std(AbnRet)*. The results for our main sample are in Table 5. The point estimates of the terms that involve *FundSize* generally go in the direction hypothesized earlier, in that the negative effect of fund size on performance is more pronounced for small-cap funds, high-volatility funds, and high-turnover funds. However, none of those terms, interacted or not, are even close to being statistically significant.

In contrast, Table 5 shows significant interactions between *IndustrySize* and fund characteristics. When $IndustrySize \times 1(SmlCap)$ is added to *IndustrySize* on the right-hand side (column 2), it enters negatively and significantly, indicating that industry-level decreasing returns to scale are more pronounced for small-cap funds. The coefficient loses statistical significance when fund-level variables are added (column 3), but its magnitude remains about

¹⁸Pástor and Stambaugh (2012) also argue that industry-level returns to scale are induced by illiquidity.

the same as in column 2. The other two interaction terms are even more significant. The term $IndustrySize \times Std(AbnRet)$ enters negatively (columns 5 and 6), indicating more steeply decreasing returns to scale for more volatile funds. Similarly, columns 8 through 10 indicate steeper decreasing returns to scale for funds with higher turnover.

To summarize, Table 5 shows that industry size depresses fund performance especially for funds that have high volatility and high turnover. The negative relation between industry size and fund performance is also marginally stronger for small-cap funds. The results in Table 6, which are based on the extended sample, lead to the same conclusions.

The literature reports mixed evidence on the relation between fund performance and turnover. For example, Carhart (1997) finds a negative relation whereas Wermers (2000) finds a marginally positive relation. Tables 5 and 6 show a significant positive association between fund turnover and performance. Unlike the earlier studies, we include fund fixed effects in our panel regressions. Therefore, our results derive their power from the time series, not from the cross section. We do not find that higher-turnover funds perform better; instead, we find that a given fund tends to perform better when it trades more.

4.5. The evolution of fund skill

We now analyze our estimates of fund skill. A natural measure of skill in our framework is the fund fixed effect from equation (4). This measure represents the average benchmark-adjusted gross fund return that is further adjusted for any potential fund-level and industry-level returns to scale. Therefore, our fixed-effect measure of skill can be thought of as the standard gross alpha adjusted for returns to scale, or alpha earned on the fund's first dollar when there is no competition from the industry. We focus on the all-inclusive specification of regression (4) whose results are reported in the last column of Table 6. We choose the estimates based on the extended sample, because we want to plot the time series of fund skill over the longest possible sample period.

Figure 2 shows how the cross-sectional distribution of fund skill varies over time. For each month in our extended sample, January 1979 through December 2011, the figure plots the average as well as the percentiles of the estimated fund fixed effects across all funds operating during that month. The most distinctive feature of the figure is the upward trend in the distribution of skill. For example, the mean fixed effect grows from -5 bp per month at the beginning of the sample to +13 bp per month at the end. The growth is more pronounced at the top: the 90th percentile grows from 51 bp to 88 bp per month. This evidence suggests

that funds have become more skilled over time, especially the above-average funds.

Does this improvement in skill translate into better performance? The answer is no, according to Figure 3. This figure plots two-year moving averages of equal-weighted average fund returns. The solid red line shows the average benchmark-adjusted gross return (*GrossR*). This line does not exhibit any obvious trend, certainly not an upward trend, suggesting that fund performance has failed to improve over time.

How can we reconcile the upward trend in skill (Figure 2) with the lack of a trend in performance (solid line in Figure 3)? Major clues appear in two results discussed earlier: the upward trend in industry size (Figure 1) and the negative relation between industry size and fund performance (Table 3). Taken together, these two results imply that the growing industry size makes it more difficult for managers to outperform despite their improving skill. To illustrate this point, we plot one more line in Figure 3. The dashed black line is analogous to the solid red line except that it adds 0.0326 times *IndustrySize* in each month. This adjustment represents compensation for the adverse effect of the growing industry size on realized fund performance. The coefficient 0.0326 is minus the estimated slope from the regression of *GrossR* on fund fixed effects and *IndustrySize* (column 5 in Panel A of Table 3). The gap between the two lines in Figure 2 grows from 6 bp per month in 1979 to roughly 50 bp per month in 2011. These estimates suggest that average benchmark-adjusted fund returns in 2011 would have been 44 bp per month higher if industry size had counterfactually stayed at its 1979 level instead of growing, all else equal. Consistent with rising skill, the dashed line even trends upward somewhat, especially between 1985 and 2000, when our estimated average skill increases the most.

Why has average fund skill risen over time? We can rule out the explanation that a given fund's skill improves over time, perhaps as a result of learning on the job. We can rule it out because our measure of a fund's skill, the fund fixed effect, is time-invariant. For that reason, the answer must involve changes in the composition of the fund universe. A natural explanation is that the new funds entering the industry are more skilled, on average, than the existing funds. The higher skill of the new arrivals could result from better education of the new managers, for example, or from their superior mastery of new technology. In addition to new funds being more skilled, it is also possible that the funds exiting the industry are less skilled, on average, but this exit-based effect is unlikely to play a leading role since fund entry has far exceeded exit over time (Figure 1).

To summarize, we find that funds have become more skilled over time, yet this improvement in skill has failed to boost fund performance. This evidence is consistent with the

observed gradual growth in industry size, which has had an adverse effect on fund performance due to decreasing returns to scale.

4.6. Performance erosion over a fund’s lifetime

If industry size grows while a fund’s ability stays constant, then performance should erode over the fund’s lifetime. To test this prediction, we run a panel regression of benchmark-adjusted gross fund returns (*GrossR*) on fund fixed effects as well as fund age dummy variables. For any given fund and age t , $t = 1, 2, \dots, 20$ years, the age dummy variable is equal to one in the fund’s t -th year of operations and zero otherwise. We measure fund age by the number of years since the fund’s first offer date (from CRSP) or, if missing, since the fund’s inception date (from Morningstar). We estimate this panel regression in our main sample and show the results in Figure 4.¹⁹ Specifically, Figure 4 plots the fund age fixed effects along with their 95% confidence interval. For any age t , $t = 1, 2, \dots, 20$ years, the age- t fixed effect measures the difference between a fund of age t and the same fund at age > 20 years in terms of their average *GrossR*.

Figure 4 shows that fund performance declines over a typical fund’s lifetime. This new result is almost monotonic for fund ages up to 12 years. The point estimates of the age fixed effects decline in an approximately linear fashion from 37 bp per month at age one to zero at age 12, after which they are roughly flat. These estimates are positive and statistically significant up to age six, indicating that up to this age, fund performance is significantly higher than it is at ages exceeding 20 years. Since the regression includes fund fixed effects, the decline observed in Figure 4 represents a within-fund rather than across-fund pattern.²⁰ In short, as funds get older, their performance tends to suffer.

Further support for this negative age-performance relation comes from a panel regression of *GrossR* on fund age. Whereas Figure 4 uses age fixed effects to measure the nonparametric relation between *GrossR* and age, the regression assumes a linear relation. We include fund fixed effects in the regression, as before, to focus on the variation in performance over a given fund’s lifetime. The results are presented in column 1 of Table 7. Again, we find a negative and significant relation between a fund’s age and its performance, with t -statistics equal to -3.0 in the main sample and -4.0 in the extended sample. The point estimate of the slope coefficient on fund age in the main sample indicates that one additional year of age reduces

¹⁹The results based on the extended sample are very similar; they are not reported here.

²⁰Chen et al. (2004), Yan (2008), and Ferreira et al. (2013a) control for fund age in their performance regressions. None of these studies find a significant coefficient on fund age for U.S. funds, but they do not include fund fixed effects. Their datasets are also different; for example, they do not use Morningstar data.

the fund’s gross benchmark-adjusted return by 1.23 bp per month, or 15 bp per year.

One potential concern is that the negative relation between fund age and performance could be driven by the incubation bias documented by Evans (2010). Incubation is a strategy that some families follow to initiate new funds. A family might start multiple funds privately, with a limited amount of capital. At the end of an evaluation period, it might open only some of these funds, often those with better performance, to the public. Evans finds that the incubated funds outperform the non-incubated funds during the incubation period.

