Internet Appendix for "Skill, Scale, and Value Creation in the Mutual Fund Industry"*

*Citation Format: Barras, Laurent, Patrick Gagliardini, and Olivier Scaillet, 2021, Internet Appendix for "Skill, Scale, and Value Creation in the Mutual Fund Industry", Journal of Finance. Please note: Wiley is not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the authors of the article. Barras is at McGill University (Desautels Faculty of Management), Gagliardini is at the Università della Svizzera italiana (USI Lugano) and the Swiss Finance Institute (SFI), and Scaillet is at the University of Geneva (Geneva Finance Research Institute (GFRI)) and the SFI.

This appendix is divided in six sections. Section I provides a description of the methodology for estimating the kernel density and the distribution characteristics (moments, proportion, and quantile). It also contains the proofs of the propositions discussed in the paper. Section II examines the asymptotic properties of the estimated distribution characteristics under an alternative analytical approach (as opposed to the numerical approach used in the paper). Section III provides a detailed discussion of the Error-in-Variable (EIV) bias. Section IV describes our extensive Monte-Carlo analysis. It also presents simulation results under the assumption that skill and scalability are uncorrelated. Section V describes the construction of the data set and different fund groups. Finally, Section VI explains the construction of our new formal specification test. It also reports additional empirical results on (i) the validity of the panel specification, (ii) the impact of survivorship and reverse survivorship bias, (iii) the use of alternative asset pricing models, (iv) the analysis based on daily fund returns, and (v) the introduction of variables that capture changes in the economic conditions.

I Methodology

A Estimation Procedure

To begin the presentation of the methodology, we explain how to estimate the measure m_i for each fund, where $m_i \in \{a_i, b_i, va_i, va_i(s)\}$. The estimation procedure explicitly controls for the small-sample bias in the time-series regression $r_{i,t} = a_i - b_i q_{i,t-1} + \beta_i' f_t + \varepsilon_{i,t}$. This bias, which disappears asymptotically, arises because the mutual fund error term $\varepsilon_{i,t}$ is positively correlated with the innovation in size $\varepsilon_{q_i,t}$, i.e., $\varepsilon_{i,t} = \psi_i \varepsilon_{q_i,t} + \upsilon_{i,t}$, where ψ_i is positive. Specifically, $\varepsilon_{q_i,t}$ denotes the size innovation projected onto the space spanned by the factors f_t: $\varepsilon_{q_i,t} = e_{q_i,t} - \beta_{q_i}' x_t$, where $x_t = (1, f_t')'$ and $e_{q_i,t}$ is the innovation of the size regression $q_{i,t} = \theta_{q_i} + \rho_{q_i} q_{i,t-1} + e_{q_i,t}$. Failing to adjust for the small-sample bias produces values for the skill and scale coefficients that are too high.[1]

As noted by Amihud and Hurvich (2004), adding the regressor $\varepsilon_{q_i,t}$ eliminates the small-sample bias. To see this point, we can replace $\varepsilon_{i,t}$ with $\psi_i \varepsilon_{q_i,t} + \upsilon_{i,t}$ to obtain

$$r_{i,t} = a_i - b_i q_{i,t-1} + \beta_i' f_t + \psi_i \varepsilon_{q_i,t} + \upsilon_{i,t}, \tag{A1}$$

and verify that strict exogeneity holds, i.e., $E[\upsilon_i | X_i] = 0$, where $\upsilon_i = (\upsilon_{i,1}, ..., \upsilon_{i,T})'$, X_i

[1] Using the analysis of Stambaugh (1999), we have $E[\hat{b}_i^{wb} - b_i] = -E[\hat{\rho}_{q_i} - \rho_{q_i}]\psi_i > 0$ and $E[\hat{a}_i^{wb} - a_i] = E[\hat{b}_i^{wb} - b_i]E[q_{i,t-1}] > 0$, where $\hat{a}_i^{wb}$ and $\hat{b}_i^{wb}$ denote the estimators of a_i and b_i without the small-sample bias correction.

is the $T_i \times (K_f + 3)$ matrix of the available observations of $x_{i,t} = (1, -q_{i,t-1}, f_t', \varepsilon_{q_i,t})'$, and K_f is the number of factors. Of course, we do not observe the true projected innovation $\varepsilon_{q_i,t}$. Therefore, we use the procedure proposed by Amihud and Hurvich (2004) and Avramov, Barras, and Kosowski (2013) to compute a proxy for $\varepsilon_{q_i,t}$ denoted by $\varepsilon^c_{q_i,t}$.

This four-step procedure is applied to each fund i individually ($i = 1, .., n$). First, we run the size regression to obtain the estimated coefficients $\hat{\theta}_{q_i}$ and $\hat{\rho}_{q_i}$. Second, we compute the adjusted size innovation as

$$e^c_{q_i,t} = q_{i,t} - (\hat{\theta}^c_{q_i} + \hat{\rho}^c_{q_i} q_{i,t-1}), \tag{A2}$$

where the second-order coefficients corrected for the small-sample bias are given by $\hat{\rho}^c_{q_i} = \min(\hat{\rho}_{q_i} + (1 + 3\hat{\rho}_{q_i})/T_i + 3(1 + 3\hat{\rho}^2_{q_i})/T_i^2, 0.999)$ and $\hat{\theta}^c_{q_i} = (1 - \hat{\rho}_{q_i})\overline{q}_i$. Third, we regress $e^c_{qi,t}$ on the factors to obtain $\varepsilon^c_{q_i,t} = e^c_{q_i,t} - \hat{\beta}'_{q_i} x_t$. Finally, we insert $\varepsilon^c_{q_i,t}$ in Equation (A1) to obtain

$$r_{i,t} = a_i - b_i q_{i,t-1} + \beta_i' f_t + \psi_i \varepsilon^c_{q_i,t} + \upsilon_{i,t}. \tag{A3}$$

From this regression, we can obtain estimated values for m_i that are adjusted for the small-sample bias.

B Asymptotic Properties of the Kernel Density

Proof of Proposition III.1 (Asymptotic properties). In this section, we provide a proof of the asymptotic properties of the kernel density $\hat{\phi}(m)$ for each measure m_i. To this end, we initially focus on the skill coefficient, i.e., $m_i = a_i$. We allow for weak serial dependence in the error terms (i.e., temporal mixing). To simplify the presentation and avoid unnecessary technicalities related to spatial mixing conditions, we assume that the error terms are cross-sectionally independent. To further ease the presentation, we do not explicitly include the small-sample bias correction of the previous section because it has no impact on the asymptotic analysis when T becomes large.[2]

[2] The inclusion of the estimated variable $\varepsilon^c_{q_i,t}$ in the set of regressors does not change the asymptotic properties of the nonparametric density kernel estimator because the estimation error in $\varepsilon^c_{q_i,t}$ only affects the higher order terms beyond T^{-1}.

From the OLS estimation of the regression $r_{i,t} = a_i - b_i q_{i,t-1} + \beta_i' f_t + \varepsilon_{i,t}$, we have:

$$
\begin{aligned}
\hat{m}_i &= e_1' \hat{Q}_{x,i}^{-1} \frac{1}{T_i} \sum_t I_{i,t} x_{i,t} r_{i,t} = m_i + e_1' \hat{Q}_{x,i}^{-1} \frac{1}{T_i} \sum_t I_{i,t} x_{i,t} \varepsilon_{i,t} \\
&= m_i + \frac{1}{\sqrt{T}} \tau_{i,T} e_1' \hat{Q}_{x,i}^{-1} \left(\frac{1}{\sqrt{T}} \sum_t I_{i,t} x_{i,t} \varepsilon_{i,t} \right) \equiv m_i + \frac{1}{\sqrt{T}} \hat{\eta}_{i,T}, \qquad \text{(A4)}
\end{aligned}
$$

where $x_{i,t} = (1, -q_{i,t-1}, f_t')'$. Moreover, let us write

$$
\hat{\eta}_{i,T} = \eta_{i,T} + \frac{1}{\sqrt{T}} \hat{v}_{i,T}, \qquad \text{(A5)}
$$

where $\eta_{i,T} = \tau_i \frac{1}{\sqrt{T}} \sum_t I_{i,t} u_{i,t}$, $u_{i,t} = e_1' Q_{x,i}^{-1} x_{i,t} \varepsilon_{i,t}$, $\hat{v}_{i,T} = (\tau_{i,T} - \tau_i) \sum_t I_{i,t} e_1' \hat{Q}_{x,i}^{-1} x_{i,t} \varepsilon_{i,t} + \tau_i \sum_t I_{i,t} e_1' (\hat{Q}_{x,i}^{-1} - Q_{x,i}^{-1}) x_{i,t} \varepsilon_{i,t}$, $\tau_i = \text{plim}_{T\to\infty} \tau_{i,T}$, and $\tau_{i,T} = T/T_i$. The term $\hat{\eta}_{i,T}/\sqrt{T}$ corresponds to the estimation error of $\hat{m}_i$. It is equal to the sum of $\eta_{i,T}/\sqrt{T}$ and $\hat{v}_{i,T}/T$, where the second component captures the errors due to estimating the matrix $Q_{x,i}$ and the random sample size T_i. We can write $\hat{\phi}(m) - \phi(m) = I_1 + I_2 + I_3 + I_4$, where:

$$
\begin{aligned}
I_1 &= \frac{1}{h} E\left[K\left(\frac{m_i - m}{h} \right) \right] - \phi(m), \\
I_2 &= \frac{1}{h} E\left[K\left(\frac{m_i + \eta_{i,T}/\sqrt{T} - m}{h} \right) \right] - \frac{1}{h} E\left[K\left(\frac{m_i - m}{h} \right) \right], \\
I_3 &= \frac{1}{nh} \sum_i \left\{ K\left(\frac{m_i + \eta_{i,T}/\sqrt{T} - m}{h} \right) - E\left[K\left(\frac{m_i + \eta_{i,T}/\sqrt{T} - m}{h} \right) \right] \right\}, \\
I_4 &= \frac{1}{nh} \sum_i \left[\mathbf{1}_i^{\chi} K\left(\frac{m_i + \eta_{i,T}/\sqrt{T} + \hat{v}_{i,T}/T - m}{h} \right) \right] \\
&\quad - \frac{1}{nh} \sum_i \left[K\left(\frac{m_i + \eta_{i,T}/\sqrt{T} - m}{h} \right) \right]. \qquad \text{(A6)}
\end{aligned}
$$

The first term I_1 is the smoothing bias, the second term I_2 is the Error-In-Variable (EIV) bias, and I_3 is the main stochastic term. The remainder term I_4 is associated with $\hat{v}_{i,T}/T$ and is negligible with respect to the others. We now characterize the first three dominating terms.

(i) From standard results in kernel density estimation, the smoothing bias is such that $I_1 = \frac{1}{2} \phi^{(2)}(m) K_2 h^2 + O(h^3)$, with $K_2 = \int u^2 K(u) du$.

(ii) By a Taylor expansion of the kernel function K we have

$$I_2 = \sum_{j=1}^{\infty} \frac{1}{j!T^{j/2}h^{j+1}} E\left[K^{(j)}\left(\frac{m_i - m}{h}\right)(\eta_{i,T})^j\right]. \tag{A7}$$

We can then apply j times partial integration and a change of variable to obtain

$$\begin{aligned}\frac{1}{h^{j+1}} E\left[K^{(j)}\left(\frac{m_i - m}{h}\right)(\eta_{i,T})^j\right] &= \frac{1}{h^{j+1}} \int K^{(j)}\left(\frac{u-m}{h}\right)\psi_{T,j}(u)du \\ &= (-1)^j \frac{1}{h} \int K\left(\frac{u-m}{h}\right)\psi_{T,j}^{(j)}(u)du \\ &= (-1)^j \int K(u)\,\psi_{T,j}^{(j)}(m+hu)du, \end{aligned} \tag{A8}$$

where $\psi_{T,j}(m) = E[(\eta_{i,T})^j | m_i = m]\phi(m)$ for $j = 1, 2, \ldots$. We have $\psi_{T,1}(m) = 0$ and $\lim_{T\to\infty}\psi_{T,2}(m) = E[S_i | m_i = m]\phi(m) \equiv \psi(m)$ where S_i is equal to $\tau_i^2 \text{plim}_{T\to\infty} \frac{1}{T}\sum_{t,s} I_{i,t}I_{i,s}u_{i,t}u_{i,s}$. By weak serial dependence of the error terms, functions $\psi_{T,j}(m)$ for $j > 2$ are bounded with respect to T. Thus, we get: $I_2 = \frac{1}{2T}\psi^{(2)}(m) + O(1/T^{3/2} + h^2/T)$.

(iii) Let us now consider term I_3. For expository purpose, we treat the factor values f_t as given constants. Then:

$$V[I_3] = \frac{1}{nh^2} V\left[K\left(\frac{m_i + \eta_{i,T}/\sqrt{T} - m}{h}\right)\right]. \tag{A9}$$

From the above arguments, we have $\frac{1}{h}E\left[K\left(\frac{m_i+\eta_{i,T}/\sqrt{T}-m}{h}\right)\right] = \phi(m) + o(1)$ and

$$\begin{aligned}\frac{1}{h}E\left[K\left(\frac{m_i + \eta_{i,T}/\sqrt{T} - m}{h}\right)^2\right] &= \int K(u)^2 du \frac{1}{h} E\left[\bar{K}\left(\frac{m_i + \eta_{i,T}/\sqrt{T} - m}{h}\right)\right] \\ &= \phi(m)\int K(u)^2 du + o(1), \end{aligned} \tag{A10}$$

where $\bar{K}(u) = K(u)^2 / \int K(u)^2 du$. Therefore:

$$V[I_3] = \frac{1}{nh}\phi(m)\int K(u)^2 du + o\left(\frac{1}{nh}\right). \tag{A11}$$

Under regularity conditions, we can apply an appropriate central limit theorem

(CLT) to obtain $\sqrt{nh}I_3 \Rightarrow N(0, \phi(m)K_1)$, where $K_1 = \int K(u)^2 du$. Grouping the different elements completes the proof.[3] QED

We can apply the same arguments for all the other measures used in the paper: (i) the scale coefficient ($m_i = b_i$), (ii) the value added ($m_i = va_i$), and (iii) the subperiod value added ($m_i = va_i(s)$). The only required change is to use the appropriate definition for $\hat{m}_i$ and $u_{i,t}$ given in the paper.

C Optimal Bandwidth

Proof of Proposition III.1 (Optimal bandwidth) We now prove the remaining part of Proposition III.1 by solving for the optimal bandwidth h^* that minimizes the Asymptotic Mean Integrated Squared Error (AMISE) of the density $\hat{\phi}(m)$. From the arguments in Section B above, we get the asymptotic expansion of the bias $bs(m)$ of the estimator $\hat{\phi}(m)$ with leading terms,

$$
\begin{aligned}
bs_1(m) &= \frac{1}{2}h^2 K_2 \phi^{(2)}(m), &\text{(A12)}\\
bs_2(m) &= \frac{1}{2T}\psi^{(2)}(m). &\text{(A13)}
\end{aligned}
$$

where $bs_1(m)$ denotes the smoothing bias and $bs_2(m)$ denotes the EIV bias.[4] We also get the asymptotic expansion of the variance of the estimator $\hat{\phi}(m)$ with leading terms $\sigma^2(m) = \frac{1}{nh}\phi(m)K_1$. Combining these elements, we can write the AMISE as

$$
\begin{aligned}
AMISE(h) &= \int [\sigma^2(u) + bs(u)^2]du = \int [\sigma^2(u) + (bs_1(u) + bs_2(u))^2]du \\
&= \frac{1}{nh}K_1 + \frac{h^4 K_2^2}{4}\int [\phi^{(2)}(u)]^2 du \\
&\quad + \frac{h^2 K_2}{2T}\int \phi^{(2)}(u)\psi^{(2)}(u)du + \frac{1}{4T^2}\int \left[\psi^{(2)}(u)\right]^2 du, \qquad \text{(A14)}
\end{aligned}
$$

[3] Okui and Yanagi (2020) also consider a kernel estimator for the density of the mean and autocorrelation of random variables. However, their distributional results differ from our regression-based results aimed at measuring fund skill.

[4] From Equations (A12) and (A13), the integral of $bs_1^r(m)m$ and $bs_2^r(m)m$ is equal to zero if $\phi^{(1)}$ and $\psi^{(1)}$ vanish at the boundary of the support (which is the case in our Gaussian reference model). Hence it implies that the bias adjusted density (Equation (19) in the paper) integrates to one by construction.

where we assume that $\int \phi^{(2)}(u)\psi^{(2)}(u)du \geqslant 0$ so that the AMISE is convex. The optimal bandwidth h^* minimizes the AMISE and solves the equation:

$$\begin{aligned} & -\frac{1}{nh^2} + c_1 h^3 + c_2 \frac{h}{T} = 0 \\ \iff \quad & 1 = c_1 n h^5 + c_2 \frac{nh^3}{T}, \end{aligned} \tag{A15}$$

where $c_1 = K_2^2 \int [\phi^{(2)}(m)]^2 dm / K_1$ and $c_2 = \frac{K_2}{K_1} \int \phi^{(2)}(m)\psi^{(2)}(m)dm$ (with c_1, $c_2 > 0$).

The analytical approximation of the optimal bandwidth h^* depends on the relative increase of n and T. If (i) nh^3/T tends to a nonzero constant and (ii) nh^5 tends to zero, Equation (A15) implies that asymptotically

$$h^* = c_2^{-\frac{1}{3}} \left(\frac{n}{T}\right)^{-\frac{1}{3}} = \left(\frac{K_2}{K_1} \int \phi^{(2)}(m)\psi^{(2)}(m)dm\right)^{-\frac{1}{3}} \left(\frac{n}{T}\right)^{-\frac{1}{3}}. \tag{A16}$$

This solution is admissible (i.e., it satisfies $nh^5 \to 0$) if the sample sizes n and T are such that $n^{2/5}/T \to \infty$ or, put differently, if T is small relative to n. QED

We now consider the asymptotic distribution of the kernel density obtained with the optimal bandwidth h^*. We can check that $\sqrt{nh^*}(h^{*3} + h^{*2}/T + 1/T^{3/2}) = o(1)$ if $n/T^4 \to 0$. Replacing $bs(m)$ with its asymptotic approximation we have:

$$\sqrt{nh^*}\left(\hat{\phi}(m) - \phi(m) - \frac{1}{2}\phi^{(2)}(m)K_2 h^{*2} - \frac{1}{2T}\psi^{(2)}(m)\right) \Rightarrow N\left(0, \phi(m)K_1\right), \tag{A17}$$

where the smoothing bias is negligible and the dominant component is the EIV bias of order $O(1/T)$ (because we have $n^{2/5}/T \to \infty$ and $Th^{*2} \to 0$).

Note that if (i) nh^3/T tends to zero and (ii) nh^5 tends to a nonzero constant, Equation (A15) produces a different optimal bandwidth of the form $h^* \sim c_1^{-\frac{1}{5}} n^{-\frac{1}{5}}$ (i.e., the usual Silverman rule). This solution is admissible (i.e., it satisfies $nh^3/T \to 0$) if the sample sizes n and T are such that $n^{2/5}/T \to 0$ or, put differently, if T is large relative to n.[5]

Our Monte-Carlo analysis in Section IV reveals that given our actual sample size, the optimal bandwidth in Equation (A16) produces the best results. Motivated by these results, we therefore use it in our baseline specification. We also verify that the empirical results are remarkably similar under the two bandwidth choices.

[5] In the special case where $n^{2/5}/T \to \rho$, with $\rho > 0$, the two rates of convergence $n^{-1/5}$ and $(n/T)^{-1/3}$ coincide. Then, Equation (A15) has a solution such that $h^* \sim \bar{c}^{1/5} n^{-1/5}$, where $\bar{c}$ solves the equation $1 = c_1 \bar{c} + c_2 \rho \bar{c}^{3/5}$. Therefore, the optimal bandwidth remains proportional to $n^{-1/5}$ (similar to the Silverman rule).

