Skill, Scale, and Value Creation in the Mutual Fund Industry

Laurent Barras, Patrick Gagliardini, and Olivier Scaillet*

This version, July 8, 2020

JEL Classification: G11, G12, C14, C33, C58.
Keywords: Mutual funds, skill, scale, value added, nonparametric estimation, large panel, error-in-variable bias

*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. We thank Yacine Ait-Sahalia, David Ardia, Jonathan Berk, Sebastien Betermier, Yong Chen, Jean-David Fermanian, Francesco Franzoni, Marcelo Fernandes, Laurent Fresard, Mila Getmansky, Michel Habib, Alexander Kempf, Gaelle Le Fol, Peter Limbach, Jordan Martel, David Schumacher, Soňa Sidalova, Nick Roussanov, Benoit Udekem, Florian Weigert, participants at the 2018 EC2 Meeting on Big Data Econometrics, the 2018 Meeting of the International Society for Nonparametric Statistics, the 2018 McGill/HEC Spring Workshop, the 2018 Paris December Conference, the 2018 Swiss Finance Institute Research Days, the 2019 Cambridge Big Data Workshop, the 2019 China International Conference in Finance, the 2019 Cologne Colloquium on Financial Markets, the 2019 European Meeting of the Econometric Society, the 2019 Meeting of the European Finance Association (EFA), the 2019 Meeting of Quantitative Finance and Financial Econometrics, the 2019 Workshop on Financial Markets and Nonlinear Dynamics, the 2019 International Workshop in Financial Econometrics, the 2020 Review of Asset Pricing Studies (RAPS) Winter Conference, as well as seminar participants at Carlos III, Emory, Florida International University, Georgia State, Georgia Tech, McGill, McMaster, Singapore Management University (SMU), National University of Singapore (NUS), Universite d'Orleans, Universite Libre de Bruxelles (ULB), and the Universities of Geneva, Miami, Lugano, and Luxemburg for their comments. The first author thank the Social Sciences and Humanities Research Council of Canada (SSHRC) for its financial support. The third author acknowledges financial support by the Geneva Finance Research Institute. He did part of this research when visiting the University of Cergy-Pontoise. A previous version circulated under the title: "The Cross-Sectional Distribution of Fund Skill Measures".

ABSTRACT

We develop a novel approach to jointly examine skill, scale, and value added across individual funds. We find that the value added is (i) positive for the vast majority of funds, and (ii) close to its optimal level after an adjustment period possibly due to investors' learning. We also show that skill and scale (i) vary substantially, and (ii) are strongly correlated across funds—two features that shape the value added distribution. These results are consistent with theoretical models in which funds have bargaining power over investors, and highlight the importance of the fund industry in improving market efficiency.

I Introduction

The academic literature on mutual funds has largely focused on performance, i.e., whether investors earn positive alphas when they buy mutual fund shares.[1] However, we know far less about value creation, i.e., whether funds extract value from capital markets through their investment decisions.[2] The analysis of value creation is important because it sheds new light on the optimal size of the mutual fund industry and the financial sector as a whole (e.g., Cochrane (2013), Greenwood and Scharfstein (2013)). This analysis is also informative about the role of the fund industry in improving the allocation of resources in the economy (e.g., Pedersen (2018))—a role that is independent of how the value created is split between funds and investors.

The study of value creation, or value added, is pioneered by Berk and van Binsbergen (2015; BvB hereafter). They define the fund's value added as the product of its gross alpha and size. As such, the value added is similar to the notion of economic rent defined as the markup price times the quantity sold. Empirically, BvB show that the gross alpha gives a distorted view of value creation—measuring the value added across funds, they find that its variation is mostly driven by cross-sectional differences in fund size.

In this paper, we provide the first fund-level analysis of skill, scale, and value added. Put simply, we propose an alternative decomposition of the value added—instead of focusing on size and gross alpha as in BvB, we focus on skill and scale. Our analysis builds on the premise that the value created by each fund ultimately depends on (i) skill, i.e., the fund's ability to identify profitable investment ideas, and (ii) scale, i.e., the limit on the scalability of these ideas as the fund grows in size. The extensive panel evidence documented by Zhu (2018) confirms the presence of such diseconomies of scale.