Our age result does not seem to be driven by the incubation bias. First of all, our \$15 million fund size screen eliminates many incubated funds. More important, Evans (2010) reports that “removing the first three years of return data for all funds eliminates the bias.” Motivated by this finding, we rerun the regression of *GrossR* on fund age after excluding the returns of funds younger than three years. The results, which are reported in column 4 of Table 7, are quite similar to those from column 1 in which all funds are included. The estimates remain statistically significant in both panels, and they are only slightly smaller in magnitude compared to column 1. In addition, recall from Figure 4 that the estimates of the age fixed effects are positive and statistically significant up to age six. In the extended sample, the estimates are significantly positive up to age eight (not plotted). All of this evidence makes the incubation bias an unlikely explanation for our fund age result.

This result is also unlikely to be due to risk. Equity mutual funds’ market betas are well known to be close to each other as well as close to one (e.g., Chen et al, 2004), and any risk associated with size or value exposures is likely to be removed by our benchmark adjustment. Moreover, for risk to explain why funds tend to perform better when younger, risk exposure would have to decline with age. In contrast, Chevalier and Ellison (1999) argue that younger managers have an incentive to avoid risk because they are more likely to be fired for bad performance. Consistent with this argument, Chevalier and Ellison find that younger managers tend to hold less risky and more conventional portfolios.

Why, then, does fund performance erode over a typical fund’s lifetime? Our earlier results offer the following narrative. New funds entering the industry tend to be more skilled than the existing funds (Figure 2). Therefore, new funds tend to outperform their benchmarks initially. However, as a given fund ages, the industry keeps growing (Figure 1), and the continued arrival of skilled competition depresses the fund’s performance (Table 3).

According to this narrative, the source of the erosion in fund performance is the growth in industry size during the fund’s lifetime. The negative age-performance relation should thus disappear after controlling for *IndustrySize*. To test this idea, we add *IndustrySize*

on the right-hand side of the OLS FE regressions in Table 7 (see columns 2 and 5). Indeed, we find that adding *IndustrySize* annihilates the negative relation between *GrossR* and fund age. While *IndustrySize* enters with a negative coefficient, the slope on fund age is no longer significantly negative; in fact, it turns marginally positive. We reach the same conclusions when we control for *FundSize* (columns 3 and 6). These results lend further support to the narrative from the previous paragraph.²¹

4.7. Age-based investment strategies

Since funds born more recently appear to exhibit more skill, we expect younger funds to outperform older funds in a typical month. We find evidence supporting this prediction after examining the returns on age-based investment strategies. Each month, we assign funds to four portfolios based on fund age: $[0, 3]$, $(3, 6]$, $(6, 10]$, and > 10 years. We calculate the portfolios' equal-weighted average benchmark-adjusted gross and net returns over the following month, at the end of which we rebalance. Table 8 shows the average returns of the age-sorted portfolios, along with return differences across the portfolios.

Table 8 shows that younger funds tend to outperform older funds, especially based on gross returns. All six young-minus-old differences in gross returns are positive, and four of them are statistically significant. The youngest funds (aged ≤ 3 years) outperform the oldest funds (aged > 10 years) by a statistically significant 7.2 bp per month, or 0.9% per year, in the main sample. While the returns of the youngest funds might potentially be boosted by the incubation bias, the funds aged between three and six years, which are immune to this bias according to Evans (2010), also outperform the oldest funds, by the statistically significant amount of over 0.5% per year. For the extended sample, the return difference between the $(3, 6]$ and > 10 portfolios is even larger than that between the $[0, 3]$ and > 10 portfolios (9.6 bp versus 7.5 bp per month), suggesting that the incubation bias is not responsible for our results. For that sample, even the return difference between the $(6, 10]$ and > 10 portfolios is statistically significant. Finally, an F -test, whose p -values are reported in the last column of Table 8, rejects the null hypothesis that the average benchmark-adjusted returns are equal across the four age-sorted portfolios. We thus conclude that younger funds tend to outperform older funds based on gross benchmark-adjusted returns.

Based on net returns, the young-minus-old portfolio differences tend to be smaller. The point estimates are positive in five of the six cases, but they are statistically significant only

²¹Also supporting the narrative is the fact that if we add *IndustrySize* as a control on the right-hand side of the regression underlying Figure 4, the downward trend in the figure disappears and none of the age FEs are significantly positive. The results are not plotted, to save space.

between the $[0, 3]$ and > 10 portfolios. Therefore, we cannot reliably conclude that investors buying young funds outperform those buying old funds. The younger funds appear to be able to capture some of their higher skill by charging higher fees. Indeed, the average annual expense ratios in the four age-sorted portfolios in our main sample are 1.35% (youngest funds), 1.30%, 1.27%, and 1.17% (oldest funds).

Recall that the age results in Table 7 are obtained from regressions that include fund fixed effects. Therefore, those earlier results compare the performance of a typical fund at different ages. In contrast, Table 8 compares the performance of young and old funds in a typical month. The negative age-performance relation thus seems to hold not only within but also across funds. Both the within-fund and across-fund performance differences are consistent with our narrative. The across-fund results rhyme well with the notion that the new funds entering the industry tend to be more skilled than the older funds. The within-fund results are consistent with the idea that as a fund grows older, its performance deteriorates as a result of industry growth and the related arrival of skilled competition.

4.8. Robustness

In this subsection, we examine the robustness of our main result regarding the nature of returns to scale. First, we add family size to the right-hand side of the baseline regression specification from Table 3. For any given fund, we compute *FamilySize* by adding up *FundSize* across all funds belonging to the fund’s family, as classified by Morningstar. Median *FamilySize* is \$12.2 billion (Table 2). The addition of *FamilySize* is motivated by Chen et al. (2004), who find a positive relation between a fund’s performance and the size of its fund family. Ferreira et al. (2013a) find a similar result in international funds.

Table 9 is the counterpart of Table 3 for our main sample, with *FamilySize* added to each regression specification. The Fama-MacBeth regressions in Chen et al. (2004) and Ferreira et al. (2013a) are most comparable to our simple OLS estimates without fixed effects. Like those authors, we find a positive relation between returns and family size in those specifications, although the relation is only statistically significant in specification (7). The relation flips to negative and usually insignificant when we include fund fixed effects. The coefficient on *FundSize* is negative but insignificant in the relevant RD specifications, just like in Table 3. Finally, *IndustrySize* continues to enter with a significantly negative slope, as in Table 3. In short, the addition of family size does not alter any of our conclusions.

The same conclusions continue to hold when we replace *FamilySize* with fund age. Recall

from Table 7 that when fund age is included in the regression together with *IndustrySize* and *FundSize*, the coefficient on *FundSize* is negative but insignificant, whereas the coefficient on *IndustrySize* is significantly negative, just like in Table 3.

Table 10 uses different functional forms of *FundSize*, specifically the natural logarithm of *FundSize* and *FundSize* squared, in the baseline regression specification from Table 3. The motivation is that while *FundSize* is not significantly related to fund returns in Table 3 under the bias-free RD procedure, a different functional form of *FundSize* could potentially matter for performance. However, Table 10 shows that none of the functional forms of *FundSize* enters with a significant coefficient. In contrast, *IndustrySize* remains significant in all specifications (albeit marginally so in column 2).

Finally, Table 11 adds sector size on the right hand side of the baseline regression. The idea is that if decreasing returns to scale are driven by competition with other funds, then funds in the same sector, which presumably follow similar investment strategies, should matter more than funds in other sectors. A fund’s performance should therefore be more closely related to the size of the fund’s sector than to the size of the entire industry. To evaluate this idea, we measure *SectorSize* by adding up fund sizes across all funds within a given sector, divided by the total market value of all stocks belonging to that sector.

We consider two versions of *SectorSize*. The first version uses the nine sectors corresponding to Morningstar’s 3 × 3 stylebox (small growth, mid-cap value, etc.).²² The number of funds in these sectors ranges from 126 in small value to 653 in large growth. The second version uses only three size-based sectors: large-cap, mid-cap, and small-cap.²³ This simpler version of *SectorSize* is coarser than the first version but it is immune to the difficulties associated with labeling funds as following value, growth, and blend styles.