D Adjustment of the Density Bias

Proof of Proposition III.2. We now prove the second proposition of the paper which provides closed form expressions for the two bias components $bs_1(m)$ and $bs_2(m)$ and the optimal bandwidth h^*. We use a Gaussian reference model in which m_i and $s_i = \log(S_i)$ follow a bivariate Gaussian distribution with mean parameters μ_m, μ_s, variance parameters σ_m^2, σ_s^2, and correlation parameter ρ.[6] We also use a standard Gaussian kernel $K(u) = \frac{1}{\sqrt{2\pi}} \exp\left(-u^2/2\right)$ with $K_1 = \int K(u)^2 du = \frac{1}{2\sqrt{\pi}}$ and $K_2 = \int u^2 K(u) du = 1$. The constants c_1 and c_2 are given by:

$$c_1 = 2\sqrt{\pi} \int [\phi^{(2)}(u)]^2 du, \tag{A18}$$

$$c_2 = 2\sqrt{\pi} \int \phi^{(2)}(u)\psi^{(2)}(u) du = 2\sqrt{\pi} \int \phi^{(4)}(u)\psi(u) du, \tag{A19}$$

where we use twice partial integration for c_2.

Let us now compute the two integrals appearing in these formulas. We have $\phi(m) = \frac{1}{\sigma_m}\varphi\left(\frac{m-\mu_m}{\sigma_m}\right)$ where $\varphi(z) = \frac{1}{\sqrt{2\pi}} \exp\left(-z^2/2\right)$ is the standard Gaussian density. We have

$$\phi^{(1)}(m) = -\frac{1}{\sigma_m}\left(\frac{m-\mu_m}{\sigma_m}\right)\phi(m), \tag{A20}$$

$$\phi^{(2)}(m) = \frac{1}{\sigma_m^2}\left(\left(\frac{m-\mu_m}{\sigma_m}\right)^2 - 1\right)\phi(m). \tag{A21}$$

Therefore, the first integral is equal to

$$\begin{aligned} \int [\phi^{(2)}(u)]^2 du &= \frac{1}{\sigma_m^5}\int (z^2-1)^2 \frac{1}{2\pi}\exp(-z^2) dz = \frac{1}{2\sqrt{\pi}\sigma_m^5}\int (v^2/2-1)^2\varphi(v) dv \\ &= \frac{3}{8\sqrt{\pi}\sigma_m^5}, \end{aligned} \tag{A22}$$

with the changes of variables from u to $z = (u-\mu_m)/\sigma_m$, and from z to $v = \sqrt{2}z$.

We can write the second integral as

$$\begin{aligned} \int \phi^{(4)}(m)\psi(m) dm &= \frac{\exp\left(\mu_s + \frac{1}{2}\sigma_s^2(1-\rho^2)\right)}{\sigma_m^5}\int \varphi^{(4)}(z)\exp(\rho\sigma_s z)\varphi(z) dz \\ &= \frac{\exp\left(\mu_s + \frac{1}{2}\sigma_s^2(1-\rho^2)\right)}{2\sqrt{\pi}\sigma_m^5}\int (v^4/4 - 3v^2 + 3)\exp(\lambda v)\varphi(v) dv, \end{aligned} \tag{A23}$$

[6] The Gaussian marginal density of m_i implies that our reference model nests the standard model underlying the derivation of the Silverman rule for kernel smoothing.

where $\psi(m) = E[\exp(s_i)|m_i = m]\phi(m) = \exp\left(\mu_s + \rho\sigma_s\left(\frac{m-\mu_m}{\sigma_m}\right) + \frac{1}{2}\sigma_s^2(1-\rho^2)\right)\phi(m)$, $\lambda = \rho\sigma_s/\sqrt{2}$ by using the same changes of variables as above, and $\varphi^{(4)}(z) = (z^4 - 6z^2 + 3)\varphi(z)$. To compute the integral in Equation (A23), we can exploit the following equality that applies to a standard Gaussian random variable Z: $\int z^k \exp(\lambda z)\varphi(z)dz = E[Z^k \exp(\lambda Z)] = \frac{\partial^k}{\partial\lambda^k}E[\exp(\lambda Z)]$ with $E[\exp(\lambda Z)] = \exp(\lambda^2/2)$. This yields $\int(v^4/4 - 3v^2 + 3)\exp(\lambda v)\varphi(v)dv = \left(\frac{1}{4}\frac{\partial^4}{\partial\lambda^4} - 3\frac{\partial^2}{\partial\lambda^2} + 3\right)\exp(\lambda^2/2) = \frac{1}{4}(\lambda^4 - 6\lambda^2 + 3)\exp(\lambda^2/2)$. Therefore, we obtain

$$\int \phi^{(4)}(m)\psi(m)dm = \frac{3\exp\left(\mu_s + \frac{1}{2}\sigma_s^2\left(1 - \frac{\rho^2}{2}\right)\right)}{8\sqrt{\pi}\sigma_m^5}(\rho^4\sigma_s^4/12 - \rho^2\sigma_s^2 + 1). \quad \text{(A24)}$$

Using these results, we obtain the optimal bandwidth

$$h^* = \left[\frac{3(\rho^4\sigma_s^4/12 - \rho^2\sigma_s^2 + 1)}{4\sigma_m^5}\exp\left(\mu_s + \frac{1}{2}\sigma_s^2\left(1 - \frac{\rho^2}{2}\right)\right)\right]^{-\frac{1}{3}}(n/T)^{-1/3}, \quad \text{(A25)}$$

where $c_2 \geq 0$ when either $\rho^2\sigma_s^2 \leq 6 - 2\sqrt{6}$, or $\rho^2\sigma_s^2 \geq 6 + 2\sqrt{6}$.

Finally, we can use the Gaussian reference model to obtain closed form expressions of the smoothing bias and the EIV bias. Differentiating $\psi(m)$ twice, we obtain[7]

$$\begin{aligned}\psi^{(2)}(m) &= \exp\left(\mu_s + \rho\sigma_s\left(\frac{m-\mu_m}{\sigma_m}\right) + \frac{1}{2}\sigma_s^2(1-\rho^2)\right)\phi(m) \\ &\quad \times\left\{\left(\frac{\sigma_s\rho}{\sigma_m}\right)^2 - 2\frac{\sigma_s\rho}{\sigma_m^2}\left(\frac{m-\mu_m}{\sigma_m}\right) + \frac{1}{\sigma_m^2}\left[\left(\frac{m-\mu_m}{\sigma_m}\right)^2 - 1\right]\right\} \\ &= \exp\left(\mu_s + \frac{1}{2}\sigma_s^2\right)\frac{1}{\sigma_m^2}\left(\left(\frac{m-\mu_m-\rho\sigma_s\sigma_m}{\sigma_m}\right)^2 - 1\right) \\ &\quad \times\frac{1}{\sigma_m}\varphi\left(\frac{m-\mu_m-\rho\sigma_s\sigma_m}{\sigma_m}\right). \quad \text{(A26)}\end{aligned}$$

Using Equations (A21) and (A26), we can replace $\phi^{(2)}(m)$ and $\psi^{(2)}(m)$ in Equations (A12) and (A13) to obtain the two bias terms under the reference model:

$$bs_1^r(m) = \frac{1}{2}h^2K_2\phi^{(2)}(m) = \left[\frac{1}{2}h^2\frac{1}{\sigma_m^2}(\bar{m}_1^2 - 1)\right]\frac{1}{\sigma_m}\varphi(\bar{m}_1), \quad \text{(A27)}$$

$$bs_2^r(m) = \frac{1}{2T}\psi^{(2)}(m) = \left[\frac{1}{2T}\exp(\mu_s + \frac{1}{2}\sigma_s^2)\frac{1}{\sigma_m^2}(\bar{m}_2^2 - 1)\right]\frac{1}{\sigma_m}\varphi(\bar{m}_2), \quad \text{(A28)}$$

[7] Alternatively, we can directly derive Equation (A26) by (i) rewriting $\psi(m)$ as a recentered Gaussian density up to a multiplicative constant, i.e., $\psi(m) = \exp\left(\mu_s + \frac{1}{2}\sigma_s^2\right)\frac{1}{\sigma_m}\varphi\left(\frac{m-\mu_m-\rho\sigma_s\sigma_m}{\sigma_m}\right)$, and (ii) differentiating this expression twice.

where $\bar{m}_1 = \frac{m-\mu_m}{\sigma_m}$, and $\bar{m}_2 = \frac{m-\mu_m-\rho\sigma_m\sigma_s}{\sigma_m}$. In our implementation, the parameters of the bivariate Gaussian distribution are estimated by the sample moments of $\hat{m}_i$ and $\hat{s}_i = \log \hat{S}_i$. QED

E Estimators of the Distribution Characteristics

To compute the characteristics of the distribution, we use a numerical approach based on the bias-adjusted density $\tilde{\phi}(m)$. For the moments, we simply use the respective definitions of the standard deviation, skewness, and kurtosis:

$$SD = V^{\frac{1}{2}} = \left(\int \phi(u)(u-M)^2 du\right)^{\frac{1}{2}}, \tag{A29}$$

$$Sk = \frac{\int \phi(u)(u-M)^3 du}{V^{\frac{3}{2}}}, \tag{A30}$$

$$Ku = \frac{\int \phi(u)(u-M)^4 du}{V^2}, \tag{A31}$$

where V denotes the variance of the distribution. To obtain the bias-adjusted estimators $\widetilde{SD}$, $\widetilde{Sk}$, and $\widetilde{Ku}$, we replace $\phi(u)$ with the bias-adjusted density estimator $\tilde{\phi}(u)$ in the above expressions We also compute the mean $\tilde{M}$ as the empirical average of the estimated measures which does not suffer from the EIV bias: $\tilde{M} = \hat{M} = \frac{1}{n_\chi}\sum_i \hat{m}_i \mathbf{1}_i^\chi$.[8] Once we have the bias-corrected estimates, we can approximate the asymptotic variance of the mean, standard deviation, skewness, and kurtosis using the delta method to conduct statistical inference:

(i) For the estimated mean, we have the asymptotic variance:

$$V\left[\tilde{M}\right] = \frac{V}{n}, \tag{A32}$$

which only requires a consistent estimator of the variance of the distribution V.

(ii) For the estimated volatility, we have:

$$V\left[\widetilde{SD}\right] = \frac{E\left[\left((2SD)^{-1}\Psi_2\right)^2\right]}{n}, \tag{A33}$$

[8] The integrals of the bias terms $bs_1^\tau(m)m$ and $bs_2^\tau(m)m$ are equal to zero (footnote 4), which implies that the empirical average is the same as the average obtained via a numerical integration of $\tilde{\phi}(m)$.

where $\Psi_2 = (m_i - E[m_i])^2 - E\left[(m_i - E[m_i])^2\right]$.

(iii) For the estimated skewness, we have:

$$V\left[\widetilde{Sk}\right] = \frac{E\left[\left(SD^{-3}\Psi_3 - \frac{3}{2}SD^{-2}Sk\Psi_2 - 3SD^{-1}\Psi_1\right)^2\right]}{n}, \tag{A34}$$

where $\Psi_3 = (m_i - E[m_i])^3 - E\left[(m_i - E[m_i])^3\right]$, and $\Psi_1 = (m_i - E[m_i]) - E\left[(m_i - E[m_i])\right]$ (see Bai and Ng (2005)).

(iv) For the estimated kurtosis $\widetilde{Ku}$, we have:

$$V\left[\widetilde{Ku}\right] = E\left[\left(SD^{-4}\Psi_4 - 2SD^{-2}Ku\Psi_2 - SD^{-1}Sk\Psi_1\right)^2\right], \tag{A35}$$

where $\Psi_4 = (m_i - E[m_i])^4 - E\left[(m_i - E[m_i])^4\right]$ (see Bai and Ng (2005)).

We also use a numerical approach to compute the proportion and quantile estimators. We denote the proportion of funds with a measure m_i below the threshold m by $\Phi(m) = P[m_i \leq m]$ and the quantile at any given percentile level $p \in (0,1)$ by $Q(p) = \Phi^{-1}(p)$, where Φ is the cdf. We obtain bias-adjusted estimators of $\Phi(m)$ and $Q(p)$ via a numerical integration of the density, i.e., we have

$$\Phi(m) = \int_{-\infty}^{m} \phi(u)du, \tag{A36}$$

$$\int_{-\infty}^{Q(p)} \phi(u)du = p, \tag{A37}$$

where we replace $\phi(u)$ with the bias-adjusted density estimator $\tilde{\phi}(u)$ (for the quantile, we use an iterative procedure until Equation (A37) holds). We can then use the bias-corrected estimated proportion and quantile to estimate their asymptotic variances

$$V[\tilde{\Phi}(m)] = \frac{\Phi(m)(1-\Phi(m))}{n}, \tag{A38}$$

$$V[\tilde{Q}(p)] = \frac{\frac{p(1-p)}{\phi(Q(p))^2}}{n}, \tag{A39}$$

where ϕ is the normal density obtained from the Gaussian reference model.

II Overview of the Analytical Approach

A Moments

An alternative to the numerical approach described above is to estimate the distribution characteristics using an analytical approach. Asymptotically, both approaches (numerical and analytical) are equivalent.

To begin, we consider the estimation of the cross-sectional expectation $E[g(m_i)]$, where g is a given smooth function of m_i. We investigate the convergence properties of the cross-sectional estimator $\frac{1}{n}\sum_{i=1}^{n} g(\hat{m}_i)\mathbf{1}_i^{\chi}$ based on the OLS estimates $\hat{m}_i$ of the non-trimmed assets. The following proposition proves the asymptotic normality of the estimator under the baseline specification $r_{i,t} = a_i - b_i q_{i,t-1} + \beta_i' f_t + \varepsilon_{i,t}$.

Proposition A.1. As $n, T \to \infty$, such that $n = o(T^3)$,

$$\sqrt{n}\left(\frac{1}{n}\sum_i g(\hat{m}_i)\mathbf{1}_i^{\chi} - E\left[g(m_i)\right] - \mathcal{B}_T\right) \Rightarrow N\left(0, V\left[g(m_i)\right]\right), \tag{A40}$$

where $\mathcal{B}_T = \frac{1}{2T}E[g^{(2)}(m_i)S_i]$ and $V\left[g(m_i)\right]$ is the cross-sectional variance of $g(m_i)$.

Proof of Proposition A.1. Equation (A4) yields the mean value expansion

$$g(\hat{m}_i) = g(m_i) + g^{(1)}(\bar{m}_i)\frac{1}{\sqrt{T}}\hat{\eta}_{i,T} + g^{(2)}(\bar{m}_i)\frac{1}{2T}\hat{\eta}_{i,T}^2, \tag{A41}$$

where $\bar{m}_i$ lies between $\hat{m}_i$ and m_i. Then, we get

$$\begin{aligned}
&\sqrt{n}\left(\frac{1}{n}\sum_i g(\hat{m}_i)\mathbf{1}_i^{\chi} - E\left[g(m_i)\right] - \mathcal{B}_T\right) \\
= \;& \frac{1}{\sqrt{n}}\sum_i \left(g(m_i) - E\left[g(m_i)\right]\right) - \frac{1}{\sqrt{n}}\sum_i g(m_i)(1-\mathbf{1}_i^{\chi}) + \frac{1}{\sqrt{nT}}\sum_i \mathbf{1}_i^{\chi} g^{(1)}(\bar{m}_i)\hat{\eta}_{i,T} \\
& + \frac{1}{2T}\frac{1}{\sqrt{n}}\sum_i \left(\mathbf{1}_i^{\chi} g^{(2)}(\bar{m}_i)\hat{\eta}_{i,T}^2 - E\left[g^{(2)}(m_i)S_i\right]\right) \\
\equiv \;& I_{21} + I_{22} + I_{23} + I_{24}.
\end{aligned} \tag{A42}$$

We have $I_{22} = o_p(1)$ and $I_{23} = O_p(1/\sqrt{T}) = o_p(1)$ using similar arguments as in Lemma 2 of Gagliardini, Ossola, and Scaillet (2016). The remainder term $I_{24} = O_p(\sqrt{n/T^3} + \sqrt{n}/T^2 + 1/T)$, which gives $I_{24} = o_p(1)$ if $n = o(T^3)$.[9] Therefore, the asymptotic distribution in Equation (A38) depends on the first term $I_{21} \Rightarrow N(0, V[g(m_i)])$ from the

[9] The condition $n = o(T^3)$ is used to control the remainder term in the Taylor expansion of the function g and the bias term.

standard CLT. QED

The distribution results in Equation (A38) reveal that we have an asymptotic bias $\mathcal{B}_T$ of order $1/T$ which comes from the estimation error of $\hat{m}_i$ (EIV contribution). To compute the bias-adjusted estimated mean, standard deviation, skewness, and kurtosis, we can use an analytical approach (based on the delta method) and replace the unknown moments with consistent estimators based on empirical averages:

(i) The mean is given by $M = E[m_i]$. Therefore, the asymptotic bias $\mathcal{B}_T$ is zero because $g^{(2)}(m) = 0$. For this particular case, we do not need the condition $n = o(T^3)$ for the above proposition to hold.

(ii) The variance is given by $V = E\left[(m_i - E[m_i])^2\right]$. To obtain the bias of the standard deviation $SD = V^{\frac{1}{2}}$, we apply the delta method:

$$\mathcal{B}_T(SD) = (2SD)^{-1}\mathcal{B}_T(E[m_i^2]), \tag{A43}$$

where the asymptotic bias of the second moment is given by

$$\mathcal{B}_T(E[m_i^2]) = \frac{1}{2T}E[2S_i]. \tag{A44}$$

(iii) The skewness is given by $Sk = E\left[(m_i - E[m_i])^3\right] / E\left[(m_i - E[m_i])^2\right]^{3/2}$. Applying the delta method, we obtain

$$\mathcal{B}_T(Sk) = (\nabla_3 Sk)\mathcal{B}_T(E[m_i^3]) + (\nabla_2 Sk)\mathcal{B}_T(E[m_i^2]), \tag{A45}$$

where $\nabla_j Sk$ denotes the derivative of Sk w.r.t. $E[m_i^j]$ and the different terms are given by

$$\begin{aligned}
\mathcal{B}_T(E[m_i^3]) &= \frac{1}{2T}E[6m_iS_i], \\
\nabla_3 Sk &= V[m_i]^{-3/2}, \\
\nabla_2 Sk &= -3E[m_i]V[m_i]^{-3/2} + E[m_i^3](\frac{-3}{2})V[m_i]^{-5/2} \\
&\quad + \left\{-3E[m_i^2]E[m_i] + 2E[m_i]^3\right\}(\frac{-3}{2})V[m_i]^{-5/2}.
\end{aligned} \tag{A46}$$

(iv) The kurtosis is given by $Ku = E\left[(m_i - E[m_i])^4\right] / E\left[(m_i - E[m_i])^2\right]^2$. Applying the delta method, we obtain

$$\mathcal{B}_T(Ku) = (\nabla_4 Ku)\mathcal{B}_T(E[m_i^4]) + (\nabla_3 Ku)\mathcal{B}_T(E[m_i^3]) + (\nabla_2 Ku)\mathcal{B}_T(E[m_i^2]), \quad \text{(A47)}$$

where the different terms are given by

$$\begin{aligned}
\mathcal{B}_T(E[m_i^4]) &= \frac{1}{2T}E[12m_i^2 S_i], \\
\nabla_4 Ku &= V[m_i]^{-2}, \\
\nabla_3 Ku &= -4E[m_i]V[m_i]^{-2}, \\
\nabla_2 Ku &= 6E[m_i]^2 V[m_i]^{-2} + \left\{E[m_i^4] - 4E[m_i^3]E[m_i]\right\}(-2)V[m_i]^{-3} \\
&\quad + \left\{6E[m_i^2]E[m_i]^2 - 3E[m_i]^4\right\}(-2)V[m_i]^{-3}. \qquad \text{(A48)}
\end{aligned}$$

B Proportion and Quantile

We now turn to the analysis of the proportion estimator inferred from the cumulative distribution function (cdf) and the associated quantile. The proportion estimator is the cross-sectional average of the indicator function $g(\hat{m}_i) = \mathbf{1}\{\hat{m}_i \leq m\}$ based on the OLS estimates $\hat{m}_i$ for the non-trimmed assets, $\hat{\Phi}(m) = \frac{1}{n_\chi}\sum_i \mathbf{1}\{\hat{m}_i \leq m\}\mathbf{1}_i^\chi$, while the quantile estimator is the inverse function $\hat{Q}(p) = \hat{\Phi}^{-1}(p)$.