Our joint analysis of skill, scale, and value added contributes to the literature along several dimensions. First, we quantify how many funds create value and determine whether they do so with more profitable or more scalable ideas. Second, we examine whether funds create more value over time as investors learn about skill and scale. Third, we measure how far the value added of each fund is from its optimal value predicted by skill and scale. Finally, we examine whether the industry deliver negative alphas to investors because it is populated by "charlatans" without investment skills, or by funds that scale their size too far.

[1] A non-exhaustive list of papers on performance includes Barras, Scaillet, and Wermers (2010), Carhart (1997), Elton et al. (1993), Harvey and Liu (2018), Jensen (1968), Kosowski et al. (2006), Pastor and Stambaugh (2002), Roussanov, Ruan, and Wei (2020), and Wermers (2000).

[2] Active funds create value when they buy undervalued stocks by (i) using their superior information, (ii) correcting for mispricing caused by noise traders, and (iii) providing liquidity to forced sellers.

To address these issues, we introduce two key innovations. First, we develop a new measure of value added. The standard measure proposed by BvB is equal to the average value created by the fund over its entire life. As such, it averages across each size level at which the fund operates. In contrast, our new measure is defined as the value added when the fund operates at its average size. To the extent that the average captures the size of the fund as it gets older, our measure provides a long-run estimate of the value added. Comparing both measures allows us to control for the size adjustment period, which arises naturally if investors need time to learn about skill and scale and optimize their fund allocation (Pastor and Stambaugh (2012)).

Second, we develop a new estimation approach to infer the cross-sectional distributions of skill, scale, and value added. This approach allows us to determine the number of funds with positive value added, and incorporate the suspected vast heterogeneity in skill and scale. The estimation of each distribution is flexible and bias-free. It is flexible because we use a nonparametric approach which does not require to specify the shape of the true distribution. This flexibility is essential because misspecification risk is large—a joint specification of skill, scale, and value added is a daunting challenge for which theory offers little guidance. The estimation is also bias-free because we explicitly control for the Error-in-Variable (EIV) bias. This bias arises because we use estimated measures (instead of the true ones) to infer each distribution.[3]

In our baseline specification, we follow Berk and Green (2004) and model the fund's gross alpha as $\alpha_{i,t} = a_i - b_i q_{i,t-1}$, where $q_{i,t-1}$ is the lagged fund size. This choice provides simple measures of skill and scale that are specific to each fund. Skill is measured with a_i—the gross alpha on the first dollar invested in the fund. Scale is measured with b_i—the coefficient that captures diseconomies of scale. We then estimate the average value added as $va_i = E\left[\alpha_{i,t} q_{i,t-1}\right] = E\left[(a_i - b_i q_{i,t-1}) q_{i,t-1}\right]$ (the standard measure of BvB), and the long-run (lr) value added as $va_i^{lr} = E\left[\alpha_{i,t}\right] E[q_{i,t-1}] = E\left[a_i - b_i q_{i,t-1}\right] E[q_{i,t-1}]$ (our new measure). To compute the different measures, we use the four-factor model of Cremers, Petajisto, and Zitzewitz (2012). This index-based model includes the SP500 and Russell indices which are tradable and widely used as benchmarks.

Our analysis of US equity funds over the period 1975-2018 uncovers several new insights about skill and scale. First, skill is widespread and economically large—the first dollar alpha is positive for 87.1% of the funds and equal to 3.2% per year on average. Second, funds are highly sensitive to diseconomies of scale—on average, a one

[3] This bias is reminiscent of the well-known EIV bias in the two-pass regression in which we use the estimated betas instead of the true betas (e.g., Jagannathan, Skoulakis, and Wang (2013), Kan, Robotti, and Shanken (2013), Shanken (1992)).

standard deviation increase in size reduces the gross alpha by 1.5% per year. Third, skill and scale vary substantially, both in the whole population and within investment categories. The cross-sectional volatility of a_i and b_i is typically larger than the mean—a finding that contradicts the extensively-used panel regression which assumes that scale is constant ($b_i = b$).[4] Fourth, skill and scale are strongly positively correlated. In other words, great investment ideas are difficult to scale up.