Table 11 shows that neither version of *SectorSize* exhibits a negative and significant relation with fund performance. In fact, the estimated slope coefficients on *SectorSize* are positive, especially when *IndustrySize* is also included in the regression. While *FundSize* is insignificant throughout, *IndustrySize* is negative and significant. Interestingly, the addition

²²Details are in the Data Appendix. We allocate all CRSP stocks to the Morningstar 3×3 matrix using Ken French’s 10×10 portfolios sorted by size and book-to-market. We use a 3-4-3 split for firm size and a 5-5 split for growth vs. value. Note that the three Morningstar size categories (small, mid-cap, and large) are mutually exclusive, whereas the “blend” category overlaps with both growth and value. For example, *SectorSize* for small-growth funds equals [size of small-growth funds + (1/2) size of small-blend funds] / [mkt. cap. of the 15 Fama-French portfolios in the bottom-3 size and bottom-5 B/M portfolios]. As another example, *SectorSize* for small-blend funds equals [size of all small-cap funds (blend+growth+value)]/[mkt. cap. of the 30 Fama-French portfolios in the bottom 3 size portfolios].

²³For example, *SectorSize* for small-cap funds is the total size of all small-cap funds divided by the total market cap of all stocks in the bottom 3 Fama-French size deciles.

of *SectorSize* makes the slope coefficient on *IndustrySize* even more negative compared to Table 3. These results indicate that decreasing returns to scale operate at the industry level rather than sector level.

To summarize, while the idea of sector-level decreasing returns to scale seems sensible, we do not find support for it in the data. Of course, this negative result may very well stem from a measurement problem—our proxies for sector size may not accurately measure the size of a fund’s competition. For example, a large-cap growth fund might compete with certain mid-cap growth funds as well. Our simple measures of sector size might just be too noisy to be useful. A more accurate measurement of a fund’s competition could potentially reveal decreasing returns to scale operating at the sector level. We hope that such a task will be undertaken by future research.

5. Conclusions

We empirically analyze the interaction between skill and scale in active mutual fund management. We identify two econometric biases that plague the OLS estimates of fund-level returns to scale. Our alternative bias-free estimates do not reveal decreasing returns to scale at the fund level. In contrast, we do find strong evidence of decreasing returns at the industry level. This negative relation between industry size and fund performance is stronger for funds with higher turnover, higher volatility, as well as small-cap funds.

Our results on returns to scale shape our assessment of fund manager skill. We measure skill by the fund’s gross alpha before its erosion by returns to scale. We find that newly arrived fund managers tend to be more skilled than existing ones. Despite this rise in average skill, average fund performance has failed to improve. These two facts can be reconciled by industry-level decreasing returns to scale: the observed steady growth in industry size has impeded the funds’ performance despite their improving skill. The growing industry size also helps explain our finding that performance typically declines over a fund’s lifetime. We find that young funds tend to outperform their older peers, consistent with the new entrants being more skilled. However, as funds grow older, their performance tends to deteriorate due to continued industry growth and the associated arrival of skilled competition.

Our study raises interesting new questions. First, what are the sources of the observed upward trend in average skill? We conjecture that this trend is driven by new fund managers who are better educated or better acquainted with new technology, but we provide no direct evidence. Second, it would be useful to shed more light on the mechanism through which

industry size impairs fund performance. Third, despite our progress on the methodology, the evidence on fund-level returns to scale remains inconclusive. Finally, our simple econometric model can be extended in various ways, such as by allowing an individual fund's skill to vary over time. We leave these challenges for future research.

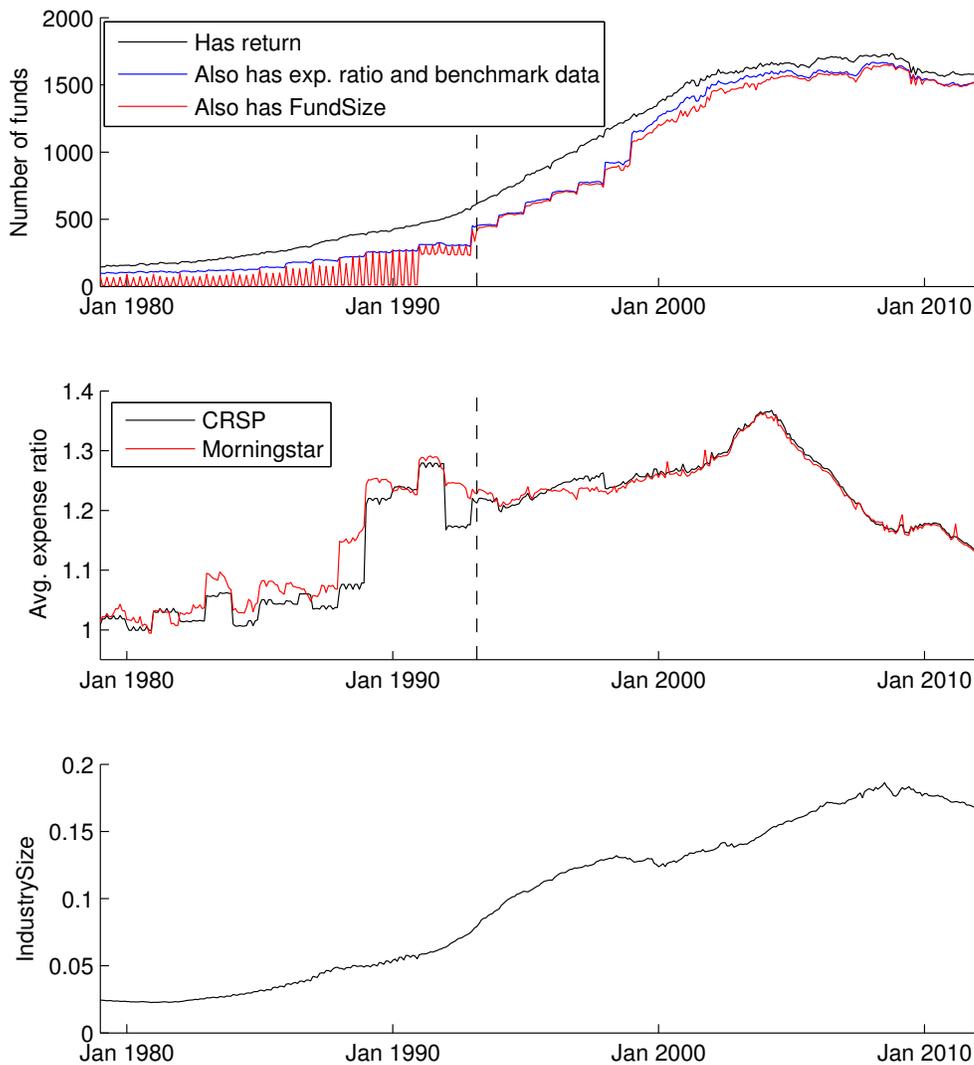


Figure 1. Sample properties. The top panel shows how the number of funds with non-missing observations changes over time. The black line plots the number of funds that have a non-missing cross-checked net return. The blue line plots the number of funds that also have a non-missing CRSP expense ratio and Morningstar benchmark return. The red line plots the number of funds that also have non-missing cross-checked FundSize. The dashed vertical line in the top two panels marks March 1993. The middle panel shows the time series of average CRSP and Morningstar expense ratios, in percent per year, for funds that have expense ratios from both sources. The bottom panel shows *IndustrySize*, the sum of funds' AUM divided by the total market value of CRSP stocks.