The next proposition extends Proposition A.1 to the proportion and quantile.

Proposition A.2. As $n, T \to \infty$, such that $n = o(T^3)$,

$$\sqrt{n}\left(\hat{\Phi}(m) - \Phi(m) - \mathcal{B}_T(m)\right) \Rightarrow N\left(0, V[\hat{\Phi}(m)]\right), \quad \text{(A49)}$$

$$\sqrt{n}\left(\hat{Q}(p) - Q(p) + \frac{\mathcal{B}_T(Q(p))}{\phi(Q(p))}\right) \Rightarrow N\left(0, V[\hat{Q}(p)]\right), \quad \text{(A50)}$$

where $\mathcal{B}_T(m) = \frac{1}{2T}\psi^{(1)}(m)$, $V[\hat{\Phi}(m)] = \Phi(m)(1 - \Phi(m)$, and $V[\hat{Q}(p)] = \frac{p(1-p)}{\phi(Q(p))^2}$.

Proof of Proposition A.2. The proof builds on our previous analysis. From Equation (A4), we have $E\left[1\{\hat{m}_i \leq m\}\right] = P\left[m_i + \frac{1}{\sqrt{T}}\hat{\eta}_{i,T} \leq m\right]$. By using the results in Gourieroux, Laurent, and Scaillet (2000), Martin and Wilde (2001), and Gagliardini and Gourieroux (2011), we obtain:

$$\begin{aligned}
P\left[m_i + \frac{1}{\sqrt{T}}\hat{\eta}_{i,T} \leq m\right] &= \Phi(m) - \frac{1}{\sqrt{T}}\phi(m)E[\hat{\eta}_{i,T}|m_i = m] \\
&\quad + \frac{1}{2T}\frac{d}{dm}(\phi(m)E[\hat{\eta}_{i,T}^2|m_i = m]) + o(1/T). \qquad \text{(A51)}
\end{aligned}$$

From Equation (A47), the bias expansion is such that: $E[\hat{\Phi}(m)] - \Phi(m) = \mathcal{B}_T(m) +$

$E\left[1\{\hat{m}_i \leq m\}(1-\mathbf{1}_i^{\chi})\right] + o(1/T)$. We deduce the asymptotic normality of the proportion estimator by controlling the different terms and applying the CLT. To deduce the asymptotic normality of the quantile estimator, we use the Bahadur expansion for the quantile estimator at level $u \in (0,1)$: $\hat{Q}(p) - Q(p) = -\frac{1}{\phi(Q(p))}\left(\hat{\Phi}(Q(p)) - p\right)$. QED

As in the previous section, we can approximate the asymptotic bias using the Gaussian reference model.[10] With our bivariate Gaussian reference model, the term $\psi^{(1)}(m)$ in the bias is equal to

$$\begin{aligned}
\psi^{(1)}(m) &= \exp\left(\mu_s + \rho\sigma_s\left(\frac{m-\mu_m}{\sigma_m}\right) + \frac{1}{2}\sigma_s^2(1-\rho^2)\right)\phi(m)\left(\frac{\sigma_s\rho}{\sigma_m} - \frac{m-\mu_m}{\sigma_m^2}\right) \\
&= \exp\left(\mu_s + \frac{1}{2}\sigma_s^2\right)\frac{-1}{\sigma_m}\left(\frac{m-\mu_m-\rho\sigma_s\sigma_m}{\sigma_m}\right) \\
&\quad \times\frac{1}{\sigma_m}\varphi\left(\frac{m-\mu_m-\rho\sigma_s\sigma_m}{\sigma_m}\right). \qquad \text{(A52)}
\end{aligned}$$

III Analysis of the EIV Bias Adjustment

In this section, we provide additional information on the EIV bias adjustment obtained with the Gaussian reference model. As explained in the paper, this approach is appealing because the bias adjustment is available in closed form. It is also precisely estimated because of parsimony—it only depends on the five parameters of the normal distribution $\theta = (\mu_m, \sigma_m, \mu_s, \sigma_s, \rho)'$. These benefits are not shared by a fully nonparametric approach in which the bias is estimated via a nonparametric estimation of the second-order derivatives $\phi^{(2)}$ and $\psi^{(2)}$.[11]

An important question is whether the EIV bias obtained with the normal reference model provides a good approximation of the true bias (i.e., whether $bs_2^r(m) \approx bs_2(m)$). Two compelling arguments show that it is the case. First, Proposition III.1 shows that the true bias $bs_2(m)$ is a function of the second-order derivative of the true density ϕ. As long as ϕ peaks around its mean, this derivative takes negative values in the center and positive values in the tails—exactly like the function $bs_2^r(m)$. The two terms $bs_2(m)$ and $bs_2^r(m)$ only differ if ϕ is a mixture of distributions whose components have means

[10] The asymptotic bias takes the same form as the one in Jochmans and Weidner (2018) where they consider n parameters of interest directly drawn from a Gaussian distribution whose measurement errors decrease at a parametric rate $\sqrt{T}$. In their setting, they use other arguments based on the behaviour of the probability integral transform for their proofs. In a different context, Okui and Yanagi (2019) also derive an estimator of the cdf to examine the mean and autocorrelation of random variables.

[11] We can estimate the rth-derivative of a density ϕ by kernel smoothing (Bhattacharya (1967)). The rate of consistency of the derivative estimator equals $\sqrt{nh^{2r+1}}$ and is much slower than the rate $\sqrt{nh}$ for the density estimator. In other words, the higher-order derivatives are imprecisely estimated because the rate of consistency decreases with the derivative's order r.

extremely far away from one another. In this case, we have a trough instead of a peak around the mean.

Second, our extensive Monte-Carlo analysis calibrated on the data reveals that the bias-adjusted density captures the true density well (see Section IV). Our Monte-Carlo analysis resonates with the one performed by Silverman (1986) for the standard non-parametric density estimation without the EIV problem. He shows that the rule of thumb for the bandwidth choice, which relies on a normal reference model, is quite robust to departures from normality.

The reference model allows us to conduct a comparative static analysis of the EIV bias. As shown in Equation (A28), there are three key parameters that determine $bs_2^r(m)$: (i) the variance of the true measure σ_m^2, (ii) the average across funds of the variance of the estimated measure, measured as $\sigma_{\hat{m}}^2 = \frac{1}{T}E[S_i] = \frac{1}{T}\exp(\mu_s + \frac{1}{2}\sigma_s^2)$, and (iii) the correlation ρ between the true measure and estimation variance.

A higher value of σ_m^2 reduces the magnitude of the EIV bias because it makes the cross-sectional variation of the estimated measure more aligned with that of the true measure (i.e., the relative importance of m_i over noise increases). On the contrary, a higher value of $\sigma_{\hat{m}}^2$ makes the EIV bias more severe because the estimated measure becomes more volatile (i.e., the relative importance of noise over m_i increases). Finally, a higher value of $|\rho|$ keeps the shape of the bias unchanged, but creates asymmetry.

In Figure A1, we quantify these changes for the skill coefficient a_i. To begin, we compute $bs_2^r(m)$ in the benchmark case where the parameters of the reference model are obtained from our sample. The mean μ_m is set equal to 0.24% per month, the variance terms σ_m^2 and $\sigma_{\hat{m}}^2$ are equal to $\frac{0.0017}{100}$ and $\frac{0.0011}{100}$, and the correlation ρ reaches 0.21. Plugging these parameter values in Equation (A28), we find that the EIV bias adjustment requires a transfer of probability mass from the tails to the center equal to 15%. This proportion is obtained by integrating $bs_2^r(m)$ over the area for which $bs_2^r(m)$ takes negative values.

Next, we sequentially increase the values of (i) σ_m^2 from $\frac{0.0017}{100}$ to $\frac{0.0037}{100}$, (ii) $\sigma_{\hat{m}}^2$ from $\frac{0.0011}{100}$ to $\frac{0.0031}{100}$, and (iii) ρ from 0.21 to 0.44. We find that changes in the variance terms have a significant impact on the shape of the EIV bias. Panel A shows that increasing σ_m^2 reduces the transfer of probability from 15% to just 7%, while Panel B shows that increasing $\sigma_{\hat{m}}^2$ implies an increase in probability transfer from 15% to 26%. Finally, Panel C shows that increasing ρ implies that 87% of the probability transfer (0.13/0.15) is at the right of the mean (versus 75% in the baseline case (0.10/0.15)).

Please insert Figure A1 here

IV Monte-Carlo Simulations

A The Setup

We now conduct a Monte-Carlo analysis to evaluate the finite-sample properties of the estimated skill and scale distributions obtained with our nonparametric approach. We consider a hypothetical population of n funds with T return observations ($n = 1,000$, 2,500, 5,000, and 10,000; $T = 100$, 250, 500, and $1,000$). To model the fund return $r_{i,t}$ and its lagged size $q_{i,t-1}$, we use the baseline specification

$$r_{i,t} = a_i - b_i q_{i,t-1} + \beta_i' f_t + \varepsilon_{i,t}, \tag{A53}$$

along with an AR(1) model for the log size $lq_{i,t-1} = log(q_{i,t-1})$ to ensure the positivity of fund size,

$$lq_{i,t} = \theta_{lq_i} + \rho_{lq} lq_{i,t-1} + e_{lq_i,t}, \tag{A54}$$

where f_t is the vector of four factors (market, size, value, and momentum), $\theta_{lq_i} = \mu_{lq_i}(1 - \rho_{lq})$, and $\mu_{lq_i} = E[lq_{i,t-1}]$. The residual terms $\varepsilon_{i,t}$ and $e_{lqi,t}$ are drawn from a bivariate normal: $\varepsilon_{i,t} \sim N(0, \sigma^2_{\varepsilon_i})$, $e_{lq_i,t} \sim N(0, \sigma^2_{e_{lq_i}})$, where $\sigma^2_{e_{lq_i}} = (1-\rho^2_{lq})\sigma^2_{lq_i}$ and $\sigma^2_{lq_i}$ is the variance of $lq_{i,t-1}$. We also account for the positive correlation between the fund residual and the innovation in fund size by setting $corr(\varepsilon_{i,t}, e_{lqi,t})$ equal to ρ.

To determine the values for the fund-specific parameters $\{a_i,\ b_i,\ \beta_i',\ \mu_{lq_i},\ \sigma^2_{lq_i}\}$, we randomly draw from the estimated vectors observed in our sample $\{\hat{a}_i,\ \hat{b}_i,\ \hat{\beta}_i',\ \hat{\mu}_{lq_i},\ \hat{\sigma}^2_{lq_i}\}$. This approach allows us to maintain the correlation structure between the different parameters, in particular between the skill coefficient, the scale coefficient, and the size parameters: $\mu_{lq_i} = \mu_{lq_i}(a_i, b_i)$, $\sigma^2_{lq_i} = \sigma^2_{lq_i}(a_i, b_i)$.[12] The remaining parameters are calibrated using the median values in the data, which yields $\rho_{lq} = 0.97$, $\rho = 0.20$, and $\sigma^2_{\varepsilon_i} = 0.0167^2$.

To reproduce the salient features of the skill and scale distributions, we rescale the estimated values of $\hat{a}_i$ and $\hat{b}_i$ to match the cross-sectional volatility reported in Table II of the paper (4.1% and 1.7% per year for a_i and b_i). The true distributions of the skill and scale coefficients are both non normal (the skewness is equal to 0.7 and 0.9, and the kurtosis is equal to 11.7 and 12.1). Therefore, our Monte-Carlo setting allows us to

[12] In particular, we capture the strong correlation between the skill and scale coefficients. Interestingly, this correlation has implications for modeling the prior distributions of a_i and b_i in an empirical Bayes setting. For instance, Pastor and Stambaugh (2012) elicit the joint prior distribution of a_i and b_i by setting their correlation equal to zero. Therefore, investors in their model believe that the variance of $\alpha_{i,t}$ is higher than the one inferred from an empirical Bayes prior. This initial belief implies a lower allocation to active funds which could persist for a long time.

examine the properties of the estimators when the Gaussian reference model (used for the EIV bias adjustment) differs from the true distributions.

Conditional on the values $\{\hat{a}_i,\, \hat{b}_i,\, \hat{\beta}_i',\, \hat{\mu}_{lq_i},\, \hat{\sigma}^2_{lq_i}\}$ taken by each fund, we examine the properties of the estimators. For each iteration s ($s = 1, ..., 500$), we build the return and size time-series of each fund as follows. First, we draw the initial value of $lq_{i,0}(s)$ from its unconditional distribution: $lq_{i,0}(s) \sim N(\mu_{lq_i}, \sigma^2_{lq_i})$. Second, we draw the vector $f_1(s)$ from the realized values in the sample, and the innovations $\varepsilon_{i,1}(s)$ and $e_{lq_i,1}(s)$ from the bivariate normal. Third, we construct the fund gross return and log size at time 1 as

$$\begin{aligned} r_{i,1}(s) &= a_i - b_i q_{i,0}(s) + \beta_i' f_1(s) + \varepsilon_{i,1}(s), \\ lq_{i,1}(s) &= \theta_{lq,i} + \rho_{lq} lq_{i,0}(s) + e_{lq_i,1}(s), \end{aligned} \tag{A55}$$

where $q_{i,0}(s) = \exp(lq_{i,0}(s))$. Fourth, we repeat the two previous steps for each time t ($t = 2, ..., T$), we obtain the entire time-series for the fund gross return and size: $r_{i,1}(s), ..., r_{i,T}(s),\ q_{i,0}(s), ..., q_{i,T-1}(s)$. Fifth, we apply our nonparametric approach to compute the bias-adjusted density $\tilde{\phi}(s)$ and a set of several distribution characteristics that include the mean, volatility, skewness, and the proportion of funds with a positive measure $\pi^+ = 1 - \Phi(0)$. Finally, we repeat the entire procedure across all S iterations.

To assess the performance of the bias-adjusted density $\tilde{\phi}$, we compute the Mean Integrated Squared Error (MISE) defined as

$$MISE = \int [\sigma^2(m) + bs(m)^2] dm, \tag{A56}$$

where the bias and variance functions are given by

$$bs(m) = \frac{1}{S} \sum_{s=1}^{S} \tilde{\phi}(m; s) - \phi(m), \tag{A57}$$

$$\sigma^2(m) = \frac{1}{S} \sum_{s=1}^{S} \left(\tilde{\phi}(m; s) - \frac{1}{S} \sum_{s=1}^{S} \tilde{\phi}(m; s) \right)^2. \tag{A58}$$

For the moment/proportion estimator $\tilde{\varphi}$ ($\tilde{\varphi} = \tilde{M},\, \widetilde{SD},\, \widetilde{Sk},\, \tilde{\pi}^+$), we compute the Mean Squared Error (MSE) as

$$MSE(\tilde{\varphi}) = \sigma^2(\tilde{\varphi}) + bs^2(\tilde{\varphi}), \tag{A59}$$

where the bias and the variance terms are given by

$$bs(\tilde{\varphi}) = \frac{1}{S}\sum_{s=1}^{S}\tilde{\varphi}(s) - \varphi, \tag{A60}$$

$$\sigma^2(\tilde{\varphi}) = \frac{1}{S}\sum_{s=1}^{S}\left(\tilde{\varphi}(s) - \frac{1}{S}\sum_{s=1}^{S}\tilde{\varphi}(s)\right)^2. \tag{A61}$$

B Main Results

In Table AI, we report the MISE and its two components (integrated squared bias and variance) for the skill distribution. In Panel A, we compute the MISE of the bias-adjusted density $\tilde{\phi}(m)$ for the baseline choice of the optimal bandwidth $h^* \sim c_2^{-1/3}(n/T)^{-1/3}$ (shown in Equation (A16)). In Panel B, we repeat the analysis for the alternative choice of the optimal bandwidth under which $h^* \sim c_1^{-1/5}n^{-1/5}$. Finally, Panel C reports the MISE of the estimated density $\hat{\phi}(m)$ obtained with the standard approach which does not adjust for the bias.

Our analysis reveals two main insights. First, accounting for the EIV bias improves the estimation of the true distribution $\phi(m)$. To illustrate, we consider the scenario where $n = 2,500$ and $T = 250$, which is representative of our actual sample after trimming (i.e., $\sum_{i=1}\mathbf{1}_i^{\chi} \approx 2,500$ and $\frac{1}{n_\chi}\sum_{i=1}\mathbf{1}_i^{\chi}T_i \approx 250$) . We find that the MISE of $\hat{\phi}(m)$ is nearly two times larger than the level observed for $\tilde{\varphi}(m)$ with the baseline bandwidth (4.96 vs 9.06). Second, our nonparametric approach yields a stronger performance under the baseline choice for the optimal bandwidth—in all scenarios, using the alternative bandwidth choice produces a higher MISE.

In Table AII, we examine the performance of the moment and proportion estimators for the skill distribution. Panel A shows the MSE and its two components (bias and standard deviation) of each bias-adjusted estimator obtained via a numerical integration of $\tilde{\phi}(m)$ (using the baseline bandwidth). Panel B reports the same statistics for the bias-adjusted estimators obtained with the analytical approach described in Section II. For comparison, Panel C reports the bias-unadjusted estimators (obtained via a numerical integration of $\hat{\phi}(m)$).

The results show that the bias-adjusted estimators perform better when the numerical integration is used. In most cases, it produces a lower MSE than the one obtained with the analytical formulas. We also find that the unadjusted estimators are markedly biased. When $n = 2,500$ and $T = 250$, the bias for the volatility and the unadjusted proportion is equal to 1.08% per year and -5.52%, respectively. In contrast, our nonparametric approach reduces the bias for all quantities. Overall, these findings highlight the importance of controlling for the bias.

Next, we turn to the analysis of the scalability distribution. Tables AIII and AIV report the MISE of the estimated density and the MSE of the moment and proportion estimators. Similar to the skill coefficient, we find that $\tilde{\varphi}(m)$ outperforms $\hat{\phi}(m)$. If $n = 2,500$ and $T = 250$, the difference in MISE between the two estimated densities is equal to 11.47 (28.58 vs 17.01). The bias adjustment is also important for the other estimators. For instance, the standard approach underestimates the proportion of funds with a positive scale coefficient by 6.8% (in absolute terms).

To sum up, the Monte Carlo analysis yields three main insights. First, the EIV bias has a notable impact on the different estimators and thus cannot be ignored. Second, the baseline choice for the optimal bandwidth produces a lower MISE for the bias-adjusted density. Third, the numerical approach generally outperforms the analytical approach. These results justify the use of the optimal bandwidth in Equation (A16) and the numerical approach for the empirical analysis of the paper.

Please insert Tables AI to AIV here

C Simulations with Uncorrelated Skill and Scalability

The asymptotic distribution of the OLS estimators $\hat{a}_i$ and $\hat{b}_i$ implies that they are correlated at the fund level. If, for simplicity, we omit the factors f_t, we have

$$\sqrt{T}\begin{bmatrix} \hat{a}_i - a_i \\ \hat{b}_i - b_i \end{bmatrix} \Rightarrow N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \frac{E[q_{i,t-1}^2]}{V[q_{i,t-1}]}\sigma_{\varepsilon_i}^2 & \frac{E[q_{i,t-1}]}{V[q_{i,t-1}]}\sigma_{\varepsilon_i}^2 \\ \frac{E[q_{i,t-1}]}{V[q_{i,t-1}]}\sigma_{\varepsilon_i}^2 & \frac{1}{V[q_{i,t-1}]}\sigma_{\varepsilon_i}^2 \end{bmatrix}\right), \tag{A62}$$

where $cov(\sqrt{T}\hat{a}_i, \sqrt{T}\hat{b}_i) = \frac{E[q_{i,t-1}]}{V[q_{i,t-1}]}\sigma_{\varepsilon_i}^2 > 0$. Therefore, one concern is that the strong cross-sectional correlation between $\hat{a}_i$ and $\hat{b}_i$ observed in the data is mechanically driven by the fund-level correlation between $\hat{a}_i$ and $\hat{b}_i$. To address this concern, we consider a world where a_i and b_i are uncorrelated across funds. Consistent with this assumption, we show that the cross-sectional correlation between $\hat{a}_i$ and $\hat{b}_i$ is equal to zero (even though the fund-level correlation between $\hat{a}_i$ and $\hat{b}_i$ is positive).