These results shape the cross-sectional distribution of the average value added. In line with their investment skills, 60.9% of the funds create value over the sample period. This result helps reconcile the seemingly puzzling evidence in BvB showing that the average value added is positive, whereas the majority of funds destroy value. This discrepancy arises because we adjust for the EIV bias. Intuitively, the unadjusted distribution is contaminated by estimation noise, which biases the fund proportion estimators and largely inflates the tail probabilities. Our improved estimation has therefore implications for the debate on the size of the active industry because funds that destroy value are unambiguously too large.

The evidence is even stronger with our new long-run measure of value added as we uncover systematic variation in size. Whereas fund size is initially small, its ending value is generally close to the average size. Therefore, our measure captures the value added at the end of the sample period. The long-run measure is well above the standard measure—on average, the gap equals \$6.4M per year (\$7.8M vs \$1.4M). This positive gap is consistent with theory—the standard value added is smaller because it averages over times when size and value created are particularly low, i.e., we have $va_i^{lr} - va_i = b_i V[q_{i,t-1}] > 0$, where $V[q_{i,t-1}]$ is the variance of $q_{i,t-1}$. This expression affords a simple interpretation in the context of models with learning effects (Berk and Green (2004), Pastor and Stambaugh (2012)). As investors learn about skill and scale, the amount they want to invest varies over time. Combined with strong diseconomies of scale, such variations can produce a large difference between va_i^{lr} and va_i.

With correlated skill and scale, funds with the highest value added are not those with the best investment ideas. Instead, their investment strategies balance skill and scale—their first dollar alphas are slightly above average, while their size coefficients are slightly below average. We observe the same trade-off across investment categories. For instance, small cap funds buy illiquid stocks which likely implies both higher mispricing and trading costs. With this particular skill-scale combination—high a_i and high b_i—, we find that small cap funds create more value than large cap funds.

[4] An non-exhaustive list includes Chen et al. (2004), Pastor, Stambaugh, and Taylor (2015), Yan (2008), and Zhu (2018).

Our results show that only a minority of funds destroy value. However, they do not imply that the size of the fund industry reflects sensible market forces (Cochrane (2013)) or, more formally, that the value created is consistent with a rational model for fund investment. To address this issue, we examine the equilibrium predictions of the Berk and Green model in which funds have bargaining power over investors and thus set their size to maximize profits, or, equivalently, the value added. Consistent with this prediction, we find that, in the long run, funds extract 79.5% of the optimal value added va_i^*. In contrast, Zhu (2018) finds a proportion below 1% after estimating va_i^* with a panel specification for log size ($b_i = b$). This striking difference arises because we focus on the long-run value added instead of the average value over the entire period—the latter can be far from optimal if size changes over time. In addition, a panel specification with a homogeneous scale coefficient yields inflated values for the optimal value (i.e., $E[va_i^*(b_i = b) > E[va_i^*])$. Therefore, accounting for the heterogeneity in skill and scale is essential to uncover the ability of the Berk and Green model to fit the data.

Through the value creation process, active funds make asset prices more efficient and contribute to the allocation of resources in the economy. For instance, their trading activity in secondary markets are likely to reduce the cost of capital in primary markets (Pedersen (2018)) and produce real effects on the economy (Bond, Edmans, and Goldstein (2012)). As a result, the fund industry performs a social function, regardless of how funds and investors share the value added. Interestingly, our results have offsetting effects on market efficiency. On the one hand, the proximity of the value added to its optimal value suggests that funds ration the amount of active management. Therefore, prices should be less efficient than in a fully competitive market for asset management. On the other hand, the correlation between skill and scale implies that the industry is not heavily concentrated—the top 5% of the funds only capture 29.1% of the total value added. This result tends to increase market efficiency because funds are too small to internalize the impact on their trading activities on aggregate mispricing.