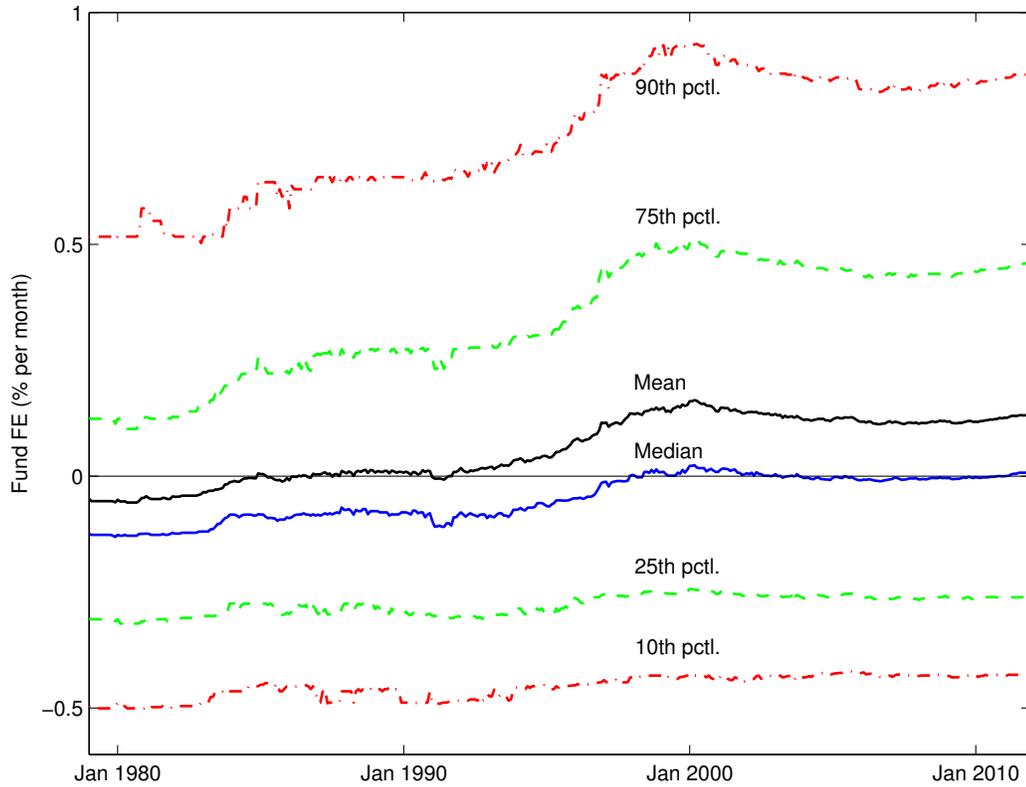


Figure 2. Distribution of fund skill over time. The figure plots each month’s mean and percentiles of estimated fund fixed effects across all funds operating during that month. The fixed effects are estimated by using specification (10) in Table 6.

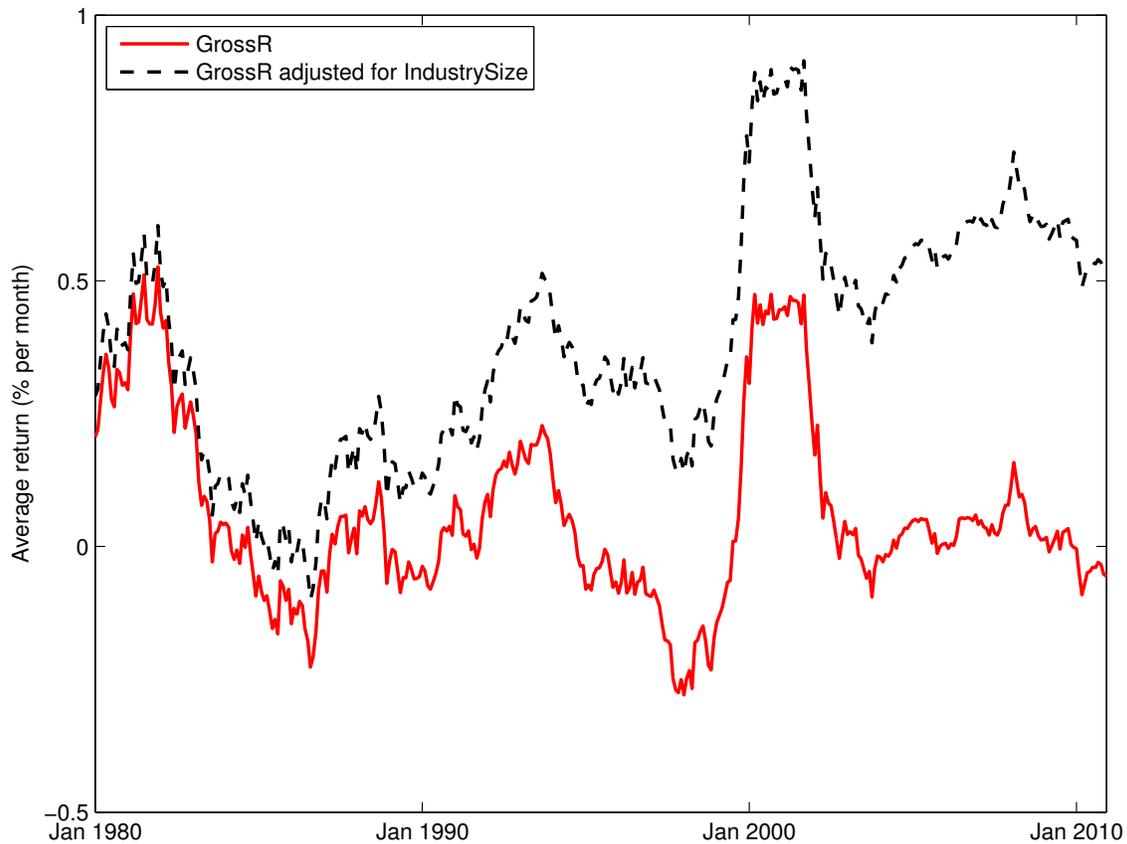


Figure 3. Average fund returns over time. The figure plots two-year moving averages of fund returns. The solid red line shows equal-weighted average benchmark-adjusted gross returns ($GrossR$). The dashed black line is the same as the red line except that it adjusts funds' $GrossR$ by adding 0.0326 times $IndustrySize$. The value 0.0326 is minus the estimated slope from an OLS FE regression of $GrossR$ on fund fixed effects and $IndustrySize$ (Table 3 Panel A column 5).