We consider a simple modification of the Monte-Carlo setup in which a_i and b_i are uncorrelated across funds. We draw the true coefficients a_i and b_i of each fund i ($i = 1, ..., 2,500$) independently from the estimated vectors observed in our sample (rescaled to match the cross-sectional volatility in Table II of the paper). If a_i and b_i are positive, we assume that the average size $E[q_{i,t-1}]$ is equal to $\frac{a_i}{2b_i}$ (as in the model of Berk and Green (2004), and that $\sigma_{lq_i}^2$ is proportional to μ_{lq_i} (by a factor k calibrated on the data). These two assumptions provide a simple way to model the link that exists between skill, scale, and size. Specifically, we have $\mu_{lq,i} + \frac{1}{2}\sigma^2_{lq_i} = \log\left(\frac{a_i}{2b_i}\right) \Rightarrow \left(1 + \frac{1}{2}k\right)\mu_{lq,i} = \log\left(\frac{a_i}{2b_i}\right) \Rightarrow \mu_{lq,i} = \log\left(\frac{a_i}{2b_i}\right) / \left(1 + \frac{1}{2}k\right)$. With a log-normally distributed size, we can then compute the parameters of the asymptotic distribution in Equation (A60) as $E[q^2_{i,t-1}] = e^{2\mu_{lq,i}+2k\mu_{lq,i}}$ and $V[q_{i,t-1}] = E[q^2_{i,t-1}] - (E[q_{i,t-1}])^2$. Otherwise, if a_i and b_i are negative, we measure $E[q_{i,t-1}]$, $E[q^2_{i,t-1}]$, and $V[q_{i,t-1}]$ as the median values among funds for which $\hat{a}_i$ or $\hat{b}_i$ are negative.

For each iteration s ($s = 1, ..., 500$), we draw $[\hat{a}_i(s), \hat{b}_i(s)]'$ from the asymptotic distribution of each fund in Equation (A60). We then compute the average fund-level correlation (FLC) between $\hat{a}_i$ and $\hat{b}_i$ as

$$FLC(\hat{a}_i, \hat{b}_i) = \frac{1}{n}\sum_i \left(\frac{1}{S}\sum_s \left(\hat{a}_i(s) - a_i\right)\left(\hat{b}_i(s) - b_i\right)\right), \tag{A63}$$

and the average cross-sectional correlation (CSC) as

$$CSC(\hat{a}_i, \hat{b}_i) = \frac{1}{S}\sum_s \left(\frac{1}{n}\sum_i \left(\hat{a}_i(s) - \bar{a}(s)\right)\left(\hat{b}_i(s) - \bar{b}(s)\right)\right), \tag{A64}$$

where $\bar{a}(s) = \frac{1}{n}\sum_i \hat{a}_i(s)$ and $\bar{a}(s) = \frac{1}{n}\sum_i \hat{b}_i(s)$. Consistent with the theoretical predictions, we find that $FLC(\hat{a}_i, \hat{b}_i)$ is equal to 0.18, whereas $CSC(\hat{a}_i, \hat{b}_i)$ is essentially equal to zero (i.e., $CSC(\hat{a}_i, \hat{b}_i)$ =0.00004).

V Mutual Fund Dataset

A Construction of the Dataset

We now provide additional information on the construction of the mutual fund dataset. To begin, we collect monthly data on net returns and size, as well as annual data on fees, turnover, and investment objectives from the CRSP database between January 1975 and December 2019 (540 observations). We measure the monthly gross return of each fund as the sum of its monthly net return and fees. The net return is computed as a value-weighted average of the net returns across all shareclasses using their beginning-of-month total net asset values. The monthly fees are defined as the value-weighted average of the most recently reported annual fees across shareclasses divided by 12. We eliminate the monthly gross return observation when (i) the monthly net return is below -100% or above 100%, or when (ii) the monthly fees are below 2.5 bps (0.3% per year) or above 83 bps (10% per year). We measure fund size by taking the sum of the beginning-of-month net asset values across all shareclasses. We apply a linear interpolation to fill in missing observations when funds report size on a quarterly basis. We also adjust size for inflation by expressing all numbers in January 1, 2000 dollars (see Berk and van Binsbergen (2015)). Finally, we correct for reporting errors for the TNA.[13]

We apply a set of filters before conducting the empirical analysis. First, we remove all funds that are classifed as passive or closed for more than a third of the observations using (i) the index fund indicator (letter B, D, or E), (ii) the ETF indicator (letter F or N), (iii) and the closed fund indicator (letter N). Therefore, our sample focuses on open-end, actively managed funds with a well-defined equity style (as described below), and a weight invested in equities above 80%. Second, we eliminate funds if they are tiny, i.e., if their size is below minimum size of $15 million for more than a third of the observations (similar to Chen et al. (2004), Pastor, Stambaugh, and Taylor (2015)). Third, we delete the following-month return after a missing return observation because CRSP fills this with the cumulated return since the last nonmissing return. Fourth, we run a correlation analysis to eliminate duplicates, i.e., funds for which the return correlation is above 0.99 (using a minimum of 12 monthly observations).

To benchmark each fund, we use the four-factor model of Cremers, Petajisto, and Zitzewitz (2012) which includes the vector $f_t = (r_{m,t}, r_{smb,t}, r_{hml,t}, r_{mom,t})'$, where $r_{m,t}$, $r_{smb,t}$, $r_{hml,t}$, and $r_{mom,t}$ capture the excess returns of the market, size, value, and momentum factors. This model departs from the traditional model of Carhart (1997) in two respects: (i) the market factor is proxied by the excess return of the SP500 (instead of the CRSP index), and (ii) the size and value factors are index-based and measured as the return difference between the Russell 2000 and SP500, and between the Russell 3000 Value and Russell 3000 Growth. Because the index-based returns for size and value are not available between January 1975 and December 1978, we replace them with the values of the size and value factors obtained from Ken French's website (focusing on the period January 1979-December 2018 does not change our main results). For the momentum factor, we use data obtained from Ken French's website.

The motivation for using this model is that it correctly assigns a zero alpha to the SP500 and Russell 2000. Both indices cover about 85% of the total market capitalization and are widely used as benchmarks by mutual funds. On the contrary, the Carhart model fails to price these indices—for one, the Russell 2000 has a negative alpha of -2.4% per

[13] For instance, we find more than 1,500 observations in CRSP for which the TNA of a given shareclass jumps (or is reduced) by a factor higher than 3 in a given month before going right back to the same value the following month.

year over the period 1980-2005 (Cremers, Petajisto, and Zitzewitz (2012)). Therefore, small cap funds that use this index as a benchmark are likely to be classified as unskilled under the Carhart model.

We obtain our final universe of funds after applying the selection rule in Equation (8) of the paper. We follow Gagliardini, Ossola, and Scaillet (2016) and select funds for which the condition number of the matrix $\hat{Q}_{x,i}$ is below 15 and the number of monthly observations is above 60. These selection criteria produce a final universe of 2,427 funds. To apply our nonparametric approach, we compute the asymptotic variance of each fund measure using a lag of three months ($L = 3$). To mitigate the impact of outliers on the vector $\hat{\theta}$ of estimated parameters in the reference model, we also exclude the values for $\hat{m}_i$ and $\hat{s}_i$ whose cross-sectionally standardized values are above 10.

B Construction of the Fund Groups

To classify funds into the small cap and large cap groups, we proceed as follows. At the start of each month, we classify each fund in different style groups using the style information provided by Lipper. If this information is missing, we use the investment objectives reported by Strategic Insight, Wiesenberger, and CRSP in a sequential manner. Table AV provides the list of the 32 styles across the different data providers which are used for forming our final universe of equity funds. In addition, it shows the mapping between the 32 styles and the small/large cap dimensions. A value of: (i) 1 refers to a small cap fund, (ii) 2 refers to a mid cap fund, and (iii) 3 refers to a large cap fund. A fund is included in a given group (small cap, large cap) if its style corresponds to that of the group for the majority of its monthly observations.

Please insert Table AV here

For the turnover groups, we sort funds in three categories (low, medium, and high turnover) based on their average monthly turnover. To measure the monthly turnover of each fund, we follow Pastor, Stambaugh, and Taylor (2018) and use the most recently observed ratio of min(buys,sells) on fund size.

Finally, we construct the set of broker and direct sold funds using the procedure proposed by Del Guercio and Reuter (2014) and Sun (2020). At the start of each month, we only select shareclasses that are sold to retail investors. We consider each shareclass as direct sold if it charges no front or back load and has an annual distribution fee (12b-1 fees) of no more than 0.25% per year. Otherwise, we consider it as broker sold. Aggregating across shareclasses, the fund is then considered as broker sold (direct sold) for that particular month if at least 75% of its assets are broker sold (direct sold). A fund is included in a given group (broker sold, direct sold) if it belongs to it for the majority of its monthly observations.

VI Additional Results

A Derivation of the Specification Test

In this section, we derive a new specification test to confirm the validity of our empirical results. Our objective is to test the null hypothesis $H_{0,i}$ that our baseline linear specification $\alpha_{i,t} = a_i - b_i q_{i,t-1}$ is correct for each fund. Our specification test follows the strategy of a Hausman test which evaluates the difference beween two consistent estimators under the null hypothesis of well specification (Hausman (1978)). In our context, we compare the linear estimator of the gross alpha $\hat{\alpha}_{i,t}$ with its model-free version proposed by Berk and van Binsbergen (2015) and denoted by $\hat{\alpha}_{i,t}^{bvb}$. Whereas the two estimators converge to the same quantity under the null hypothesis $H_{0,i}$, they converge to different quantities under the alternative hypothesis of misspecification.

We consider our baseline model

$$r_{i,t} = a_i - b_i q_{i,t-1} + \beta_i' f_t + \varepsilon_{i,t}, \tag{A65}$$

and want to test this specification against the extended time-series regression model

$$r_{i,t} = a_i - b_i q_{i,t-1} + c_i' p_{i,t-1} + \beta_i' f_t + \varepsilon_{i,t}, \tag{A66}$$

where $p_{i,t-1}$ is a vector of variables that are omitted in our baseline specification and orthogonal to the factors f_t and error $\varepsilon_{i,t}$. To ease the presentation, we do not explictly include the small-sample bias correction which has no impact on the asymptotic analysis (see Section I.B).

The linear estimator of the gross alpha under Equation (A65) is

$$\hat{\alpha}_{i,t} = \hat{a}_i - \hat{b}_i q_{i,t-1}, \tag{A67}$$

where $\hat{a}_i = e_1'\hat{\gamma}_i$, $\hat{b}_i = e_2'\hat{\gamma}_i$, $\hat{\gamma}_i = \hat{Q}_{x,i}^{-1}\frac{1}{T_i}\sum_t I_{i,t}x_{i,t}r_{i,t}$, $\hat{Q}_{x,i} = \frac{1}{T_i}\sum_t I_{i,t}x_{i,t}x_{i,t}'$, $x_{i,t} = (1, -q_{i,t-1}, f_t')'$, and e_1 (e_2) is a vector with one in the first (second) position and zeros elsewhere. The model-free estimator of the gross alpha is given by

$$\hat{\alpha}_{i,t}^{bvb} = r_{i,t} - \hat{\beta}_i' f_t, \tag{A68}$$

where $\hat{\beta}_i = E_2'\tilde{Q}_x^{-1}\frac{1}{T_i}\sum_t I_{i,t}x_t r_{i,t}$, E_2 is a selection matrix that selects the lower K_f-subvector of coefficients, $\tilde{Q}_{x,i} = \frac{1}{T_i}\sum_t I_{i,t}x_t x_t'$, and $x_t = (1, f_t')'$.

We build the difference $\Delta\hat{\alpha}_{i,t} = \hat{\alpha}_{i,t}^{bvb} - \hat{\alpha}_{i,t}$ when $I_{i,t} = 1$ and let $\Delta\hat{\alpha}_i$ be the $T_i \times 1$ vector of such differences for fund i at dates with $I_{i,t} = 1$. We then select a $p \times 1$ vector of variables $w_{i,t-1}$ and regress it onto $x_{i,t}$ to obtain the residuals $\tilde{w}_{i,t-1}$. Let us consider the auxiliary time-series regression:

$$\Delta\hat{\alpha}_{i,t} = \tilde{w}_{i,t-1}'d_i + e_{i,t}, \tag{A69}$$

where d_i is the parameter vector and $e_{i,t}$ is the error term. The R^2 of this regression is

$$R_i^2 = 1 - \frac{\Delta\hat{\alpha}_i' M_{\tilde{W}_i}\Delta\hat{\alpha}_i}{\Delta\hat{\alpha}_i'\Delta\hat{\alpha}_i} = \frac{\Delta\hat{\alpha}_i' P_{\tilde{W}_i}\Delta\hat{\alpha}_i}{\Delta\hat{\alpha}_i'\Delta\hat{\alpha}_i} \tag{A70}$$

where $P_{\tilde{W}_i} = \tilde{W}_i(\tilde{W}_i'\tilde{W})_i^{-1}\tilde{W}_i' = I_{T_i} - M_{\tilde{W}_i}$ and $\tilde{W}_i$ is the $T_i \times p$ matrix of the available values for $\tilde{w}_{i,t-1}$.[14]

We use the quantity $T_i R_i^2$ as the test statistic for the null hypothesis $H_{0,i}$. We now derive the asymptotic distribution under Equation (A65) or (A66). Suppose first that Equation (A65) holds in the data, i.e., the linear specification with lagged size is correctly specified. Then $\hat{\alpha}_{i,t}^{bvb} = a_i - b_i q_{i,t-1} + \varepsilon_{i,t} - f_t'(\hat{\beta}_i - \beta_i)$ and

$$\Delta\hat{\alpha}_{i,t} = \hat{\alpha}_{i,t}^{bvb} - \hat{\alpha}_{i,t} = \varepsilon_{i,t} - (\hat{a}_i - a_i) + (\hat{b}_i - b_i)q_{i,t-1} - f_t'(\hat{\beta}_i - \beta_i) = \varepsilon_{i,t} - x_{i,t}'(\tilde{\gamma}_i - \gamma_i), \tag{A71}$$

where $\tilde{\gamma}_i = (\hat{a}_i, \hat{b}_i, \hat{\beta}_i')'$ is a consistent estimator of γ_i. Hence, we have in vector notation: $\Delta\hat{\alpha}_i = \varepsilon_i - X_i(\tilde{\gamma}_i - \gamma_i)$, where X_i is the $T_i \times (K_f + 2)$ matrix of the available values for $x_{i,t}$. Using $\tilde{W}_i'X_i = 0$ and assuming conditional homoscedasticity for the error term $\varepsilon_{i,t}$, we obtain:

$$T_i R_i^2 = \frac{(\frac{1}{\sqrt{T_i}}\varepsilon_i'\tilde{W}_i)(\frac{1}{T_i}\tilde{W}_i'\tilde{W}_i)^{-1}(\frac{1}{\sqrt{T_i}}\tilde{W}_i'\varepsilon_i)}{\frac{1}{T_i}\Delta\hat{\alpha}_i'\Delta\hat{\alpha}_i} \Rightarrow \chi^2(p), \tag{A72}$$

which holds because we have $plim\frac{1}{T_i}\Delta\hat{\alpha}_i'\Delta\hat{\alpha}_i = plim\frac{1}{T_i}\varepsilon_i'\varepsilon_i = \sigma_i^2$, $\frac{1}{\sqrt{T_i}}\tilde{W}_i'\varepsilon_i \Rightarrow N(0, \sigma_i^2 Q_{\tilde{W}_i})$ and $Q_{\tilde{W}_i} = plim\frac{1}{T_i}\tilde{W}_i'\tilde{W}_i = plim\frac{1}{T_i}\sum_t I_{i,t}\tilde{w}_{i,t-1}\tilde{w}_{i,t-1}'$.

Suppose now that the linear model is misspecified and data are generated according to the model in Equation (A66). Consider the linear projection of $p_{i,t-1}$ onto the constant and $q_{i,t-1}$, with residual $\tilde{p}_{i,t-1}$, and let $c_i'p_{i,t-1} = \xi_i + \rho_i q_{i,t-1} + c_i'\tilde{p}_{i,t-1}$ (by our assumption, $\tilde{p}_{i,t-1}$ is also the residual in the regression of $p_{i,t-1}$ onto $x_{i,t}$). By plugging into Equation

[14] Note that Equation (A67) does not include the constant, and the R^2 is defined accordingly. Including a constant and modifying the definition of R^2 does not change the behaviour of the test statistic under the null hypothesis and its consistency under the alternative.

(A66), we get $r_{i,t} = a_i^* - b_i^* q_{i,t-1} + c_i' \tilde{p}_{i,t-1} + \beta_i' f_t + \varepsilon_{i,t}$, with pseudo-true parameter values $a_i^* = a_i + \xi_i$ and $b_i^* = b_i - \rho_i$. Then, we have

$$\hat{\alpha}_{i,t}^{bvb} = a_i^* - b_i^* q_{i,t-1} + c_i' \tilde{p}_{i,t-1} + \varepsilon_{i,t} - f_t'(\hat{\beta}_i - \beta_i), \tag{A73}$$

$$\hat{\alpha}_{i,t} = \hat{a}_i^l - \hat{b}_i^l q_{i,t-1}, \tag{A74}$$

where $\hat{a}_i^l = e_1'(X_i'X_i)^{-1}X_i'(X_i\gamma_i^* + \varepsilon_i + \tilde{P}_i c_i) = a_i^* + e_1'(X_i'X_i)^{-1}X_i'(\varepsilon_i + \tilde{P}_i c_i)$, $\hat{b}_i^l = b_i^* + e_2'(X_i'X_i)^{-1}X_i'(\varepsilon_i + \tilde{P}_i c_i)$, $\gamma_i^* = (a_i^*, b_i^*, \beta_i')'$, and $\tilde{P}_i$ is the matrix of the observations of the variables $\tilde{p}_{i,t-1}$ when $I_{i,t} = 1$. Combining Equations (A73)-(A74), we have

$$\Delta\hat{\alpha}_{i,t} = \varepsilon_{i,t} - (\hat{a}_i^l - a_i^*) + (\hat{b}_i^l - b_i^*)q_{i,t-1} - f_t'(\hat{\beta}_i - \beta_i) + c_i'\tilde{p}_{i,t-1}, \tag{A75}$$

or, in vector notation, $\hat{\Delta}\alpha_i = \varepsilon_i - X_i(\tilde{\gamma}_i - \gamma_i^*) + \tilde{P}_i c_i$, where $\tilde{\gamma}_i$ is a consistent estimator of γ_i^*. Then, we have:

$$plim\ \frac{1}{T_i}\tilde{W}_i'\Delta\hat{\alpha}_i = plim\ \frac{1}{T_i}\tilde{W}_i'\varepsilon_i + (plim\ \frac{1}{T_i}\tilde{W}_i'\tilde{P}_i)c_i = Q_{\tilde{W}_i}\Lambda_i c_i, \tag{A76}$$

where $\Lambda_i = plim\ (\frac{1}{T_i}\tilde{W}_i'\tilde{W}_i)^{-1}\frac{1}{T_i}\tilde{W}_i'\tilde{P}_i$ is the regression coefficient matrix of $\tilde{p}_{i,t-1}$ onto $\tilde{w}_{i,t-1}$, i.e., Λ_i is the coefficient vector associated with $z_{i,t-1}$ in a regression of $p_{i,t-1}$ onto $x_{i,t}$ and $w_{i,t-1}$: $p_{i,t-1} = \delta_i' x_{i,t} + \Lambda_i' w_{i,t-1} + \eta_{i,t}$, where $\eta_{i,t}$ is the regression error with variance V_{η_i}. By using $plim\ \tilde{\gamma}_i - \gamma_i^* = 0$, we have

$$\begin{aligned} plim\ \frac{1}{T_i}\Delta\hat{\alpha}_i'\Delta\hat{\alpha}_i &= plim \frac{1}{T_i}[\varepsilon_i + \tilde{P}_i c_i]'[\varepsilon_i + \tilde{P}_i c_i] = \sigma_i^2 + c_i'(plim\ \frac{1}{T_i}\tilde{P}_i'\tilde{P}_i)c_i \\ &= \sigma_i^2 + c_i'(\Lambda_i' Q_{\tilde{W}_i}\Lambda_i + V_{\eta_i})c_i, \end{aligned} \tag{A77}$$

which implies that

$$plim\ R_i^2 = \frac{c_i'(\Lambda_i' Q_{\tilde{W}_i}\Lambda_i)c_i}{\sigma_i^2 + c_i'(\Lambda_i' Q_{\tilde{W}_i}\Lambda_i + V_{\eta_i})c_i}. \tag{A78}$$

Based on Equation (A78), we know that the test based on $T_i R_i^2$ is consistent, namely $T_i R_i^2$ diverges in large samples and the power of the test approaches one asymptotically under misspecification. As long as $\Lambda_i c_i \neq 0$, the vector $w_{i,t-1}$ captures the effect of the omitted component $p_{i,t-1}' c_i$ and is therefore informative about the source of misspecification of the linear specification $\alpha_{i,t} = a_i - b_i q_{i,t-1}$. If the omitted variables $p_{i,t-1}$ coincide with the chosen $w_{i,t-1}$ (i.e., $p_{i,t-1} = w_{i,t-1}$), we automatically obtain that $\Lambda_i c_i \neq 0$ if $|corr(p_{i,t-1}' c_i, q_{i,t-1})| < 1$, and the consistency of the test follows.