An important question for investors is whether they benefit from the value created by funds. Our new approach combined with the index-based version of the four-factor model produces a more optimistic performance evaluation than previous studies. However, we still find that 61.8% of the funds exhibit negative net alphas. Consistent with the concern that low-skilled funds exploit unsophisticated investors (e.g., Gruber (1996)), we find that the worst performing funds mostly belong to the group of charlatans (Berk and van Binsbergen (2019)), i.e., funds without investment ideas ($a_i < 0$). However, these funds only represent 12.9% of the population. In other words, a sizable number of funds exhibit both a positive value added and a negative alpha. A simple explanation in the context of the Berk and Green model is that investors tend to overestimate mutual fund skill. An alternative explanation is that information frictions prevent investors from searching for cheaper funds when they pay excessive fees (Roussanov, Ruan, and Wei (2020)). If these mechanisms are at play, the high proportion of underperforming funds should decrease over time as investors sharpen their views and technological progress reduces information frictions.

The remainder of the paper is as follows. Section II presents our framework for measuring skill, scale, and value added. Section III describes our nonparametric approach. Section IV presents the mutual fund dataset. Section V contains the empirical analysis, and Section VI concludes. The appendix provides additional information regarding the methodology, the data, and the empirical results.

II The Measures of Skill, Scale, and Value Added

A Skill and Scale

We consider a population of n funds, where each fund is denoted by the subscript i ($i = 1, ..., n$). To measure fund's skill and scale, we use the specification proposed by BvB. For each fund, the total benchmark-adjusted revenue from active management is given by $TR_{i,t} = a_i q_{i,t-1}$, where $q_{i,t-1}$ denotes the lagged fund size. The total cost of trading is modeled as a convex function of fund size, $TC_{i,t} = b_i q^2_{i,t-1}$. Taking the difference $TR_{i,t} - TC_{i,t}$ and dividing by $q_{i,t-1}$, we obtain a linear specification for the fund gross alpha,

$$\alpha_{i,t} = a_i - b_i q_{i,t-1}, \tag{1}$$

which varies over time in response to changes in fund size.

We measure skill using the first dollar (fd) alpha a_i. This coefficient isolates the profitability of the fund's investment ideas by measuring the gross alpha when $q_{i,t-1} = 0$. As noted by Perold and Salomon (1991), we can interpret a_i as a "paper" return that is unencumbered by the drag of real world implementation.[5] We measure scale using the size coefficient b_i. This coefficient gauges the sensitivity of the gross alpha to changes in fund size. In other words, b_i captures the diseconomies of scale when the fund deploys more capital on its investment ideas.

A key feature of our framework is that we allow both a_i and b_i to be fund specific. To do so, we treat a_i and b_i not as fixed parameters, but as random realizations from the

[5] Several papers find that the gross alpha varies with business cycle indicators z_{t-1} (e.g., Avramov, Barras, and Kosowski (2013), Kacperczyk, van Nieuwerburgh, and Veldkamp (2014)). In this case, we can simply interpret a_i as the fd alpha measured at the average value $E[z_{t-1}]$.

cross-sectional skill distributions $\phi(a)$ and $\phi(b)$.[6] This contrasts with previous studies which typically impose restrictions on Equation (1). For instance, it is common to use a panel specification, which assumes homogeneity of the scale coefficient, i.e., $b_i = b$ (e.g., Chen et al. (2004), Pastor, Stambaugh, and Taylor (2015), Yan (2008), Zhu (2018)).

Whereas this pooling assumption reduces estimation errors, it raises several issues. First, it is unclear from an economic perspective why scale should be constant because funds invest and trade differently. Consistent with this view, the panel specification is strongly rejected in our sample (as discussed in Section V.A). Second, the estimation of the panel regression (with fund fixed effects) is subject to the standard incidental parameter problem (Neymann and Scott (1948), Nickel (1981)). In such a setting, we need to apply a debiasing method like recursive demeaning (e.g., Hjalmarsson (2010), Moon and Philipps (2000)) to obtain an asymptotically unbiased pooled estimate $\hat{b}$. In our case, this recursive demeaning method is unnecessary because we depart from the panel approach (i.e., we work on a fund-by-fund basis).

B Value Added

Next, we use the value added to assess the economic value associated with skill and scale. The value added determines the value that the fund extracts from capital markets. It is equal to the product of the fund's gross alpha and its size: $va_{i,t} = \alpha_{i,t} q_{i,t-1}$. As such, the value added has an intuitive interpretation—if the fund has bargaining power over investors (Berk and Green (2004)), it is similar to the profits of a monopolist measured as the markup price of the good multiplied by the total quantity sold.