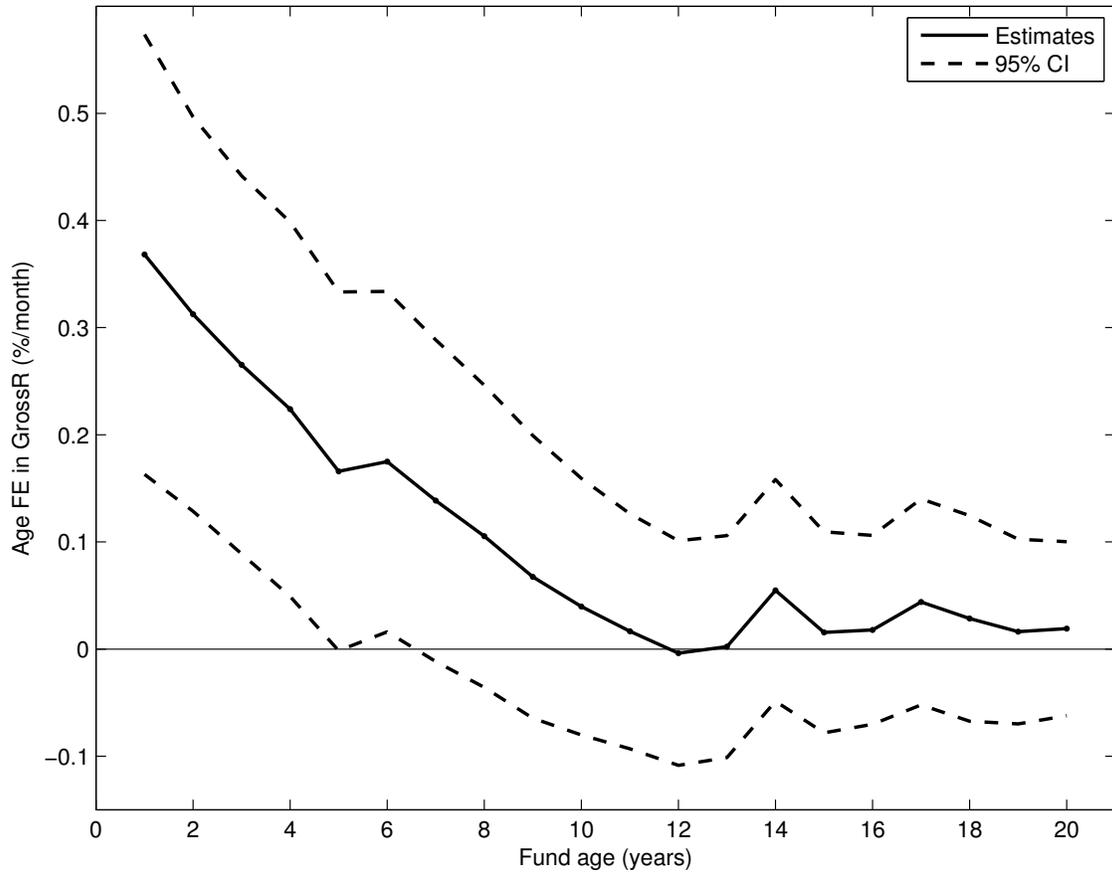


Figure 4. Fund age vs. performance. The figure plots fund age fixed effects in benchmark-adjusted gross returns ($GrossR$) along with their 95% confidence intervals. We obtain these fixed effects from a panel regression of $GrossR$ on fund fixed effects and dummy variables for fund age = 1, 2, ..., 20 years. The age- t fixed effect measures the average difference between $GrossR$ for a fund of age t and the same fund at age 21+ years. Standard errors are clustered by sector \times month. The sample covers March 1993 to December 2011.

Table 1
Simulation Exercise Comparing OLS, OLS FE, and RD Estimators

We simulate 10,000 samples of mutual fund gross returns (R) and size (q). Simulated returns follow $R_{it} = a_i + \beta q_{it-1} + \varepsilon_{it}$, and size follows $q_{it}/q_{it-1} - 1 = c + \gamma R_{it} + v_{it}$. In each sample, we estimate β using the OLS, OLS FE, and RD estimators. Panel A (B) shows means (medians) of the β estimates across the simulated samples. Panel C shows the fraction of samples in which we reject the null hypothesis $\beta = 0$.

Panel A: Mean estimated β ($\times 10^5$)									
β ($\times 10^5$)	OLS, no FE			OLS with FE			RD		
	$\gamma = 0.8$	$\gamma = 0.9$	$\gamma = 1.0$	$\gamma = 0.8$	$\gamma = 0.9$	$\gamma = 1.0$	$\gamma = 0.8$	$\gamma = 0.9$	$\gamma = 1.0$
0	0.85	0.84	0.82	-0.39	-0.38	-0.37	0.00	0.00	0.00
-1	0.30	0.37	0.42	-1.69	-1.73	-1.77	-0.97	-0.87	-1.03
-3	-0.96	-0.78	-0.62	-4.17	-4.28	-4.37	-2.34	-3.00	-3.08
-10	-7.51	-7.38	-7.27	-10.81	-10.82	-10.83	-10.03	-10.02	-10.02

Panel B: Median estimated β ($\times 10^5$)									
β ($\times 10^5$)	OLS, no FE			OLS with FE			RD		
	$\gamma = 0.8$	$\gamma = 0.9$	$\gamma = 1.0$	$\gamma = 0.8$	$\gamma = 0.9$	$\gamma = 1.0$	$\gamma = 0.8$	$\gamma = 0.9$	$\gamma = 1.0$
0	0.85	0.84	0.82	-0.39	-0.38	-0.37	0.00	0.00	0.00
-1	0.30	0.37	0.42	-1.68	-1.73	-1.76	-1.02	-1.01	-1.01
-3	-0.96	-0.78	-0.62	-4.17	-4.27	-4.37	-2.95	-2.95	-2.95
-10	-7.51	-7.38	-7.27	-10.81	-10.82	-10.83	-10.01	-10.01	-10.01

Panel C: Fraction reject the null hypothesis ($\beta = 0$)									
β ($\times 10^5$)	OLS, no FE			OLS with FE			RD		
	$\gamma = 0.8$	$\gamma = 0.9$	$\gamma = 1.0$	$\gamma = 0.8$	$\gamma = 0.9$	$\gamma = 1.0$	$\gamma = 0.8$	$\gamma = 0.9$	$\gamma = 1.0$
0	1.00	1.00	1.00	0.98	0.99	0.99	0.06	0.06	0.06
-1	0.77	0.90	0.96	1.00	1.00	1.00	0.13	0.14	0.16
-3	1.00	0.99	0.94	1.00	1.00	1.00	0.21	0.20	0.19
-10	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

Table 2
Summary Statistics

This table shows summary statistics for our sample of active equity mutual funds from 1979–2011. The unit of observation is the fund/month. All returns and expense ratios are in units of fraction per month. Net return is the return received by investors. Benchmark-adjusted net return equals net return minus the return on Morningstar’s chosen benchmark portfolio. *GrossR* is the benchmark-adjusted net return plus 1/12th of the annual expense ratio. *FundSize* is the fund’s total AUM aggregated across share classes, divided by the total stock market capitalization in the same month, then multiplied by the total stock market capitalization at the end of 2011. *IndustrySize* is the sum of all funds’ AUM divided by the total stock market capitalization in the same month, imputing *FundSize* when missing. *Turnover* is in units of fraction per year. $1(SmlCap)$ is an indicator for a small-cap fund, as defined by Morningstar’s Category variable. $Std(AbnRet)$ is the standard deviation of a fund’s abnormal returns, defined as residuals from a regression of excess gross fund returns on excess benchmark portfolio returns. *Fund age* is the number of years since the fund’s first offer date. *FamilySize* is the sum of *FundSize* across funds in the same family. *SectorSize* is the sum of AUM (in nominal dollars) across all funds within a given sector, divided by the total market value of all CRSP stocks belonging to that sector. The first version uses the nine sectors in Morningstar’s 3×3 StyleBox. The second version uses three sectors: small cap, mid cap, and large cap.

35

	# of fund/ months	Mean	Stdev.	Percentiles				
				1%	25%	50%	75%	99%
Net return	365,182	0.0069	0.0529	-0.1518	-0.0197	0.0103	0.0371	0.1320
Benchmark-adjusted net return	345,921	-0.0005	0.0234	-0.0642	-0.0106	-0.0009	0.0091	0.0680
Benchmark-adj. gross return (<i>GrossR</i>)	314,580	0.0005	0.0229	-0.0627	-0.0095	0.0001	0.0100	0.0680
Expense ratio (% per month)	334,767	0.0010	0.0004	0.0003	0.0008	0.0010	0.0012	0.0021
<i>FundSize</i> (in 2011 \$millions)	319,454	1,564	5,779	16	84	265	921	24,371
<i>IndustrySize</i>	381,192	0.1334	0.0441	0.0232	0.1227	0.1391	0.1714	0.1833
<i>Turnover</i>	328,642	0.8529	0.7351	0.0200	0.3499	0.6500	1.1200	4.0900
$1(SmlCap)$	340,061	0.1939	0.3954	0	0	0	0	1
$Std(AbnRet)$	363,338	0.0183	0.0084	0.0056	0.0125	0.0171	0.0222	0.0448
<i>Fund age</i> (years)	377,876	13.16	14.27	0.17	3.92	8.58	16.17	69.00
<i>FamilySize</i> (in 2011 \$millions)	315,894	61,821	149,667	23	1,768	12,181	46,217	825,507
<i>SectorSize</i> (v.1)	339,762	0.1561	0.0989	0.0220	0.0908	0.1290	0.2004	0.4788
<i>SectorSize</i> (v.2)	339,762	0.1390	0.0681	0.0210	0.1106	0.1318	0.1564	0.3550

Table 3
Relation Between Size and Fund Performance

The dependent variable in each regression model is *GrossR*, the fund's benchmark-adjusted gross return. *FundSize* is the fund's total AUM at the end of the previous month, inflated to millions of 2011 dollars using the total market cap of stocks in CRSP. *IndustrySize* is the total AUM of all active equity mutual funds divided by the total market cap of all stocks in CRSP. The OLS FE estimator includes fund fixed effects. The RD estimator recursively forward-demeans all variables and instruments for forward-demeaned *FundSize* using backward-demeaned *FundSize*. We multiply the slopes on *FundSize* by 10^6 to make them easier to read. The reported slopes on *FundSize* thus equal the change in *GrossR*, in units of bp per month, associated with a \$100 million increase in *FundSize*. Heteroskedasticity-robust *t*-statistics clustered by sector \times month are in parentheses. We also cluster by fund in the RD specifications.