Applying this theoretical framework, we consider two sets of variables for $w_{i,t}$. First, we include the ratio of industry size to total market capitalization to capture changes in industry competition, and aggregate turnover to capture changes in the level of mispricing in capital markets (see Pastor, Stambaugh, and Taylor (2015, 2018)). Second, we include higher order terms of fund size ($q_{i,t-1}^2$ and $q_{i,t-1}^3$) to capture nonlinearities in the relationship between the gross alpha and fund size. For each fund, we then test the null hypothesis $H_{0,i}$ that the linear specification $\alpha_{i,t} = a_i - b_i q_{i,t-1}$ is correct.

For the first set of variables, we reject the null hypothesis only 13.0% of the times at the 5%-significance level. In other words, $T_i R_i^2$ is larger than the 95%-quantile of the $\chi^2(p)$ distribution for only 12.7% of the funds. Furthermore, we find that 29.3% of these funds can be classified as false discoveries ($H_{0,i}$ is rejected whereas it is true) using the approach proposed by Barras, Scaillet, and Wermers (2010). Turning to the analysis of the second set of variables, we obtain similar results—we reject $H_{0,i}$ for 14.1% of the funds (at the 5% level), among which more than 26.8% are false discoveries.

B Validity of the Panel Approach

In this section, we formally test whether the panel approach that imposes a constant scale coefficient across funds ($b_i = b$) is consistent with the data. To this end, we use the test of slope homogeneity developed by Pesaran and Yamagata (2008) for large panels. The null hypothesis is H_0: $b_i = b$ for $i = 1, ..., n$ against the alternative hypothesis H_1: $b_i \neq b_j$ for a non-zero fraction of pairwise slopes for $i \neq j$.

We denote by r_i and q_i the T_i-vectors of the fund gross excess returns and lagged fund sizes and by Z_i the $T_i \times (K_f + 1)$ matrix of available values for $x_t = (1, f_t')'$. The idea of the test is to investigate the dispersion of individual slope estimates from a suitable pooled estimate. We define the weighted sum of squared deviations:

$$\hat{S} = \sum_i (\hat{b}_i - \hat{b}_{WFE})^2 \frac{q_i' M_i q_i}{\hat{\sigma}_i^2}, \tag{A79}$$

where $M_i = I_{T_i} - Z_i(Z_i'Z_i)^{-1}Z_i'$ is the projection matrix, I_{T_i} is the $T_i \times T_i$ identity matrix, $\hat{b}_i = (q_i' M_i q_i)^{-1} q_i' M_i r_i$ is the estimated scale coefficient of each fund, $\hat{b}_{WFE} = \left(\sum_i \frac{q_i' M_i q_i}{\hat{\sigma}_i^2}\right)^{-1} \left(\sum_i \frac{q_i' M_i r_i}{\hat{\sigma}_i^2}\right)$ is the weighted fixed effect pooled estimate, $\hat{\sigma}_i^2$ is the variance estimate defined as $\frac{(r_i - \hat{b}_{FE} q_i)' M_i (r_i - \hat{b}_{FE} q_i)}{T_i - K - 1}$, and $\hat{b}_{FE} = (\sum_i q_i' M_i q_i)^{-1} \sum_i q_i' M_i r_i$ is the standard fixed effect pooled estimate. Pesaran and Yamagata (2008) show that under the null hypothesis H_0 the test statistic

$$\hat{\Delta} = \sqrt{n}\left(\frac{\frac{1}{n}\hat{S} - 1}{\sqrt{2}}\right) \tag{A80}$$

is asymptotically distributed as a standard Gaussian random variable when $n, T \to \infty$ such that $\sqrt{n}/T_{\min}^2 \to 0$ with $T_{\min} = \min_{1 \leqslant i \leqslant n} T_i$. Therefore, we can build the chi-square test statistic $\hat{\Delta}^2$ which is asymptotically distributed as a chi-square random variable χ_1^2 with one degree of freedom.[15]

We examine two specifications for the panel regression: (i) the linear specification $\alpha_{i,t} = a_i - bq_{i,t-1}$, and (ii) the log specification $\alpha_{i,t} = a_i - b\log(q_{i,t-1})$.[16] We also conduct the test in the entire population and within each group (small/large cap, low/high turnover, broker/direct sold). Examining each fund group separately allows us to determine whether grouping funds into well-defined categories absorbs the heterogeneity. Our results reveal that the test of homogeneity is always strongly rejected, i.e., for each specification (size, log size), we reject H_0 with probability one. Furthermore, the null hypothesis of homogeneous coefficients is also rejected with probability one in each fund group. Therefore, forming groups is not sufficient to absorb the large heterogeneity in a_i and b_i.

C Survivorship and Reverse Survivorship Bias

In this section, we examine the impact of the survivorship and reverse survivorship bias. Our empirical analysis does not require that the funds remain alive until the end of the sample in 2019. In other words, our original sample includes all both living and dead funds. However, our fund selection rule requires that each fund has a minimum of 60 monthly observations ($T_{\min} = 60$) to be included in our final sample. Our results could therefore be subject to a survivorship bias if unskilled funds (a_i<0) disappear early. To examine this issue, we repeat our analysis across different thresholds for $T_{\min}$ ranging from 12 to 60. Panel A of Table AVI shows that our main results are not driven by the survivorship bias—the skill distribution remains largely unchanged as $T_{\min}$ changes from 60 to 12.

[15] The requirement on the relative rate between n and T_i, namely $n = o(T_{\min}^4)$ for the asymptotic validity of the testing procedure is weak and matches the time-series and cross-sectional sample sizes in our application since $n_\chi = 2,321$ is much smaller that $T_{\min}^4$ =60^4 = 12,960,000.

[16] In the logarithmic specification, the intercept loses its interpretation as a first-dollar alpha (i.e, it corresponds to the alpha when $q_{i,t-1} = 1$ instead of 0). In addition, the intercept depends on the measurement unit (e.g, \$1 or \$1M). The invariance to size denomination is an advantage of the linear specification.

It is a priori tempting to choose $T_{\min} = 12$ (instead of 60) to mitigate the survivorship bias and offer an improved estimation of the skill distribution. However, reducing $T_{\min}$ may not be optimal because it potentially increases the severity of the reverse survivorship bias (i.e., the reported skill could be biased downwards). The reverse survivorship bias arises because some skilled funds (a_i>0) may perform unexpectedly poorly and disappear early. For these funds, the estimated skill is lower than the true level because it is computed based on unusually low return observations (Linnainmaa (2013)). By reducing $T_{\min}$, we increase the likelihood of including these funds in the sample.

To examine this issue, we compare the skill distributions among the disappearing funds for $T_{\min} = 60$ and 108. The assumption is that unskilled funds tend to disappear early (during the first five years). In this case, the difference between the two distributions captures the impact of the reverse survivorship bias, i.e., it should decrease with $T_{\min}$ as we exclude a larger number of skilled funds that disappear after unexpected poor performance. Panel B shows that the difference in the proportion of skilled funds equals 4.1% as $T_{\min}$ decreases from 108 to 60. This number represents 85% of the proportion difference when reducing $T_{\min}$ from 60 to 12 in Panel A. This back-of-the-envelope calculation suggests that the reduction in skill observed for $T_{\min} = 12$ is mainly due to the reverse survivorship bias. Motivated by these results, we therefore choose $T_{\min} = 60$ in our baseline analysis.

Please insert Table AVI here

D Alternative Asset Pricing Models

Our empirical results potentially depend on the choice of the asset pricing model. To address this issue, we repeat our analysis using the four-factor model of Carhart (1997) which contains the same factors as the model of Cremers, Petajisto, and Zitzewitz (2012) except that the market, size, and value factors are not computed from tradable indices. We also consider the five-factor model of Fama and French (2015) which includes the market, size, value, profitability, and investment factors.[17]

Table AVII shows that the skill and scalability distributions remain qualitatively unchanged with the Carhart model. The skill coefficient is equal to 2.4% per year on average, and is positive for 78.5% of the funds (vs 3.0% and 83.1% for the baseline results). The scale coefficient is, on average, equal to 1.3% per year and 80.8% of the funds face diseconomies of scale (vs 1.3% and 82.4% for the baseline results). We observe the

[17]The size and value factors in the five-factor model are similar to ones used in the Carhart model.

main difference for the small cap group in which the skill coefficient drops from 4.6% to 3.3% on average. This sharp reduction arises because the Carhart model assigns a negative alpha to the Russell 2000 index and therefore penalizes the performance of small cap funds (consistent with the analysis by Cremers, Petajisto, and Zitzewitz (2012)). Next, Table AVIII reports the results obtained with the five-factor Fama-French model. Under this model, the scalability distribution remains largely unchanged but the proportion of funds with positive skill decreases from 83.1% to 74.0%. This reduction suggests that some funds achieve positive returns partly because they implement profitability- and investment-based strategies.

Please insert Tables AVII and AVIII here

E Analysis based on Daily Fund Returns

Our baseline specification $\alpha_{i,t} = a_i - b_i q_{i,t-1}$ assumes that the skill and scale coefficients remain constant over time. To examine the stability of these coefficients, we conduct an extensive analysis using daily fund returns. This procedure allows us to capture potential changes in the coefficients without explicitly modeling their dynamics (see Lewellen and Nagel (2006)).

To conduct this analysis, we use the daily fund return CRSP database available between January 1999 and December 2019. The CRSP database only reports the daily net return and Net Asset Value (NAV) of each shareclass, but not its daily size. To address this issue, we compute the number of shares for each shareclass at the start of the month. We can then compute the daily size of each shareclass within the month as the product between its daily NAV and the number of shares. We match the fund identifier across the daily and monthly databases to maintain our selection of open-end, actively managed funds with a well-defined equity style. We measure the daily gross return of each fund as the sum of the daily net return and fees. The daily net return is computed as the value-weighted average of the daily net returns across all shareclasses. The daily fees are defined as the value-weighted average of the most recently reported annual fees across shareclasses divided by 21·12.

We can summarize our estimation procedure in two steps. First, we run the following time-serie regression for each year a ($a = 1999, ..., 2019$):

$$r_{i,t} = \alpha_{i,a} + \beta'_{i,a} f_t + e_{i,t}, \tag{A81}$$

where $r_{i,t}$ is the fund daily gross excess return and f_t is the vector of daily factor excess returns in the model of Cremers, Petajisto, and Zitzewitz (2012). Using Equation (A81), we can extract the daily gross alpha of the fund after controling for short-term variations in factor loadings (i.e., $\beta_{i,a}$ is allowed to change on an annual basis):

$$\alpha_{i,t} = \alpha_{i,a} + e_{i,t}. \tag{A82}$$

Second, we run a regression of the daily gross alpha on lagged size to infer the time-varying skill and scale coefficients. Given the potential persistence over a small window of only one year, we estimate the regression over a non-overlapping window τ of five years:

$$\alpha_{i,t} = a_{i,\tau} - b_{i,\tau} q_{i,t-1} + \varepsilon_{i,t}, \tag{A83}$$

where each of the windows covers the years 1999-2004, 2005-2009, 2010-2014, and 2015-2019 (i.e., $\tau = 1, ..., 4$).

To examine the stability of the skill coefficient, we take the first window $\tau = 1$ as the benchmark and test the null hypothesis of constant skill $H_{0,i,\tau} : \Delta a_{i,\tau} = a_{i,1} - a_{i,\tau} = 0$ (for $\tau = 2, 3{,}4$). Overall, there is little evidence of time-variation in the skill coefficient. Using a 5%-significance threshold, we find that $H_{0,i,\tau}$ is only rejected for (i) 10.7% of the funds for $\tau = 2$, (ii) 9.6% of the funds for $\tau = 3$, and (iii) 14.1% of the funds for $\tau = 4$. We also uncover a substantial fraction of false discoveries among these funds ($H_{0,i,\tau}$ is rejected whereas it is true)—this fraction ranges between 30.1% and 45.5% across the three windows ($\tau = 2, 3{,}4$) using the approach proposed by Barras, Scaillet, and Wermers (2010).

Repeating this analysis for the scale coefficient, we test the null hypothesis of constant scale $H_{0,i,\tau} : \Delta b_{i,\tau} = b_{i,1} - b_{i,\tau} = 0$ (for $\tau = 2, 3{,}4$). The results are similar to those obtained for the skill coefficient. Using a 5%-significance threshold, we find that $H_{0,i,\tau}$ is only rejected for (i) 9.0% of the funds for $\tau = 2$, (ii) 7.8% of the funds for $\tau = 3$, and (iii) 10.3% of the funds for $\tau = 4$. Among the rejected funds, the proportion of false discoveries ($H_{0,i,\tau}$ is rejected whereas it is true) ranges between 41.2% and 56.1%.

We also find a remarkable similarity between the skill and scalability distributions measured at the daily and monthly frequencies. Specifically, we measure the annual fund skill and scale levels from daily data as $a_i^d = (\frac{1}{\tau}\sum_\tau a_{i,\tau})21 \cdot 12$ and $b_i^d = (\frac{1}{\tau}\sum_\tau b_{i,\tau})21 \cdot 12$, and examine the characteristics of the two cross-sectional distributions. We then conduct our baseline monthly analysis over the same period as the one covered by the CRSP daily database (1999-2019). The daily analysis reveals that 82.4% of the funds have a positive skill coefficient which, on average, equals 4.7% per year (vs 77.8% and 3.6% for the monthly analysis). For the scale coefficient, these numbers obtained at the daily frequency are equal to 79.5% and 1.3% per year (vs 76.8% and 1.4% for the monthly analysis).

F Impact of Changes in Economic Conditions

We now extend our baseline specification to capture the impact of changes in economic conditions. We consider two alternative specifications motivated by the recent mutual fund literature. First, we examine whether the gross alpha changes with the level of industry competition using

$$\alpha_{i,t} = a_i - a_{i,sw} sw_{t-1} - b_i q_{i,t-1}, \tag{A84}$$

where sw_{t-1} is defined as the demeaned ratio of industry size to total market capitalization (as in Pastor, Stambaugh, and Taylor (2015)). Second, we account for potential changes in aggregate mispricing using an extended version of Equation (A84)

$$\alpha_{i,t} = a_i - a_{i,sw} sw_{t-1} + a_{i,tu} tu_{t-1} - b_i q_{i,t-1}, \tag{A85}$$

where tu_{t-1} is defined as the demeaned aggregate turnover across all funds (as in Pastor, Stambaugh, and Taylor (2018)). Under both specifications, we can still interpret a_i as the alpha on the first dollar when industry competition and aggregate mispricing are equal to their average levels (i.e., $sw_{t-1} = tu_{t-1} = 0$).

The results in Table AIX show that adding the industry variable sw_{t-1} leaves the skill and scalability distributions largely unchanged. For instance, we find that 82.3% and 82.6% of the funds exhibit positive skill and scale coefficients (vs 83.1% and 82.4% for the baseline results). When the variable sw_{t-1} is used alone in the regression (i.e., $\alpha_{i,t} = a_i - a_{i,sw} sw_{t-1}$), the majority of the funds respond negatively to an increase in industry size (51.0% of the funds have a positive coefficient $a_{i,sw}$). However, this result is overturned when we include lagged size, i.e., only 46.6% of the funds have a positive coefficient $a_{i,sw}$.[18] One possible explanation for this result is that sw_{t-1} may not capture changes in industry competition with sufficient granularity (see Hoberg, Kumar, and Prabhala (2020) for a discussion).

Table AX also shows that the empirical evidence on skill and scalability remains largely unchanged under the extended model in Equation (A85). In this case, the average levels of the skill and scale coefficients are equal to 3.4% and 1.6% per year (vs 3.0% and 1.3% for the baseline results). In addition, the proportions of funds with

[18] Therefore, this result departs from the evidence in Pastor, Stambaugh, and Taylor (2015) obtained with a panel approach in which fund scale is assumed to be constant ($b_i = b$).

positive skill and scale coefficients equal 80.8% and 80.6% (vs 83.2% and 82.4% for the baseline results). Consistent with Pastor, Stambaugh, and Taylor (2018), we find that the majority of funds produce higher returns in times of higher mispricing in capital markets. The proportion of funds with a positive coefficient $a_{i,tu}$ is equal to 60.8%.

Please insert Tables AIX to AX here

References

[1] Amihud, Y., and C. M. Hurvich, 2004, Predictive Regressions: A Reduced-Bias Estimation Method, Journal of Financial and Quantitative Analysis, 813-841.

[2] Avramov D., L. Barras, and R. Kosowski, 2013, Hedge Fund Return Predictability under the Magnifying Glass, Journal of Quantitative and Financial Analysis 48, 1057-1083.

[3] Barras, L., O. Scaillet, and R. Wermers, 2010, False Discoveries in Mutual Fund Performance: Measuring Luck in Estimated Alphas, Journal of Finance 65, 179-216.

[4] Bai, J. and S. Ng, 2005, Tests for Skewness, Kurtosis, and Normality for Time Series Data, Journal of Business & Economic Statistics 23, 49-60.

[5] Berk, J. B., and J. van Binsbergen, 2015, Measuring Skill in the Mutual Fund Industry, Journal of Financial Economics 118, 1-20.

[6] Berk, J. B., and R. C. Green, 2004, Mutual Fund Flows and Performance in Rational Markets, Journal of Political Economy 112, 1269-1295.

[7] Bhattacharya, P. K., 1967, Estimation of a Probability Density Function and its Derivatives, Sankhya Series A 29, 373-382.

[8] Carhart, M. M., 1997, On Persistence in Mutual Fund Performance, Journal of Finance 52, 57-82.

[9] Chen, J., H. Hong, M. Huang, and J. D. Kubik, 2004, Does Fund Size Erode Mutual Fund Performance? The Role of Liquidity and Organization, American Economic Review 94, 1276-1302.