Panel A: Main Sample (March 1993 – December 2011)									
<i>FundSize</i>	-0.0137 (-1.87)	-0.168 (-9.38)	-0.220 (-0.62)				-0.0147 (-2.02)	-0.148 (-9.09)	-0.425 (-1.25)
<i>IndustrySize</i>				-0.0169 (-1.93)	-0.0326 (-3.60)	-0.0326 (-2.49)	-0.0165 (-1.90)	-0.0295 (-3.27)	-0.0277 (-2.14)
Constant	0.000503 (2.18)			0.00304 (2.10)			0.00300 (2.09)		
Observations	275847	275847	270556	283046	283046	283046	275847	275847	270556
Estimator	OLS no FE	OLS FE	RD	OLS no FE	OLS FE	RD	OLS no FE	OLS FE	RD
Panel B: Extended Sample (January 1979 – December 2011)									
<i>FundSize</i>	-0.0139 (-1.98)	-0.126 (-9.09)	-0.107 (-0.64)				-0.0161 (-2.30)	-0.103 (-7.91)	-0.268 (-1.69)
<i>IndustrySize</i>				-0.00671 (-1.92)	-0.0164 (-4.34)	-0.0165 (-2.66)	-0.00943 (-1.86)	-0.0165 (-3.15)	-0.0179 (-2.09)
Constant	0.000506 (2.28)			0.00144 (2.58)			0.00189 (2.26)		
Observations	290150	290150	285350	314580	314580	314580	290150	290150	285350
Estimator	OLS no FE	OLS FE	RD	OLS no FE	OLS FE	RD	OLS no FE	OLS FE	RD

Table 4
Properties of the Relation Between Industry Size and Fund Performance

The dependent variable in all regressions is *GrossR*, the fund's benchmark-adjusted gross return. *Time Trend* is the number of months since January 1979. *IndustrySize* is the total AUM of all sample funds divided by the total market cap of all stocks in CRSP. *Average Fund Size* is the average AUM across all active equity mutual funds in a given month. We inflate *Average Fund Size* to current dollars by dividing by the total stock market capitalization in the same month, then multiplying by the total stock market capitalization at the end of 2011. *Number of Funds* is a count of the sample funds operating in the given month. Note that *IndustrySize* equals *Number of Funds* times *Average Fund Size* divided by the total stock market capitalization at the end of 2011. We multiply the slopes on *Time Trend*, *Average Fund Size*, and *Number of Funds* by 10^6 to make them easier to read. All models include fund fixed effects and are estimated by OLS. Heteroskedasticity-robust *t*-statistics clustered by sector \times month are in parentheses.

Panel A: Main Sample (March 1993 – December 2011)								
<i>IndustrySize</i>	-0.0326 (-3.60)		-0.0852 (-3.04)			-0.115 (-2.60)		-0.146 (-2.87)
<i>Time Trend</i>		-10.26 (-2.99)	23.89 (2.21)					15.35 (1.56)
<i>Average Fund Size</i>				-3.862 (-3.03)	-8.885 (-3.56)	4.315 (0.73)		4.249 (0.72)
<i>Number of Funds</i>					0.450 (0.83)	-4.031 (-3.23)	8.493 (1.61)	8.234 (1.57)
Observations	283046	283046	283046	283046	283046	283046	283046	283046
Panel B: Extended Sample (January 1979 – December 2011)								
<i>IndustrySize</i>	-0.0164 (-4.34)		-0.0305 (-1.72)			-0.0629 (-1.93)		-0.0697 (-1.87)
<i>Time Trend</i>		-7.030 (-4.00)	7.102 (0.84)					4.137 (0.51)
<i>Average Fund Size</i>				-2.532 (-2.44)	-8.898 (-3.83)	-1.455 (-0.30)		-1.608 (-0.34)
<i>Number of Funds</i>					0.0184 (0.05)	-3.926 (-4.26)	3.516 (0.84)	3.352 (0.81)
Observations	314580	314580	314580	314580	314580	314580	314580	314580

Table 5
Determinants of the Size-Performance Relation
Main Sample (March 1993 – December 2011)

This table adds additional regressors to the specifications in Table 3. $1(SmlCap)$ is an indicator for whether the fund is a small-cap fund. $Std(AbnRet)$ is the fund's standard deviation of residuals from a regression of excess gross returns (in fraction per month) on excess benchmark returns. $Turnover$ is in fraction per year. We instrument for all forward-demeaned quantities that involve $FundSize$ by using the backward-demeaned values.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
<i>FundSize</i>	-0.238 (-0.74)		-0.373 (-1.20)	-0.113 (-0.18)		0.0557 (0.10)	-0.180 (-0.60)		-0.305 (-1.02)	0.0796 (0.14)
<i>FundSize * 1(SmlCap)</i>	1.304 (0.42)		-1.178 (-0.40)							-0.612 (-0.21)
<i>FundSize * Std(AbnRet)</i>				-11.46 (-0.30)		-28.74 (-0.82)				-23.13 (-0.66)
<i>FundSize * Turnover</i>							-0.0599 (-0.31)		-0.115 (-0.59)	-0.147 (-1.23)
<i>Turnover</i>							0.000957 (3.63)	0.00583 (4.13)	0.00549 (3.46)	0.00443 (2.98)
<i>IndustrySize</i>		-0.0242 (-2.34)	-0.0198 (-1.30)		0.0545 (3.35)	0.0511 (2.41)		-0.00633 (-0.69)	-0.00369 (-0.30)	0.0596 (2.40)
<i>IndustrySize * 1(SmlCap)</i>		-0.0523 (-2.09)	-0.0501 (-1.56)							-0.0341 (-1.02)
<i>IndustrySize * Std(AbnRet)</i>					-4.519 (-4.80)	-4.205 (-3.31)				-3.405 (-2.72)
<i>IndustrySize * Turnover</i>								-0.0325 (-3.67)	-0.0311 (-3.05)	-0.0224 (-2.41)
Observations	246841	260411	246841	255622	276005	255622	268236	279749	268236	232368

Table 6
Determinants of the Size-Performance Relation
Extended Sample (January 1979 – December 2011)

This table is the same as Table 5 but uses data from January 1979 to December 2011.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
<i>FundSize</i>	-0.0987 (-0.66)		-0.196 (-1.48)	0.0225 (0.03)		0.267 (0.45)	-0.00325 (-0.04)		-0.0456 (-0.45)	0.451 (0.75)
<i>FundSize * 1(SmlCap)</i>	0.273 (0.13)		-1.424 (-0.70)							-1.356 (-0.68)
<i>FundSize * Std(AbnRet)</i>				-10.40 (-0.28)		-28.21 (-0.87)				-26.70 (-0.85)
<i>FundSize * Turnover</i>							-0.0536 (-0.98)		-0.111 (-2.05)	-0.109 (-2.91)
<i>Turnover</i>							0.000986 (4.01)	0.00402 (5.17)	0.00392 (3.87)	0.00334 (3.70)
<i>IndustrySize</i>		-0.0120 (-3.04)	-0.0122 (-1.28)		0.0247 (2.91)	0.0367 (2.21)		-0.00332 (-0.69)	-0.00287 (-0.35)	0.0400 (2.15)
<i>IndustrySize * 1(SmlCap)</i>		-0.0348 (-2.67)	-0.0446 (-1.80)							-0.0337 (-1.30)
<i>IndustrySize * Std(AbnRet)</i>					-2.129 (-4.51)	-2.885 (-2.96)				-2.181 (-2.22)
<i>IndustrySize * Turnover</i>								-0.0212 (-4.34)	-0.0207 (-3.22)	-0.0151 (-2.71)
Observations	261349	291869	261349	271629	307539	271629	282803	296188	282803	247871

Table 7
Relation Between Fund Age and Performance

This table shows results from estimating a panel model with dependent variable *GrossR*, the fund's benchmark-adjusted gross return. *Fund age* is the number of years since the fund's first offer date. *IndustrySize* is the total AUM of all active equity mutual funds divided by the total market cap of all stocks in CRSP. *FundSize* is the fund's total AUM at the end of the previous month, inflated to millions of 2011 dollars using the total market cap of stocks in CRSP. Specifications that do not include *FundSize* use the OLS FE estimator; those that do include *FundSize* use the RD estimator. The first three columns in each panel use all observations; the last three columns use only funds that are at least three years old. *t*-statistics clustered by sector \times month are in parentheses.

Panel A: Main Sample (March 1993 – December 2011)						
<i>Fund age</i>	-0.000123 (-3.00)	0.000283 (2.19)	0.000277 (2.10)	-0.000102 (-2.37)	0.000281 (2.19)	0.000267 (2.02)
<i>IndustrySize</i>		-0.0845 (-3.02)	-0.0794 (-2.77)		-0.0799 (-2.86)	-0.0746 (-2.60)
<i>FundSize</i>			-0.264 (-0.78)			-0.259 (-0.77)
Observations	283046	283046	270556	248050	248050	238426
Fund ages	All	All	All	≥ 3 years	≥ 3 years	≥ 3 years
Panel B: Extended Sample (January 1979 – December 2011)						
<i>Fund age</i>	-0.0000845 (-4.00)	0.0000832 (0.82)	0.000152 (1.28)	-0.0000729 (-3.43)	0.0000934 (0.93)	0.000153 (1.29)
<i>IndustrySize</i>		-0.0302 (-1.70)	-0.0440 (-2.11)		-0.0298 (-1.69)	-0.0417 (-2.01)
<i>FundSize</i>			-0.220 (-1.48)			-0.204 (-1.39)
Observations	314580	314580	285350	277381	277381	252222
Fund ages	All	All	All	≥ 3 years	≥ 3 years	≥ 3 years

Table 8
Investment Strategies Based on Fund Age

This table shows average returns of portfolios containing mutual funds of different ages. *Fund age* is the number of years since the fund's first offer date. At the beginning of each month, we assign mutual funds to portfolios based on their age at the end of the coming month. Columns 2–5 show the portfolios' average equal-weighted benchmark-adjusted gross (*GrossR*) and net (*NetR*) returns, in percent per month. The next three columns show the average difference in benchmark-adjusted returns between portfolios. The last column contains the *p*-value from an *F*-test of whether the four age-sorted portfolios have the same average benchmark-adjusted return, clustering by calendar date.

Panel A: Main sample (March 1993 – December 2011)								
<i>Fund age</i>	Average portfolio return				Average differences			<i>F</i> - test
	[0, 3]	(3, 6]	(6, 10]	>10	[0,3] - (>10)	(3,6] - (>10)	(6,10] - (>10)	<i>p</i> -value
Avg. <i>GrossR</i>	0.084 (2.33)	0.056 (1.45)	0.020 (0.55)	0.012 (0.30)	0.072 (2.85)	0.043 (2.48)	0.008 (0.52)	0.014
Avg. <i>NetR</i>	-0.005 (-0.15)	-0.052 (-1.38)	-0.084 (-2.29)	-0.083 (-2.07)	0.077 (3.10)	0.031 (1.79)	-0.001 (-0.08)	0.008

Panel B: Extended sample (January 1979 – December 2011)								
<i>Fund age</i>	Average portfolio return				Average differences			<i>F</i> - test
	[0, 3]	(3, 6]	(6, 10]	>10	[0,3] - (>10)	(3,6] - (>10)	(6,10] - (>10)	<i>p</i> -value
Avg. <i>GrossR</i>	0.106 (2.14)	0.120 (2.58)	0.105 (2.41)	0.033 (1.03)	0.075 (1.77)	0.096 (2.75)	0.072 (2.63)	0.003
Avg. <i>NetR</i>	0.007 (0.18)	-0.018 (-0.46)	-0.033 (-0.86)	-0.057 (-1.78)	0.064 (2.69)	0.048 (1.77)	0.024 (1.06)	0.032

Table 9
Family Size and the Size-Performance Relation

This table adds *FamilySize* to the specifications in Panel A of Table 3 (Main Sample). *FamilySize* is the sum of *FundSize* across funds belonging to the same family.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
<i>FamilySize</i>	0.000548 (1.33)	-0.00415 (-1.95)	-0.00489 (-1.41)	0.000367 (0.94)	-0.00306 (-2.11)	-0.00308 (-1.84)	0.000792 (1.99)	-0.000124 (-0.08)	-0.000516 (-0.19)
<i>FundSize</i>	-0.0207 (-3.01)	-0.147 (-7.92)	-0.360 (-1.23)				-0.0247 (-3.83)	-0.147 (-7.94)	-0.436 (-1.49)
<i>IndustrySize</i>				-0.0184 (-2.08)	-0.0305 (-3.40)	-0.0300 (-2.33)	-0.0179 (-2.04)	-0.0291 (-3.27)	-0.0271 (-2.10)
Constant	0.000501 (2.13)			0.00327 (2.24)			0.00321 (2.21)		
Observations	273569	273569	268311	277838	277838	277838	273569	273569	268311
Estimator	OLS no FE	OLS FE	RD	OLS no FE	OLS FE	RD	OLS no FE	OLS FE	RD

Table 10
Size-Performance Relation with Different Functional Forms

The dependent variable is *GrossR*, the fund's benchmark-adjusted gross return. All columns use the RD estimator and data from the main sample (March 1993 – December 2011). We divide *FundSize* by 10^6 before estimation, to make the slope coefficients easier to read. We instrument for all forward-demeaned variables that involve *FundSize* by using their backward-demeaned values.

	(1)	(2)	(3)	(4)	(5)	(6)
<i>Log(FundSize)</i>	0.00156 (0.38)	-0.00229 (-0.73)				
<i>FundSize</i> ²			-2.369 (-1.22)	-3.232 (-1.56)	2.767 (0.65)	5.205 (1.26)
<i>FundSize</i>					-0.475 (-1.14)	-0.767 (-1.91)
<i>IndustrySize</i>		-0.0258 (-1.95)		-0.0291 (-2.24)		-0.0298 (-2.28)
Observations	248999	248999	270556	270556	262396	262396

Table 11
Relation Between Sector Size and Fund Performance

The dependent variable is *GrossR*, the fund's benchmark-adjusted gross return. *SectorSize* (v1) is the total AUM of funds in the same sector, divided by the total market cap of CRSP stocks in that same sector; we use the nine sectors defined by Morningstar's size \times growth StyleBox. *SectorSize* (v2) is the same as *SectorSize* (v1) but uses three sectors: small, medium, and large-cap stocks. Remaining variables and estimators are defined in Table 3. Data are from the main sample (March 1993–December 2011).