[10] Cremers, M., A. Petajisto, and E. Zitzewitz, 2012, Should Benchmark Indices Have Alpha? Revisiting Performance Evaluation, Critical Finance Review 2, 1–48.

[11] Del Guercio, D., and J. Reuter, 2014, Mutual Fund Performance and the Incentive to Generate Alpha, Journal of Finance 69,1674-1704.

[12] Fama, E. F., and K. R. French, 2015, A Five-Factor Asset Pricing Model, Journal of Financial Economics 116, 1-22.

[13] Gagliardini, P., and C. Gouriéroux, 2013, Granularity Adjustment for Risk Measures: Systematic vs Unsystematic Risks, International Journal of Approximate Reasoning 54, 717-747.

[14] Gagliardini, P., and C. Gouriéroux, 2014, Granularity Theory with Application to Finance and Insurance, Cambridge University Press.

[15] Gagliardini, P., E. Ossola, and O. Scaillet, 2016, Time-varying Risk Premium in Large Cross-Sectional Equity Datasets, Econometrica 84, 985-1056.

[16] Gourieroux, C., J.-P. Laurent, and O. Scaillet, 2000, Sensitivity Analysis of Value-at-Risk, Journal of Empirical Finance 7, 225-245.

[17] Hausman, J. A., 1978, Specification Tests in Econometrics, Econometrica 46, 1251–1271.

[18] Hoberg G., N. Kumar, and N. Prabhala, 2018, Mutual Fund Competition, Managerial Skill, and Alpha Persistence, Review of Financial Studies 31, 1896–1929.

[19] Jochmans, K., and M. Weidner, 2018, Inference on a Distribution with Noisy Draws, Working Paper.

[20] Lewellen, J., Nagel, S., 2006, The Conditional CAPM does not Explain Asset-Pricing Anomalies. Journal of Financial Economics 82, 289–314 .

[21] Linnainmaa, J., 2013, Reverse Survivorship Bias, Journal of Finance 68, 789-813.

[22] Martin, R., and T. Wilde, 2002, Unsystematic Credit Risk, Risk, 15, 123-128.

[23] Okui, R., and T. Yanagi, 2020, Kernel Estimation for Panel Data with Heterogeneous Dynamics, The Econometrics Journal 23, 156-175.

[24] Okui, R., and T. Yanagi, 2019, Panel Data Analysis with Heterogeneous Dynamics, Journal of Econometrics 212, 451-475.

[25] Pastor, L., R. F. Stambaugh, and L. Taylor, 2015, Scale and Skill in Active Management, Journal of Financial Economics 116, 23-45.

[26] Pastor, L., R. F. Stambaugh, and L. Taylor, 2018, Do Funds Make More When They Trade More?, Journal of Finance 72, 1483-1528.

[27] Pesaran, M. H., and T. Yagamata, 2008, Testing Slope Homogeneity in Large Panels, Journal of Econometrics 142, 50–93.

[28] Silverman, B. W., 1986, Density Estimation for Statistics and Data Analysis, Chapman and Hall, London.

[29] Stambaugh, R. F., 1999, Predictive Regressions, Journal of Financial Economics 54, 375–421.

[30] Sun, Y., 2020, Index Fund Entry and Financial Product Market Competition, Management Science, Forthcoming.

[31] Wilde, T., 2001, Probing Granularity, Risk 14, 103-106.

Table AI
Properties of the Estimated Density:
Skill Coefficient

Panel A shows the Mean Integrated Squared Error (MISE) and its two components (integrated squared bias and variance) for the bias-adjusted skill density under the baseline choice for the optimal bandwidth across different values for the number of funds and the number of monthly observations. Panel B repeats the analysis under the alternative choice of the optimal bandwidth. For comparison, Panel C reports the same information for the bias-unadjusted density.

Panel A: Bias Adjustment (Baseline Choice for Optimal Bandwidth)

MISE

n\T	100	250	500	750	1000
1000	20.50	5.24	1.91	1.16	0.83
2500	20.62	4.96	1.59	0.84	0.60
5000	20.76	4.82	1.39	0.69	0.45
7500	20.14	4.73	1.35	0.64	0.40
10000	20.40	4.65	1.29	0.57	0.37

Bias^2

n\T	100	250	500	750	1000
1000	19.92	4.50	1.26	0.56	0.29
2500	20.32	4.59	1.24	0.54	0.32
5000	20.58	4.59	1.18	0.50	0.27
7500	20.00	4.55	1.18	0.50	0.26
10000	20.29	4.51	1.16	0.46	0.26

Variance

n\T	100	250	500	750	1000
1000	0.58	0.74	0.65	0.60	0.53
2500	0.30	0.37	0.36	0.30	0.29
5000	0.18	0.23	0.21	0.19	0.18
7500	0.14	0.18	0.17	0.14	0.14
10000	0.11	0.14	0.13	0.11	0.11

Panel B: Bias Adjustment (Alternative Choice for Optimal Bandwidth)

MISE

n\T	100	250	500	750	1000
1000	21.76	5.66	2.06	1.24	0.83
2500	21.58	5.33	1.75	0.91	0.62
5000	21.52	5.13	1.55	0.76	0.47
7500	20.85	5.01	1.48	0.70	0.41
10000	21.02	4.92	1.42	0.63	0.39

Bias^2

n\T	100	250	500	750	1000
1000	21.53	5.33	1.74	0.90	0.52
2500	21.46	5.18	1.59	0.77	0.47
5000	21.46	5.05	1.46	0.67	0.38
7500	20.80	4.95	1.42	0.64	0.34
10000	20.99	4.87	1.37	0.58	0.33

Variance

n\T	100	250	500	750	1000
1000	0.23	0.33	0.32	0.34	0.32
2500	0.12	0.14	0.15	0.14	0.15
5000	0.06	0.08	0.09	0.09	0.09
7500	0.05	0.06	0.07	0.06	0.07
10000	0.03	0.05	0.05	0.05	0.05

Panel C: No Bias Adjustment

MISE

n\T	100	250	500	750	1000
1000	29.59	9.80	3.64	2.17	1.38
2500	28.87	9.06	3.08	1.56	1.01
5000	28.38	8.61	2.73	1.28	0.75
7500	27.68	8.31	2.55	1.15	0.63
10000	27.76	8.18	2.43	1.04	0.58

Bias^2

n\T	100	250	500	750	1000
1000	29.46	9.59	3.41	1.91	1.13
2500	28.80	8.96	2.96	1.44	0.89
5000	28.34	8.54	2.66	1.21	0.68
7500	27.65	8.26	2.50	1.10	0.57
10000	27.73	8.14	2.38	1.00	0.53

Variance

n\T	100	250	500	750	1000
1000	0.13	0.21	0.23	0.26	0.25
2500	0.07	0.10	0.12	0.11	0.12
5000	0.04	0.06	0.07	0.07	0.08
7500	0.03	0.05	0.06	0.05	0.06
10000	0.03	0.04	0.04	0.04	0.04

Table AII
Properties of the Estimated Moments and Proportion: Skill Coefficient

Panel A shows the Mean Squared Error (MSE) and its two components (bias and standard deviation) of the bias-adjusted estimators (mean and volatility (annualized), skewness, and proportion of funds with a positive skill measure) based on a numerical integration of the bias-adjusted density (under the baseline bandwidth choice) across different values for the number of funds and the number of monthly observations. Panel B reports the same information for the bias-adjusted estimators obtained with the analytical approach. For comparison, Panel C reports the same information for the bias-unadjusted estimators (obtained by integrating the bias-unadjusted density).

Panel A: Bias Adjustment (Numerical Integration)

Mean

MSE

n\T	100	250	500	750	1000
1000	1.37	0.21	0.07	0.04	0.03
2500	1.31	0.21	0.06	0.03	0.02
5000	1.35	0.20	0.05	0.02	0.02
7500	1.28	0.21	0.05	0.02	0.02
10000	1.30	0.20	0.05	0.02	0.01

Bias

n\T	100	250	500	750	1000
1000	1.15	0.43	0.23	0.14	0.12
2000	1.14	0.45	0.23	0.16	0.12
3000	1.16	0.44	0.22	0.15	0.11
4000	1.13	0.45	0.22	0.15	0.11
5000	1.14	0.44	0.22	0.15	0.11

Standard Deviation

n\T	100	250	500	750	1000
1000	0.22	0.15	0.12	0.12	0.12
2000	0.14	0.08	0.08	0.07	0.07
3000	0.09	0.06	0.05	0.05	0.05
4000	0.08	0.05	0.05	0.04	0.05
5000	0.06	0.04	0.04	0.04	0.04

Volatility

MSE

n\T	100	250	500	750	1000
1000	1.93	0.07	0.03	0.03	0.03
2500	1.88	0.03	0.01	0.01	0.01
5000	1.77	0.02	0.01	0.00	0.00
7500	1.66	0.02	0.00	0.00	0.00
10000	1.65	0.02	0.00	0.00	0.00

Bias

n\T	100	250	500	750	1000
1000	1.36	0.18	0.03	0.01	-0.00
2500	1.36	0.15	0.02	0.01	0.01
5000	1.32	0.13	0.03	0.01	0.01
7500	1.28	0.13	0.02	0.02	0.01
10000	1.28	0.13	0.02	0.01	0.01

Standard Deviation

n\T	100	250	500	750	1000
1000	0.28	0.19	0.16	0.16	0.16
2500	0.18	0.10	0.09	0.08	0.08
5000	0.13	0.08	0.06	0.06	0.06
7500	0.11	0.06	0.06	0.06	0.06
10000	0.08	0.07	0.05	0.05	0.05

Skewness

MSE

n\T	100	250	500	750	1000
1000	0.29	0.13	0.11	0.10	0.09
2500	0.18	0.06	0.05	0.04	0.04
5000	0.09	0.03	0.03	0.03	0.02
7500	0.06	0.03	0.02	0.02	0.02
10000	0.06	0.02	0.02	0.02	0.02

Bias

n\T	100	250	500	750	1000
1000	-0.01	-0.01	0.03	0.01	0.01
2000	0.14	0.01	0.03	0.04	0.04
3000	0.16	0.03	0.03	0.02	0.01
4000	0.11	0.02	0.04	0.02	0.02
5000	0.13	0.04	0.04	0.03	0.01

Standard Deviation

n\T	100	250	500	750	1000
1000	0.54	0.37	0.32	0.31	0.30
2000	0.41	0.24	0.22	0.20	0.20
3000	0.26	0.18	0.17	0.16	0.16
4000	0.23	0.17	0.15	0.14	0.14
5000	0.21	0.14	0.12	0.13	0.13

Proportion with Positive Skill Measure

MSE

n\T	100	250	500	750	1000
1000	21.85	4.06	1.46	1.47	0.95
2500	21.47	2.38	0.72	0.50	0.47
5000	19.89	1.94	0.49	0.32	0.29
7500	19.17	1.53	0.39	0.22	0.19
10000	19.23	1.59	0.28	0.16	0.13

Bias

n\T	100	250	500	750	1000
1000	-4.40	-1.39	-0.48	-0.49	-0.13
2500	-4.53	-1.26	-0.44	-0.22	-0.18
5000	-4.41	-1.25	-0.48	-0.23	-0.17
7500	-4.34	-1.12	-0.45	-0.23	-0.12
10000	-4.36	-1.19	-0.35	-0.20	-0.11

Standard Deviation

n\T	100	250	500	750	1000
1000	1.57	1.46	1.11	1.11	0.97
2500	0.99	0.89	0.72	0.67	0.66
5000	0.68	0.63	0.50	0.52	0.51
7500	0.60	0.53	0.43	0.41	0.42
10000	0.49	0.42	0.40	0.34	0.35

Table AII
Properties of the Estimated Moments and Proportion: Skill Coefficient (Continued)

Panel B: Bias Adjustment (Analytical Approach)

Mean

MSE

n\T	100	250	500	750	1000
1000	1.37	0.21	0.07	0.04	0.03
2500	1.31	0.21	0.06	0.03	0.02
5000	1.35	0.20	0.05	0.02	0.02
7500	1.28	0.21	0.05	0.02	0.02
10000	1.30	0.20	0.05	0.02	0.01

Bias

n\T	100	250	500	750	1000
1000	1.15	0.43	0.23	0.14	0.12
2500	1.14	0.45	0.23	0.16	0.12
5000	1.16	0.44	0.22	0.15	0.11
7500	1.13	0.45	0.22	0.15	0.11
10000	1.14	0.44	0.22	0.15	0.11

Standard Deviation

n\T	100	250	500	750	1000
1000	0.22	0.15	0.12	0.12	0.12
2500	0.14	0.08	0.08	0.07	0.07
5000	0.09	0.06	0.05	0.05	0.05
7500	0.08	0.05	0.05	0.04	0.05
10000	0.06	0.04	0.04	0.04	0.04

Volatility

MSE

n\T	100	250	500	750	1000
1000	1.14	0.06	0.03	0.03	0.03
2500	1.22	0.04	0.01	0.01	0.01
5000	1.15	0.03	0.01	0.00	0.00
7500	1.07	0.03	0.01	0.00	0.00
10000	1.09	0.03	0.00	0.00	0.00

Bias

n\T	100	250	500	750	1000
1000	1.02	0.17	0.04	0.02	0.01
2500	1.09	0.17	0.04	0.02	0.02
5000	1.06	0.16	0.05	0.02	0.01
7500	1.03	0.16	0.04	0.03	0.02
10000	1.04	0.17	0.04	0.02	0.02

Standard Deviation

n\T	100	250	500	750	1000
1000	0.31	0.19	0.16	0.16	0.16
2500	0.18	0.10	0.09	0.08	0.08
5000	0.13	0.08	0.06	0.06	0.06
7500	0.11	0.06	0.06	0.06	0.06
10000	0.08	0.07	0.05	0.05	0.05

Skewness

MSE

n\T	100	250	500	750	1000
1000	1.60	0.90	0.42	0.23	0.18
2500	1.63	0.81	0.34	0.20	0.14
5000	1.63	0.78	0.30	0.16	0.11
7500	1.53	0.76	0.30	0.16	0.11
10000	1.51	0.77	0.30	0.16	0.09

Bias

n\T	100	250	500	750	1000
1000	1.19	0.88	0.55	0.37	0.29
2500	1.24	0.88	0.54	0.39	0.31
5000	1.26	0.87	0.53	0.37	0.28
7500	1.23	0.86	0.53	0.37	0.29
10000	1.22	0.87	0.53	0.38	0.28

Standard Deviation

n\T	100	250	500	750	1000
1000	0.42	0.36	0.34	0.31	0.30
2500	0.31	0.21	0.22	0.21	0.21
5000	0.18	0.16	0.16	0.16	0.16
7500	0.16	0.14	0.14	0.14	0.14
10000	0.15	0.12	0.12	0.13	0.12

Proportion with Positive Skill Measure

MSE

n\T	100	250	500	750	1000
1000	14.48	3.49	1.47	1.43	0.93
2500	15.50	1.89	0.82	0.55	0.50
5000	14.14	1.65	0.53	0.33	0.30
7500	13.54	1.26	0.44	0.26	0.19
10000	13.88	1.37	0.34	0.20	0.16

Bias

n\T	100	250	500	750	1000
1000	-3.46	-1.11	-0.37	-0.39	0.05
2500	-3.80	-1.01	-0.45	-0.24	-0.14
5000	-3.70	-1.11	-0.50	-0.24	-0.15
7500	-3.63	-0.99	-0.50	-0.28	-0.14
10000	-3.69	-1.09	-0.42	-0.27	-0.15

Standard Deviation

n\T	100	250	500	750	1000
1000	1.58	1.50	1.15	1.13	0.96
2500	1.04	0.93	0.79	0.70	0.69
5000	0.68	0.65	0.53	0.52	0.53
7500	0.60	0.52	0.44	0.42	0.42
10000	0.49	0.43	0.40	0.35	0.37

Table AII
Properties of the Estimated Moments and Proportion:
Skill Coefficient (Continued)

Panel C: No Bias Adjustment

Mean

MSE

n\T	100	250	500	750	1000
1000	1.37	0.21	0.07	0.04	0.03
2500	1.31	0.21	0.06	0.03	0.02
5000	1.35	0.20	0.05	0.02	0.02
7500	1.28	0.21	0.05	0.02	0.02
10000	1.30	0.20	0.05	0.02	0.01

Bias

n\T	100	250	500	750	1000
1000	1.15	0.43	0.23	0.14	0.12
2500	1.14	0.45	0.23	0.16	0.12
5000	1.16	0.44	0.22	0.15	0.11
7500	1.13	0.45	0.22	0.15	0.11
10000	1.14	0.44	0.22	0.15	0.11

Standard Deviation

n\T	100	250	500	750	1000
1000	0.22	0.15	0.12	0.12	0.12
2500	0.14	0.08	0.08	0.07	0.07
5000	0.09	0.06	0.05	0.05	0.05
7500	0.08	0.05	0.05	0.04	0.05
10000	0.06	0.04	0.04	0.04	0.04

Volatility

MSE

n\T	100	250	500	750	1000
1000	11.00	1.18	0.25	0.12	0.07
2500	11.28	1.17	0.23	0.10	0.06
5000	11.17	1.15	0.24	0.10	0.05
7500	10.99	1.15	0.23	0.10	0.05
10000	11.06	1.16	0.23	0.09	0.05

Bias

n\T	100	250	500	750	1000
1000	3.31	1.07	0.48	0.31	0.22
2500	3.35	1.08	0.47	0.30	0.22
5000	3.34	1.07	0.48	0.30	0.22
7500	3.31	1.07	0.48	0.31	0.22
10000	3.32	1.08	0.48	0.30	0.22

Standard Deviation

n\T	100	250	500	750	1000
1000	0.27	0.16	0.15	0.15	0.15
2500	0.16	0.09	0.08	0.08	0.08
5000	0.12	0.07	0.06	0.06	0.06
7500	0.09	0.06	0.06	0.05	0.05
10000	0.07	0.06	0.05	0.05	0.05

Skewness

MSE

n\T	100	250	500	750	1000
1000	0.10	0.09	0.08	0.08	0.08
2500	0.06	0.06	0.05	0.04	0.04
5000	0.03	0.06	0.04	0.03	0.03
7500	0.04	0.06	0.04	0.03	0.02
10000	0.04	0.05	0.03	0.02	0.02

Bias

n\T	100	250	500	750	1000
1000	-0.19	-0.21	-0.14	-0.12	-0.09
2500	-0.15	-0.21	-0.15	-0.10	-0.07
5000	-0.14	-0.22	-0.15	-0.12	-0.09
7500	-0.17	-0.22	-0.15	-0.11	-0.08
10000	-0.17	-0.21	-0.15	-0.11	-0.10

Standard Deviation

n\T	100	250	500	750	1000
1000	0.25	0.22	0.25	0.26	0.26
2500	0.18	0.13	0.16	0.17	0.18
5000	0.11	0.10	0.12	0.13	0.13
7500	0.09	0.09	0.11	0.11	0.12
10000	0.08	0.08	0.09	0.10	0.11

Proportion with Positive Skill Measure

MSE

n\T	100	250	500	750	1000
1000	115.98	33.21	9.56	5.57	2.44
2500	118.94	31.23	9.49	4.31	2.52
5000	118.33	32.04	9.36	4.10	2.37
7500	117.45	30.67	9.30	4.16	2.23
10000	118.59	31.50	8.85	4.10	2.22

Bias

n\T	100	250	500	750	1000
1000	-10.69	-5.61	-2.90	-2.10	-1.25
2500	-10.87	-5.52	-2.99	-1.96	-1.44
5000	-10.86	-5.63	-3.02	-1.96	-1.45
7500	-10.83	-5.52	-3.02	-2.00	-1.44
10000	-10.88	-5.60	-2.95	-2.00	-1.45

Standard Deviation

n\T	100	250	500	750	1000
1000	1.31	1.31	1.07	1.08	0.93
2500	0.86	0.85	0.75	0.68	0.68
5000	0.54	0.58	0.50	0.50	0.52
7500	0.50	0.48	0.41	0.40	0.40
10000	0.42	0.38	0.38	0.33	0.36

Table AIII
Properties of the Estimated Density:
Scale Coefficient

Panel A shows the Mean Integrated Squared Error (MISE) and its two components (integrated squared bias and variance) for the bias-adjusted scale density under the baseline choice for the optimal bandwidth across different values for the number of funds and the number of monthly observations. Panel B repeats the analysis under the alternative choice of the optimal bandwidth. For comparison, Panel C reports the same information for the bias-unadjusted density.