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
<i>SectorSize</i> (v1)	0.00505 (1.34)	0.00579 (1.46)	0.00977 (2.28)	0.0101 (2.26)				
<i>SectorSize</i> (v2)					0.000349 (0.07)	0.00232 (0.37)	0.0185 (2.41)	0.0198 (2.28)
<i>FundSize</i>		-0.136 (-0.39)		-0.388 (-1.17)		-0.177 (-0.52)		-0.428 (-1.29)
<i>IndustrySize</i>			-0.0421 (-3.86)	-0.0379 (-2.54)			-0.0531 (-3.77)	-0.0500 (-2.73)
Observations	260411	249394	260411	249394	260411	249394	260411	249394

REFERENCES

- Angrist, Joshua D., and Jörn-Steffen Pischke, 2009, *Mostly Harmless Econometrics*, Princeton University Press, Princeton.
- Barras, Laurent, Olivier Scaillet, and Russ Wermers, 2010, False discoveries in mutual fund performance: Measuring luck in estimated alphas, *Journal of Finance* 65, 179-216.
- Berk, Jonathan B., and Jules H. van Binsbergen, 2012, Measuring managerial skill in the mutual fund industry, Working paper, Stanford University.
- Berk, Jonathan B., and Richard C. Green, 2004, Mutual fund flows and performance in rational markets, *Journal of Political Economy* 112, 1269-1295.
- Bris, Arturo, Huseyin Gulen, Padma Kadiyala, and P. Raghavendra Rau, 2007, Good stewards, cheap talkers, or family men? The impact of mutual fund closures on fund managers, flows, fees, and performance, *Review of Financial Studies* 20, 953-982.
- Carhart, Mark M., 1997, On persistence in mutual fund performance, *Journal of Finance* 52, 57-82.
- Chen, Joseph, Harrison Hong, Ming Huang, and Jeffrey Kubik, 2004, Does fund size erode mutual fund performance? *American Economic Review*, 94: 1276-1302.
- Chevalier, Judith and Glenn Ellison, 1997, Risk taking by mutual funds as a response to incentives, *Journal of Political Economy* 105: 1167-1200.
- Chevalier, Judith and Glenn Ellison, 1999, Career concerns of mutual fund managers, *Quarterly Journal of Economics* 114: 389-432.
- Cohen, Randolph, Joshua Coval, and Ľuboš Pástor, 2005, Judging fund managers by the company they keep, *Journal of Finance* 60, 1057-1096.
- Coval, Joshua, and Erik Stafford, 2007, Asset fire sales (and purchases) in equity markets, *Journal of Financial Economics* 86, 479-512.
- Cremers, Martijn, Antti Petajisto, and Eric Zitzewitz, 2013, Should benchmark indices have alpha? Revisiting performance evaluation, *Critical Finance Review*, forthcoming.
- Daniel, Kent D., Mark Grinblatt, Sheridan Titman, and Russ Wermers, 1997, Measuring mutual fund performance with characteristic-based benchmarks, *Journal of Finance* 52, 1035-1058.
- Edelen, Roger M. and Jerold B. Warner, 2001, Aggregate price effects of institutional trading: A study of mutual fund flow and market returns, *Journal of Financial Economics*, 59: 195-220.
- Edelen, Roger M., Richard Evans, and Gregory B. Kadlec, 2007, Scale effects in mutual fund performance: The role of trading costs, Working paper, University of California at Davis.
- Elton, Edwin, Martin Gruber, and Christopher Blake, 2001, A first look at the accuracy of the CRSP mutual fund database and a comparison of the CRSP and Morningstar mutual fund databases, *Journal of Finance* 56, 2415-2430.
- Elton, Edwin, Martin Gruber, and Christopher Blake, 2012, Does mutual fund size matter? The relationship between size and performance, *Review of Asset Pricing Studies* 2, 31-55.
- Evans, Richard, 2010, Mutual fund incubation, *Journal of Finance* 65, 1581-1611.

- Fama, Eugene F., and Kenneth R. French, 2010, Luck versus skill in the cross section of mutual fund returns, *Journal of Finance* 65, 1915–1947.
- Ferson, Wayne E., and Rudi W. Schadt, 1996, Measuring fund strategy and performance in changing economic conditions, *Journal of Finance* 51, 425–461.
- Ferreira, Miguel A., Aneel Keswani, António F. Miguel, and Sofia Ramos, 2013, The determinants of mutual fund performance: A cross-country study, *Review of Finance* 17, 483–525.
- Ferreira, Miguel A., Aneel Keswani, António F. Miguel, and Sofia Ramos, 2013b, Testing the Berk and Green model around the world, Working paper, Nova School of Business and Economics.
- Fung, William, David A. Hsieh, Narayan Y. Naik, and Tarun Ramadorai, 2008, Hedge funds: Performance, risk, and capital formation, *Journal of Finance* 63, 1777–1803.
- Grinblatt, Mark, and Sheridan Titman, 1989, Mutual fund performance: An analysis of quarterly portfolio holdings, *Journal of Business* 62, 393–416.
- Hjalmarsson, Erik, 2010, Predicting global stock returns, *Journal of Financial and Quantitative Analysis* 45, 49–80.
- Jensen, Michael C., 1968, The performance of mutual funds in the period 1945–1964, *Journal of Finance* 23, 389–416.
- Juhl, Ted, and Oleksandr Lugovskyy, 2010, A test for slope heterogeneity in fixed effects models, Working paper, University of Kansas.
- Kacperczyk, Marcin, Clemens Sialm, and Lu Zheng, 2005, On the industry concentration of actively managed equity mutual funds, *Journal of Finance* 60, 1983–2011.
- Kacperczyk, Marcin, Clemens Sialm, and Lu Zheng, 2008, Unobserved actions of mutual funds, *Review of Financial Studies* 21, 2379–2416.
- Kaplan, Steven N. and Antoinette Schoar, 2005, Private equity performance: Returns, persistence, and capital flows, *Journal of Finance* 55, 1791–1823.
- Khan, Mozafer, Leonid Kogan, and George Serafeim, 2012, Mutual fund trading pressure: Firm-level stock price impact and timing of SEOs, *Journal of Finance* 67, 1371–1395.
- Kosowski, Robert, Allan Timmermann, Russ Wermers, and Hal White, 2006, Can mutual fund “stars” really pick stocks? New evidence from a bootstrap analysis, *Journal of Finance* 61, 2551–2595.
- Lou, Dong, 2012, A flow-based explanation for return predictability, *Review of Financial Studies* 25, 3457–3489.
- Moon, Hyungsik R., and Peter C.B. Phillips, 2000, Estimation of autoregressive roots near unity using panel data, *Econometric Theory* 16, 927–997.
- Pástor, Ľuboš, and Robert F. Stambaugh, 2002, Mutual fund performance and seemingly unrelated assets, *Journal of Financial Economics* 63, 315–349.
- Pástor, Ľuboš, and Robert F. Stambaugh, 2012, On the size of the active management industry, *Journal of Political Economy* 120, 740–781.
- Perold, André F., and Robert S. Salomon, 1991, The right amount of assets under management, *Financial Analysts Journal* 47, 31–39.

- Pollet, Joshua, and Mungo Wilson, 2008, How does size affect mutual fund behavior? *Journal of Finance* 63, 2941–2969.
- Reuter, Jonathan, and Eric Zitzewitz, 2013, How much does size erode mutual fund performance? A regression discontinuity approach, Working paper, Boston College.
- Roberts, Michael R., and Toni M. Whited, 2012, Endogeneity in corporate finance, forthcoming in George Constantinides, Milton Harris, and René Stulz, eds. *Handbook of the Economics of Finance* Volume 2, Elsevier.
- Sensoy, Berk, 2009, Performance evaluation and self-designated benchmark indexes in the mutual fund industry, *Journal of Financial Economics* 92, 25–39.
- Sirri, Erik R. and Peter Tufano, 1998, Costly search and mutual fund flows, *Journal of Finance* 53: 1589–1622.
- Stock, James H., Jonathan H. Wright, and Motohiro Yogo, 2002, A survey of weak instruments and weak identification in generalized method of moments, *Journal of Business and Economic Statistics* 20, 518–529.
- Stambaugh, Robert F., 1999, Predictive regressions, *Journal of Financial Economics* 54, 375–421.
- Wermers, Russ, 2000, Mutual fund performance: An empirical decomposition into stock-picking talent, style, transactions costs, and expenses, *Journal of Finance* 55, 1655–1695.
- Wermers, Russ, 2003, Is money really smart? New evidence on the relation between mutual fund flows, manager behavior, and performance persistence, Working Paper, University of Maryland.
- Yan, Xuemin, 2008, Liquidity, investment style, and the relation between fund size and fund performance, *Journal of Financial and Quantitative Analysis* 43, 741–768.