Panel A: Bias Adjustment (Baseline Choice for Optimal Bandwidth)

MISE

n\T	100	250	500	750	1000
1000	77.11	18.04	5.35	2.95	2.08
2500	75.05	17.01	4.63	2.15	1.55
5000	75.13	17.03	4.09	1.89	1.18
7500	74.82	16.78	4.07	1.79	1.04
10000	74.13	16.61	4.03	1.71	1.02

Bias^2

n\T	100	250	500	750	1000
1000	75.98	16.46	3.95	1.69	0.95
2500	74.51	16.29	4.03	1.64	1.05
5000	74.54	16.66	3.80	1.63	0.89
7500	74.12	16.57	3.87	1.60	0.85
10000	73.90	16.41	3.89	1.56	0.88

Variance

n\T	100	250	500	750	1000
1000	1.13	1.58	1.40	1.26	1.14
2500	0.53	0.72	0.61	0.52	0.50
5000	0.59	0.37	0.29	0.26	0.29
7500	0.70	0.21	0.20	0.19	0.19
10000	0.23	0.19	0.14	0.15	0.14

Panel B: Bias Adjustment (Alternative Choice for Optimal Bandwidth)

MISE

n\T	100	250	500	750	1000
1000	91.92	20.42	5.85	3.16	2.12
2500	87.96	19.27	5.24	2.43	1.68
5000	86.18	18.79	4.63	2.18	1.32
7500	84.71	18.29	4.57	2.05	1.17
10000	83.13	17.98	4.48	1.95	1.15

Bias^2

n\T	100	250	500	750	1000
1000	90.72	19.42	5.14	2.45	1.45
2500	87.38	18.76	4.93	2.17	1.41
5000	85.45	18.55	4.47	2.03	1.16
7500	83.90	18.16	4.44	1.94	1.05
10000	82.88	17.85	4.38	1.85	1.06

Variance

n\T	100	250	500	750	1000
1000	1.21	1.00	0.71	0.71	0.67
2500	0.58	0.51	0.31	0.26	0.27
5000	0.73	0.24	0.17	0.15	0.17
7500	0.80	0.13	0.13	0.12	0.12
10000	0.24	0.13	0.09	0.10	0.09

Panel C: No Bias Adjustment

MISE

n\T	100	250	500	750	1000
1000	97.17	31.16	10.69	5.84	3.60
2500	92.06	28.58	9.41	4.25	2.66
5000	89.26	27.34	8.39	3.80	2.14
7500	87.56	26.59	8.14	3.52	1.86
10000	85.48	26.25	7.87	3.36	1.77

Bias^2

n\T	100	250	500	750	1000
1000	96.81	30.78	10.20	5.32	3.08
2500	91.89	28.39	9.20	4.06	2.44
5000	88.94	27.25	8.27	3.68	2.01
7500	87.13	26.52	8.05	3.43	1.76
10000	85.42	26.20	7.80	3.29	1.70

Variance

n\T	100	250	500	750	1000
1000	0.36	0.38	0.48	0.52	0.51
2500	0.18	0.19	0.22	0.20	0.22
5000	0.32	0.09	0.12	0.12	0.14
7500	0.43	0.07	0.09	0.09	0.10
10000	0.06	0.05	0.07	0.08	0.07

Table AIV
Properties of the Estimated Moments and Proportion: Scale Coefficient

Panel A shows the Mean Squared Error (MSE) and its two components (bias and standard deviation) of the bias-adjusted estimators (mean and volatility (annualized), skewness, and proportion of funds with a positive scale measure) based on a numerical integration of the bias-adjusted density (under the baseline bandwidth choice) across different values for the number of funds and the number of monthly observations. Panel B reports the same information for the bias-adjusted estimators obtained with the analytical approach. For comparison, Panel C reports the same information for the bias-unadjusted estimators (obtained by integrating the bias-unadjusted density).

Panel A: Bias Adjustment (Numerical Integration)

Mean

MSE

n\T	100	250	500	750	1000
1000	1.63	0.12	0.02	0.01	0.01
2500	1.48	0.12	0.02	0.01	0.01
5000	1.49	0.12	0.02	0.01	0.01
7500	1.40	0.11	0.02	0.01	0.01
10000	1.43	0.11	0.02	0.01	0.01

Bias

n\T	100	250	500	750	1000
1000	1.22	0.32	0.14	0.09	0.07
2000	1.20	0.34	0.14	0.10	0.08
3000	1.21	0.34	0.14	0.09	0.07
4000	1.18	0.33	0.14	0.10	0.07
5000	1.19	0.33	0.14	0.10	0.07

Standard Deviation

n\T	100	250	500	750	1000
1000	0.38	0.11	0.06	0.06	0.06
2000	0.21	0.06	0.04	0.03	0.03
3000	0.16	0.05	0.03	0.02	0.02
4000	0.14	0.04	0.02	0.02	0.02
5000	0.10	0.04	0.02	0.02	0.02

Volatility

MSE

n\T	100	250	500	750	1000
1000	14.87	0.12	0.01	0.01	0.00
2500	16.29	0.19	0.00	0.00	0.00
5000	22.82	0.09	0.00	0.00	0.00
7500	32.68	0.07	0.00	0.00	0.00
10000	16.64	0.07	0.00	0.00	0.00

Bias

n\T	100	250	500	750	1000
1000	3.54	0.28	-0.02	-0.02	-0.03
2500	3.80	0.29	-0.03	-0.03	-0.02
5000	4.10	0.27	-0.03	-0.03	-0.02
7500	4.32	0.25	-0.03	-0.03	-0.02
10000	3.94	0.25	-0.03	-0.03	-0.02

Standard Deviation

n\T	100	250	500	750	1000
1000	1.52	0.22	0.08	0.07	0.06
2500	1.37	0.32	0.05	0.04	0.04
5000	2.45	0.13	0.03	0.03	0.03
7500	3.74	0.08	0.03	0.03	0.03
10000	1.05	0.09	0.03	0.03	0.02

Skewness

MSE

n\T	100	250	500	750	1000
1000	47.90	4.04	0.17	0.10	0.10
2500	109.91	19.89	0.10	0.05	0.05
5000	206.97	10.49	0.04	0.02	0.02
7500	406.00	3.89	0.04	0.02	0.02
10000	254.93	7.25	0.03	0.02	0.01

Bias

n\T	100	250	500	750	1000
1000	1.85	0.02	0.01	-0.00	0.02
2000	4.50	1.02	0.10	0.06	0.05
3000	5.55	1.30	0.08	0.05	0.05
4000	6.36	0.94	0.10	0.07	0.06
5000	7.69	0.82	0.11	0.08	0.05

Standard Deviation

n\T	100	250	500	750	1000
1000	6.67	2.01	0.41	0.32	0.31
2000	9.47	4.34	0.30	0.22	0.21
3000	13.27	2.96	0.18	0.14	0.14
4000	19.12	1.73	0.18	0.11	0.11
5000	13.99	2.57	0.13	0.10	0.10

Proportion with Positive Skill Measure

MSE

n\T	100	250	500	750	1000
1000	65.64	4.91	1.55	1.51	1.16
2500	62.50	2.05	0.68	0.41	0.49
5000	59.38	1.01	0.20	0.28	0.28
7500	56.14	0.77	0.23	0.18	0.17
10000	55.74	0.78	0.26	0.19	0.14

Bias

n\T	100	250	500	750	1000
1000	-7.45	-0.99	-0.05	-0.27	-0.10
2500	-7.65	-0.80	0.07	0.19	0.11
5000	-7.42	-0.70	0.11	0.13	0.08
7500	-7.32	-0.67	0.13	0.17	0.13
10000	-7.33	-0.66	0.22	0.14	0.12

Standard Deviation

n\T	100	250	500	750	1000
1000	3.20	1.98	1.24	1.20	1.07
2500	1.98	1.19	0.82	0.61	0.69
5000	2.07	0.72	0.44	0.51	0.53
7500	1.60	0.57	0.46	0.38	0.40
10000	1.40	0.59	0.46	0.41	0.36

Table AIV
Properties of the Estimated Moments and Proportion:
Scale Coefficient (Continued)

Panel B: Bias Adjustment (Analytical Approach)

Mean

MSE

n\T	100	250	500	750	1000
1000	1.63	0.12	0.02	0.01	0.01
2500	1.48	0.12	0.02	0.01	0.01
5000	1.49	0.12	0.02	0.01	0.01
7500	1.40	0.11	0.02	0.01	0.01
10000	1.43	0.11	0.02	0.01	0.01

Bias

n\T	100	250	500	750	1000
1000	1.22	0.32	0.14	0.09	0.07
2000	1.20	0.34	0.14	0.10	0.08
3000	1.21	0.34	0.14	0.09	0.07
4000	1.18	0.33	0.14	0.10	0.07
5000	1.19	0.33	0.14	0.10	0.07

Standard Deviation

n\T	100	250	500	750	1000
1000	0.38	0.11	0.06	0.06	0.06
2000	0.21	0.06	0.04	0.03	0.03
3000	0.16	0.05	0.03	0.02	0.02
4000	0.14	0.04	0.02	0.02	0.02
5000	0.10	0.04	0.02	0.02	0.02

Volatility

MSE

n\T	100	250	500	750	1000
1000	6.75	0.08	0.01	0.00	0.00
2500	7.26	0.15	0.00	0.00	0.00
5000	11.97	0.06	0.00	0.00	0.00
7500	8.80	0.05	0.00	0.00	0.00
10000	6.48	0.05	0.00	0.00	0.00

Bias

n\T	100	250	500	750	1000
1000	2.10	0.21	0.03	0.02	0.01
2500	2.24	0.26	0.03	0.01	0.01
5000	2.45	0.22	0.03	0.01	0.01
7500	2.00	0.22	0.03	0.01	0.01
10000	2.24	0.22	0.03	0.01	0.01

Standard Deviation

n\T	100	250	500	750	1000
1000	1.54	0.18	0.07	0.07	0.06
2500	1.49	0.28	0.05	0.04	0.04
5000	2.44	0.12	0.03	0.03	0.03
7500	2.20	0.06	0.03	0.03	0.03
10000	1.21	0.09	0.03	0.02	0.02

Skewness

MSE

n\T	100	250	500	750	1000
1000	78.69	3.59	0.48	0.24	0.17
2500	117.73	9.57	0.45	0.21	0.14
5000	452.01	16.92	0.38	0.17	0.10
7500	549.80	4.15	0.40	0.18	0.10
10000	367.33	4.94	0.40	0.18	0.10

Bias

n\T	100	250	500	750	1000
1000	3.19	0.21	-0.05	-0.03	-0.02
2500	3.37	0.24	-0.05	-0.04	-0.03
5000	3.44	0.23	-0.07	-0.05	-0.03
7500	3.41	0.22	-0.06	-0.05	-0.03
10000	3.43	0.22	-0.06	-0.04	-0.02

Standard Deviation

n\T	100	250	500	750	1000
1000	8.28	1.88	0.69	0.48	0.42
2500	10.31	3.08	0.67	0.46	0.37
5000	20.98	4.11	0.62	0.41	0.32
7500	23.20	2.03	0.63	0.42	0.32
10000	18.86	2.21	0.63	0.42	0.31

Proportion with Positive Skill Measure

MSE

n\T	100	250	500	750	1000
1000	41.51	4.31	1.79	1.44	1.41
2500	45.02	1.68	0.84	0.51	0.51
5000	45.71	0.66	0.30	0.30	0.33
7500	47.16	0.38	0.29	0.18	0.18
10000	47.51	0.41	0.36	0.18	0.14

Bias

n\T	100	250	500	750	1000
1000	-5.61	0.15	0.35	-0.06	0.10
2500	-6.34	0.20	0.34	0.29	0.14
5000	-6.42	0.18	0.31	0.14	0.09
7500	-6.60	0.20	0.28	0.15	0.09
10000	-6.72	0.21	0.36	0.13	0.08

Standard Deviation

n\T	100	250	500	750	1000
1000	3.17	2.07	1.29	1.20	1.18
2500	2.20	1.28	0.85	0.66	0.70
5000	2.12	0.79	0.45	0.53	0.56
7500	1.91	0.58	0.46	0.39	0.42
10000	1.56	0.61	0.48	0.41	0.37

Table AIV
Properties of the Estimated Moments and Proportion: Scale Coefficient (Continued)

Panel C: No Bias Adjustment

Mean

MSE

n\T	100	250	500	750	1000
1000	1.63	0.12	0.02	0.01	0.01
2500	1.48	0.12	0.02	0.01	0.01
5000	1.49	0.12	0.02	0.01	0.01
7500	1.40	0.11	0.02	0.01	0.01
10000	1.43	0.11	0.02	0.01	0.01

Bias

n\T	100	250	500	750	1000
1000	1.22	0.32	0.14	0.09	0.07
2000	1.20	0.34	0.14	0.10	0.08
3000	1.21	0.34	0.14	0.09	0.07
4000	1.18	0.33	0.14	0.10	0.07
5000	1.19	0.33	0.14	0.10	0.07

Standard Deviation

n\T	100	250	500	750	1000
1000	0.38	0.11	0.06	0.06	0.06
2000	0.21	0.06	0.04	0.03	0.03
3000	0.16	0.05	0.03	0.02	0.02
4000	0.14	0.04	0.02	0.02	0.02
5000	0.10	0.04	0.02	0.02	0.02

Volatility

MSE

n\T	100	250	500	750	1000
1000	24.88	0.89	0.09	0.03	0.02
2500	26.67	1.00	0.09	0.03	0.01
5000	34.43	0.90	0.09	0.03	0.01
7500	46.26	0.87	0.09	0.03	0.01
10000	26.96	0.87	0.09	0.03	0.01

Bias

n\T	100	250	500	750	1000
1000	4.77	0.93	0.30	0.17	0.11
2500	4.99	0.96	0.30	0.16	0.11
5000	5.30	0.94	0.29	0.16	0.11
7500	5.52	0.93	0.30	0.16	0.11
10000	5.09	0.93	0.29	0.17	0.11

Standard Deviation

n\T	100	250	500	750	1000
1000	1.45	0.18	0.07	0.06	0.06
2500	1.31	0.30	0.04	0.04	0.04
5000	2.51	0.11	0.03	0.03	0.03
7500	3.97	0.06	0.03	0.02	0.02
10000	1.01	0.07	0.02	0.02	0.02

Skewness

MSE

n\T	100	250	500	750	1000
1000	29.55	1.90	0.13	0.10	0.09
2500	67.08	13.11	0.08	0.06	0.04
5000	147.57	3.22	0.06	0.05	0.03
7500	293.19	0.88	0.06	0.04	0.02
10000	159.53	2.10	0.05	0.03	0.02

Bias

n\T	100	250	500	750	1000
1000	1.18	-0.26	-0.23	-0.20	-0.14
2500	3.06	0.46	-0.20	-0.17	-0.12
5000	3.98	0.45	-0.22	-0.18	-0.13
7500	4.57	0.20	-0.21	-0.17	-0.13
10000	5.46	0.15	-0.21	-0.17	-0.14

Standard Deviation

n\T	100	250	500	750	1000
1000	5.31	1.35	0.28	0.25	0.26
2500	7.60	3.59	0.21	0.17	0.17
5000	11.48	1.74	0.11	0.11	0.11
7500	16.50	0.92	0.12	0.08	0.09
10000	11.39	1.44	0.09	0.08	0.08

Proportion with Positive Skill Measure

MSE

n\T	100	250	500	750	1000
1000	164.79	49.68	13.90	7.38	3.97
2500	163.81	46.76	13.13	5.01	2.98
5000	160.98	46.56	12.92	5.51	2.98
7500	163.14	47.12	13.06	5.37	2.85
10000	163.64	47.03	12.63	5.48	2.87

Bias

n\T	100	250	500	750	1000
1000	-12.75	-6.91	-3.55	-2.48	-1.65
2500	-12.76	-6.79	-3.54	-2.15	-1.59
5000	-12.67	-6.80	-3.57	-2.30	-1.64
7500	-12.76	-6.85	-3.59	-2.29	-1.64
10000	-12.78	-6.85	-3.53	-2.31	-1.66

Standard Deviation

n\T	100	250	500	750	1000
1000	1.52	1.41	1.13	1.10	1.12
2500	0.95	0.83	0.76	0.61	0.66
5000	0.58	0.57	0.40	0.48	0.54
7500	0.48	0.42	0.40	0.36	0.40
10000	0.43	0.41	0.43	0.37	0.35

Table AV
Fund Style Classification

This table provides the list of 32 styles across the different data providers of style information (Wiesbenberger, Strategic Insight, Lipper, Policy CRSP). For each style, it also shows the mapping between each style and the growth/value (GV) and small/large cap (SL) dimensions. A value of 1 refers to growth or small cap. A value of two refers to neutral fund in terms of GV or SL dimension. Finally, a value of 3 refers to value or large cap.

Wiesenberger	Symbol	Name	Style GV	Style SL
1	G	Growth	1	
2	GCI	Growth and current income	3	
3	G-I	Income	3	
4	IEQ	Equity income	3	
5	LTG	Long-term growth	1	
6	MCG	Maximum capital gains	1	
7	SCG	Small-cap growth	1	1

Strategic Insight	Symbol	Name	Style GV	Style SL
8	AGG	Aggressive growth	1	
9	GMC	Equity mid-cap		2
10	GRI	Growth and income	3	
11	GRO	Growth	1	
12	ING	Income and growth	3	
13	SCG	Small-cap		1

Table AV
Fund Style Classification (Continued)

Lipper	Symbol	Name	Style GV	Style SL
14	CA	Capital appreciation	1	
15	G	Growth	1	
16	GI	Growth and income	3	
17	LCCE	Large-cap core	2	3
18	LCGE	Large-cap growth	1	3
19	LCVE	Large-cap value	3	3
20	MC	Mid-cap		2
21	MCCE	Mid-cap core	2	2
22	MCGE	Mid-cap growth	1	2
23	MCVE	Mid-cap value	3	2
24	MLCE	Multi-cap core	2	
25	MLGE	Multi-cap growth	1	
26	MLVE	Multi-cap value	3	
27	MR	Micro-cap		1
28	SCCE	Small-cap core	2	1
29	SCGE	Small-cap growth	1	1
30	SCVE	Small-cap value	3	1
31	SG	Small-cap		1

Policy CRSP	Symbol	Name	Style GV	Style SL
32	CS	Common stock		

Table AVI
Impact of Survivorship and Reverse Survivorship Bias

Panel A contains the summary statistics of the distributions of the skill and scale coefficients for all funds in the population across different thresholds for the minimum number of return observations (ranging from 12 to 60 monthly observations). It reports the first four moments, the proportions of funds with a negative and positive skill coefficient, and the quantiles at 5% and 95%. We compute all cross-sectional estimates by integrating numerically the bias-adjusted density obtained with our nonparametric approach. Figures in parentheses denote the estimated standard deviation of each estimator. Panel B repeats the analysis for the subpopulation of funds that disappear during the sample period across two thresholds for the minimum number of return observations (60 and 108 monthly observations). This analysis provides a rough estimate of the magnitude of the reverse survivorship bias.

Panel A: All Selected Funds

	Moments				Proportions (%)		Quantiles (Ann.)	
	Mean (Ann.)	Std. Dev. (Ann.)	Skewness	Kurtosis	Negative	Positive	5%	95%
Min. Observations=12								
Skill Coefficent	2.7 (0.2)	5.4 (0.4)	0.1 (1.2)	37.6 (8.1)	21.8 (0.8)	78.2 (0.8)	-3.2 (0.2)	9.5 (0.2)
Scale Coefficient	1.3 (0.1)	1.9 (0.1)	1.2 (0.9)	26 (7.3)	20 (0.8)	80 (0.8)	-1.1 (0.1)	4.2 (0.1)
Min. Observations=36								
Skill Coefficent	2.8 (0.1)	4.8 (0.3)	1.3 (0.9)	28.8 (5.3)	19.9 (0.8)	80.1 (0.8)	-2.8 (0.1)	9.2 (0.2)
Scale Coefficient	1.3 (0.1)	1.9 (0.1)	1.8 (0.9)	25.6 (8.8)	19.2 (0.8)	80.8 (0.8)	-1 (0.1)	4.1 (0.1)
Min. Observations=60								
Skill Coefficent	3 (0.1)	4.1 (0.2)	1.6 (0.7)	23.4 (6)	16.9 (0.8)	83.1 (0.8)	-2.2 (0.1)	8.9 (0.2)
Scale Coefficient	1.3 (0.1)	1.7 (0.1)	1.6 (0.7)	16.7 (11)	17.6 (0.8)	82.4 (0.8)	-0.9 (0.1)	3.9 (0.1)

Panel B: Selected Funds that Disappear During the Sample Period

	Moments				Proportions (%)		Quantiles (Ann.)	
	Mean (Ann.)	Std. Dev. (Ann.)	Skewness	Kurtosis	Negative	Positive	5%	95%
Min. Observations=60								
Skill Coefficent	2.5 (0.1)	4.9 (0.3)	2.2 (0.7)	27 (6)	25.8 (0.9)	74.2 (0.9)	-3.6 (0.2)	9.9 (0.2)
Scale Coefficient	1.4 (0.1)	2.1 (0.1)	2.1 (0.7)	21.5 (7.3)	23.2 (0.9)	76.8 (0.9)	-1.4 (0.1)	4.6 (0.1)
Min. Observations=108								
Skill Coefficent	2.4 (0.1)	3.3 (0.1)	0.8 (0.3)	8 (1.1)	21 (0.8)	79 (0.8)	-2.4 (0.1)	7.8 (0.1)
Scale Coefficient	1.3 (0.1)	1.6 (0.1)	1.3 (0.3)	9.9 (1.5)	19.2 (0.8)	80.8 (0.8)	-0.9 (0.1)	3.7 (0.1)

Table AVII
Distributions of Skill and Scalability
Four-Factor Model

Panel A contains the summary statistics of the distribution of the skill coefficient for all funds in the population, small/large cap funds, low/high turnover funds (i.e., bottom or top tercile of funds sorted on turnover), and broker/direct sold funds (i.e., funds that are sold through brokers or funds directly sold to investors) based on the four-factor model of Carhart (1997). It reports the first four moments, the proportions of funds with a negative and positive skill coefficient, and the quantiles at 5% and 95%. We compute all cross-sectional estimates by integrating numerically the bias-adjusted density obtained with our nonparametric approach. Figures in parentheses denote the estimated standard deviation of each estimator. Panel B repeats the analysis for the scale coefficient. To ease interpretation, we standardize the scale coefficient for each fund so that it corresponds to the change in gross alpha for a one standard deviation change in size.

Panel A: Skill Coefficient

	Moments				Proportions (%)		Quantiles (Ann.)	
	Mean (Ann.)	Std. Dev. (Ann.)	Skewness	Kurtosis	Negative	Positive	5%	95%
All Funds	2.4 (0.1)	3.9 (0.3)	1.8 (1.1)	30.4 (15.4)	21.5 (0.8)	78.5 (0.8)	-2.6 (0.1)	8.1 (0.1)
Fund Groups								
Small Cap	3.3 (0.2)	4.2 (0.5)	1.9 (2.5)	27.9 (41)	19.4 (1.6)	80.6 (1.6)	-3 (0.3)	9.8 (0.3)
Large Cap	1.6 (0.1)	2.7 (0.2)	1.2 (0.6)	13.5 (3.7)	24.5 (1.4)	75.5 (1.4)	-2.1 (0.2)	5.9 (0.2)
Low Turnover	2.1 (0.2)	3.1 (0.2)	-0.2 (0.7)	13.4 (2)	18.7 (1.4)	81.3 (1.4)	-2 (0.2)	6.7 (0.2)
High Turnover	2.6 (0.2)	5 (0.5)	2.1 (1.3)	26.3 (15.4)	24.9 (1.5)	75.1 (1.5)	-3.7 (0.2)	9.8 (0.3)
Broker Sold	2.4 (0.2)	4.1 (0.4)	2.6 (1.8)	37.7 (25.9)	22.2 (1.3)	77.8 (1.3)	-2.7 (0.2)	8.3 (0.2)
Direct Sold	2.6 (0.1)	2.9 (0.2)	0.4 (0.5)	8.5 (1.8)	14.1 (1.2)	85.9 (1.2)	-1.4 (0.2)	7.4 (0.2)

Panel B: Scale Coefficient

	Moments				Proportions (%)		Quantiles (Ann.)	
	Mean (Ann.)	Std. Dev. (Ann.)	Skewness	Kurtosis	Negative	Positive	5%	95%
All Funds	1.3 (0.1)	1.7 (0.1)	1.5 (0.7)	16 (9.8)	19.2 (0.8)	80.8 (0.8)	-1 (0.1)	3.8 (0.1)
Fund Groups								
Small Cap	1.5 (0.1)	1.7 (0.1)	-0.1 (1)	8.4 (9.9)	18.4 (1.5)	81.6 (1.5)	-1.2 (0.1)	4.3 (0.1)
Large Cap	0.9 (0.1)	1.4 (0.1)	1.7 (0.6)	13.5 (4)	24.6 (1.4)	75.4 (1.4)	-1 (0.1)	3 (0.1)
Low Turnover	0.8 (0.1)	1.1 (0.1)	0.4 (0.3)	4.7 (1.3)	21.6 (1.5)	78.4 (1.5)	-0.9 (0.1)	2.7 (0.1)
High Turnover	1.7 (0.1)	2.1 (0.2)	1 (0.5)	8.9 (3.7)	18.7 (1.4)	81.3 (1.4)	-1.2 (0.1)	5.1 (0.2)
Broker Sold	1.3 (0.1)	1.8 (0.1)	1.2 (0.5)	11.1 (1.4)	19.5 (1.2)	80.5 (1.2)	-1 (0.1)	4.2 (0.1)
Direct Sold	1.3 (0.1)	1.3 (0.1)	0.6 (0.5)	7.5 (2.3)	15.3 (1.3)	84.7 (1.3)	-0.7 (0.1)	3.4 (0.1)

Table AVIII
Distributions of Skill and Scalability
Five-Factor Model

Panel A contains the summary statistics of the distribution of the skill coefficient for all funds in the population, small/large cap funds, low/high turnover funds (i.e., bottom or top tercile of funds sorted on turnover), and broker/direct sold funds (i.e., funds that are sold through brokers or funds directly sold to investors) based on the five-factor model of Fama and French (2015). It reports the first four moments, the proportions of funds with a negative and positive skill coefficient, and the quantiles at 5% and 95%. We compute all cross-sectional estimates by integrating numerically the bias-adjusted density obtained with our nonparametric approach. Figures in parentheses denote the estimated standard deviation of each estimator. Panel B repeats the analysis for the scale coefficient. To ease interpretation, we standardize the scale coefficient for each fund so that it corresponds to the change in gross alpha for a one standard deviation change in size.

Panel A: Skill Coefficient

	Moments				Proportions (%)		Quantiles (Ann.)	
	Mean (Ann.)	Std. Dev. (Ann.)	Skewness	Kurtosis	Negative	Positive	5%	95%
All Funds	2.4 (0.1)	4.4 (0.2)	1.4 (0.5)	16.8 (2.3)	26 (0.9)	74 (0.9)	-3.1 (0.1)	8.9 (0.2)
Fund Groups								
Small Cap	3.4 (0.2)	4.3 (0.3)	0.8 (0.5)	7.8 (2.8)	20.2 (1.6)	79.8 (1.6)	-3 (0.3)	10.3 (0.3)
Large Cap	1.5 (0.2)	3.5 (0.2)	1.8 (0.4)	13.3 (2.3)	31.3 (1.5)	68.7 (1.5)	-3 (0.2)	6.8 (0.2)
Low Turnover	1.5 (0.2)	3.9 (0.3)	1.4 (0.8)	18.4 (3.1)	32.8 (1.7)	67.2 (1.7)	-3.2 (0.2)	7.1 (0.2)
High Turnover	3.5 (0.2)	5.1 (0.4)	1 (0.6)	12.8 (2.7)	19.7 (1.4)	80.3 (1.4)	-3.5 (0.3)	11.5 (0.3)
Broker Sold	2.4 (0.2)	4.3 (0.3)	1 (0.6)	13 (2.8)	26.2 (1.4)	73.8 (1.4)	-3.2 (0.2)	8.9 (0.2)
Direct Sold	2.5 (0.2)	3.6 (0.2)	1.1 (0.5)	11.3 (1.7)	21.8 (1.5)	78.2 (1.5)	-2.1 (0.2)	8.1 (0.2)

Panel B: Scale Coefficient

	Moments				Proportions (%)		Quantiles (Ann.)	
	Mean (Ann.)	Std. Dev. (Ann.)	Skewness	Kurtosis	Negative	Positive	5%	95%
All Funds	1.2 (0.1)	1.7 (0.1)	1.1 (0.5)	13.9 (4)	19.9 (0.8)	80.1 (0.8)	-1 (0.1)	3.9 (0.1)
Fund Groups								
Small Cap	1.5 (0.1)	1.7 (0.1)	-0.1 (0.9)	7.3 (8.4)	18.4 (1.5)	81.6 (1.5)	-1.2 (0.1)	4.4 (0.2)
Large Cap	0.9 (0.1)	1.5 (0.1)	1.9 (0.7)	16.9 (4.2)	24.7 (1.4)	75.3 (1.4)	-1.1 (0.1)	3.2 (0.1)
Low Turnover	0.9 (0.1)	1.2 (0.1)	-0.1 (0.5)	6.3 (2.4)	24.4 (1.5)	75.6 (1.5)	-1 (0.1)	2.9 (0.1)
High Turnover	1.6 (0.1)	2.1 (0.2)	1 (0.6)	9.8 (4.3)	18.7 (1.4)	81.3 (1.4)	-1.2 (0.1)	5.1 (0.2)
Broker Sold	1.3 (0.1)	1.8 (0.1)	0.8 (0.5)	10.5 (1.5)	19.7 (1.2)	80.3 (1.2)	-1.1 (0.1)	4.2 (0.1)
Direct Sold	1.2 (0.1)	1.4 (0.1)	-0.4 (0.4)	7.5 (1.4)	16.6 (1.3)	83.4 (1.3)	-0.7 (0.1)	3.4 (0.1)

Table AIX
Distributions of Skill and Scalability
Changes in Industry Competition

Panel A contains the summary statistics of the distribution of the skill coefficient for all funds in the population, small/large cap funds, low/high turnover funds (i.e., bottom or top tercile of funds sorted on turnover), and broker/direct sold funds (i.e., funds that are sold through brokers or funds directly sold to investors) after including a proxy for industry competition in the set of variables (the ratio of the industry size on the total market capitalization). It reports the first four moments, the proportions of funds with a negative and positive skill coefficient, and the quantiles at 5% and 95%. We compute all cross-sectional estimates by integrating numerically the bias-adjusted density obtained with our nonparametric approach. Figures in parentheses denote the estimated standard deviation of each estimator. Panel B repeats the analysis for the scale coefficient. To ease interpretation, we standardize the scale coefficient for each fund so that it corresponds to the change in gross alpha for a one standard deviation change in size.

Panel A: Skill Coefficient

	Moments				Proportions (%)		Quantiles (Ann.)	
	Mean (Ann.)	Std. Dev. (Ann.)	Skewness	Kurtosis	Negative	Positive	5%	95%
All Funds	3.5 (0.1)	5.1 (0.3)	1.4 (0.9)	28 (6.2)	17.7 (0.8)	82.3 (0.8)	-2.7 (0.2)	10.5 (0.2)
Fund Groups								
Small Cap	5.4 (0.3)	5.9 (0.6)	2.4 (1.2)	25.3 (8.9)	12.1 (1.3)	87.9 (1.3)	-2.3 (0.3)	14.4 (0.4)
Large Cap	2 (0.2)	3.4 (0.2)	1.1 (0.6)	13.2 (1.7)	22.7 (1.3)	77.3 (1.3)	-2.4 (0.2)	7 (0.2)
Low Turnover	3.1 (0.2)	4.3 (0.3)	-0.7 (1)	17.1 (5.5)	18.4 (1.4)	81.6 (1.4)	-2.7 (0.2)	9 (0.2)
High Turnover	3.9 (0.3)	6.2 (0.6)	2.1 (1.1)	26.3 (6.5)	19.3 (1.4)	80.7 (1.4)	-3.4 (0.3)	12.5 (0.3)
Broker Sold	3.4 (0.2)	4.8 (0.4)	2 (1.3)	28.7 (13.8)	17.9 (1.2)	82.1 (1.2)	-2.5 (0.2)	10.6 (0.2)
Direct Sold	3.8 (0.2)	4.2 (0.3)	0.5 (0.5)	9.4 (1)	12.9 (1.2)	87.1 (1.2)	-1.5 (0.2)	10.1 (0.2)

Panel B: Scale Coefficient

	Moments				Proportions (%)		Quantiles (Ann.)	
	Mean (Ann.)	Std. Dev. (Ann.)	Skewness	Kurtosis	Negative	Positive	5%	95%
All Funds	1.6 (0.1)	2.3 (0.1)	1.1 (0.6)	17.3 (4.3)	17.4 (0.8)	82.6 (0.8)	-1.2 (0.1)	5.1 (0.1)
Fund Groups								
Small Cap	2.2 (0.1)	2.6 (0.2)	1.5 (0.5)	11.8 (1.9)	14.8 (1.4)	85.2 (1.4)	-1.3 (0.2)	6.1 (0.2)
Large Cap	1.2 (0.1)	1.6 (0.1)	1.1 (0.6)	10.3 (3.6)	20.4 (1.3)	79.6 (1.3)	-1.1 (0.1)	3.7 (0.1)
Low Turnover	1.3 (0.1)	1.7 (0.1)	0.4 (0.9)	11.6 (6.5)	18.9 (1.4)	81.1 (1.4)	-1 (0.1)	4.1 (0.1)
High Turnover	2.1 (0.1)	2.7 (0.2)	0.8 (0.8)	12.6 (5.4)	18.1 (1.4)	81.9 (1.4)	-1.6 (0.2)	6.2 (0.2)
Broker Sold	1.7 (0.1)	2.3 (0.2)	1 (0.5)	11.2 (1.8)	17.9 (1.2)	82.1 (1.2)	-1.2 (0.1)	5.2 (0.1)
Direct Sold	1.8 (0.1)	2 (0.2)	1.9 (1.4)	20.2 (21)	14.2 (1.2)	85.8 (1.2)	-0.8 (0.1)	4.9 (0.1)

Table AX
Distributions of Skill and Scalability
Changes in Industry Competition and Aggregate Mispricing

Panel A contains the summary statistics of the distribution of the skill coefficient for all funds in the population, small/large cap funds, low/high turnover funds (i.e., bottom or top tercile of funds sorted on turnover), and broker/direct sold funds (i.e., funds that are sold through brokers or funds directly sold to investors) after including a proxy for industry competition and aggregate mispricing in the set of variables (the ratio of the industry size on the total market capitalization and aggregate fund turnover). It reports the first four moments, the proportions of funds with a negative and positive skill coefficient, and the quantiles at 5% and 95%. We compute all cross-sectional estimates by integrating numerically the bias-adjusted density obtained with our nonparametric approach. Figures in parentheses denote the estimated standard deviation of each estimator. Panel B repeats the analysis for the scale coefficient. To ease interpretation, we standardize the scale coefficient for each fund so that it corresponds to the change in gross alpha for a one standard deviation change in size.

Panel A: Skill Coefficient

	Moments				Proportions (%)		Quantiles (Ann.)	
	Mean (Ann.)	Std. Dev. (Ann.)	Skewness	Kurtosis	Negative	Positive	5%	95%
All Funds	3.4 (0.2)	5.4 (0.3)	1.4 (0.8)	26.2 (5.6)	19.2 (0.8)	80.8 (0.8)	-3.1 (0.2)	10.8 (0.2)
Fund Groups								
Small Cap	5.4 (0.3)	6.1 (0.6)	2.4 (1)	23.6 (8.2)	13.6 (1.4)	86.4 (1.4)	-2.6 (0.3)	14.5 (0.4)
Large Cap	2.1 (0.2)	3.7 (0.3)	0.6 (0.6)	14 (1.5)	24.3 (1.4)	75.7 (1.4)	-2.8 (0.2)	7.3 (0.2)
Low Turnover	3.1 (0.2)	4.7 (0.4)	0.2 (1.1)	20.1 (5.9)	18.8 (1.4)	81.2 (1.4)	-2.7 (0.2)	9.6 (0.2)
High Turnover	3.8 (0.3)	6.3 (0.6)	2.1 (1)	25.1 (6.8)	21.7 (1.5)	78.3 (1.5)	-3.9 (0.3)	12.6 (0.3)
Broker Sold	3.3 (0.2)	5.1 (0.5)	2.3 (1.4)	32.6 (14.7)	18.9 (1.2)	81.1 (1.2)	-3 (0.2)	10.8 (0.3)
Direct Sold	3.8 (0.2)	4.5 (0.3)	0.7 (0.5)	10.9 (1)	14.3 (1.2)	85.7 (1.2)	-1.9 (0.2)	10.3 (0.3)

Panel B: Scale Coefficient

	Moments				Proportions (%)		Quantiles (Ann.)	
	Mean (Ann.)	Std. Dev. (Ann.)	Skewness	Kurtosis	Negative	Positive	5%	95%
All Funds	1.6 (0.1)	2.4 (0.1)	1.4 (0.5)	17.4 (3.2)	19.4 (0.8)	80.6 (0.8)	-1.3 (0.1)	5.3 (0.1)
Fund Groups								
Small Cap	2.2 (0.2)	2.9 (0.2)	1.4 (0.5)	10.9 (1.9)	17.1 (1.5)	82.9 (1.5)	-1.5 (0.2)	6.4 (0.2)
Large Cap	1.2 (0.1)	1.6 (0.1)	0.7 (0.4)	7.3 (2.1)	21.6 (1.3)	78.4 (1.3)	-1.2 (0.1)	3.7 (0.1)
Low Turnover	1.3 (0.1)	1.8 (0.2)	0.6 (1.3)	14 (14.4)	20.7 (1.4)	79.3 (1.4)	-1.1 (0.1)	4.1 (0.1)
High Turnover	2 (0.1)	2.9 (0.2)	0.9 (0.6)	10.9 (2.9)	20 (1.4)	80 (1.4)	-1.8 (0.2)	6.3 (0.2)
Broker Sold	1.7 (0.1)	2.3 (0.2)	1.2 (0.5)	11.4 (1.5)	19.4 (1.2)	80.6 (1.2)	-1.3 (0.1)	5.2 (0.1)
Direct Sold	1.8 (0.1)	2.1 (0.2)	1.8 (1)	18 (9.9)	15.8 (1.3)	84.2 (1.3)	-1 (0.1)	5.3 (0.2)

Figure A1
Comparative Static Analysis of the EIV Bias
Skill Coefficient

This figure performs a comparative static analysis of the EIV bias function for the skill coefficient. We plot the benchmark curve using the parameters of the Gaussian reference model calibrated on our sample. In Panel A, we plot the new EIV bias function after increasing the variance of the true skill coefficient by 0.002/100. In Panel B, we plot the new EIV bias function after increasing the variance of the estimated skill coefficient by 0.002/100. In Panel C, we plot the new EIV bias function after increasing the correlation between the true skill coefficient and the estimation variance by 50% in relative terms.