We consider two different measures of value added. First, we use the standard measure initially proposed by BvB which determines the average value created by the fund over the sample period:

$$va_i(a_i, b_i) = E\left[\alpha_{i,t} q_{i,t-1}\right] = a_i E[q_{i,t-1}] - b_i E[q_{i,t-1}^2], \tag{2}$$

where $E[q_{i,t-1}]$ and $E[q_{i,t-1}^2]$ denote the time-series averages of the fund size and its squared value.

Second, we introduce a new measure which determines the value added when the fund size is equal to the average $E[q_{i,t-1}]$. As such, it departs from the standard measure which averages across all the size levels at which the fund operates. To the extent that the ending value of $q_{i,t-1}$ is close to $E[q_{i,t-1}]$, our new measure captures the value added at the end of the sample period. To simplify the exposition and highlight the long-term

[6] To lighten notation, we do not subscript the density ϕ by the skill measure.

nature of this measure, we denote it by va_i^{lr} (where lr stands for long-run):

$$va_i^{lr}(a_i, b_i) = E\left[\alpha_{i,t}\right] E\left[q_{i,t-1}\right] = a_i E[q_{i,t-1}] - b_i E[q_{i,t-1}]^2. \tag{3}$$

The two measures in Equations (2) and (3) are fund-specific and inherit the randomness of a_i and b_i. Therefore, we treat them as random realizations from the cross-sectional distributions $\phi(va_i)$ and $\phi(va_i^{lr})$.

Comparing va_i and va_i^{lr} allows us to control for changes in fund size which occur naturally during the investors' learning process. When investors do not know a_i and b_i, they must learn about them using past returns (Pastor and Stambaugh (2012)). As a result, the amount of money they are willing to invest at each point in time may be quite different from the value at the end of the sample period.

It is tempting to estimate how much value the fund creates relative to its total size. This measure, which is obtained by dividing va_i^{lr} by $E[q_{i,t-1}]$, is nothing else that the average gross alpha $\alpha_i = E[\alpha_{i,t}]$. However, $E[\alpha_{i,t}]$ is a noisy measure of value creation precisely because it does not control for the cross-sectional variation in fund size (see BvB). This variation arises because funds have different levels of skill and scale, and choose different fee setting policies. For instance, Habib and Johnsen (2016) argue that funds have a preference for low fees because it allows them to manage a large asset base and mitigate several institutional constraints.[7] Therefore, using the gross alpha is akin to measuring the monopolist rent with the markup price of the goods, regardless of how much quantity is sold.

C Remarks about the Specification

Our baseline specification of the gross alpha in Equation (1) calls for some comments. First, the estimation of a_i and b_i does not require to explicitly model the determinants of skill and scale across funds. For instance, skill can vary because some funds are run by extremely talented managers or benefit from a high speed of information dissemination within their family (Cici, Jaspesen, and Kempf (2017)). Similarly, scale can vary because some funds trade more efficiently or follow specific strategies.[8] To formalize these relationships, we can simply interpret $a_i = g_{a_i}(a_i^f, a_i^m, d_i)$ and $b_i = g_{b_i}(b_i^f, b_i^m, d_i)$ as functions of the skill/scale parameters of the fund family (a_i^f, b_i^f), the skill/scale

[7] Specifically, the Investment Company Act imposes diversification rules which prevent funds from exhausting their investment opportunities if they are too small. Holding a portion of the portfolio passively managed also allows funds to hide their informed trades and obtain better prices.

[8] For instance, Dimensional Fund Advisors (DFA) highlights its ability to minimize the costs of trading small-cap stocks by buying large share blocks from sellers in need of liquidity (Cohen (2002)). Its scale should therefore reflect its unique trading approach, as well as its specific strategy (small-cap stocks).

parameters of the fund manager (a_i^m, b_i^m), and the vector d_i which captures the specific characteristics of the fund strategy such as liquidity and turnover (Pastor, Stambaugh, and Taylor (2020)).

Second, the linear specification in Equation (1) offers several advantages for our analysis. It provides simple measures a_i and b_i to study skill and scale and examine how they affect the value added. It also provides a simple closed-form expression of the optimal value added to benchmark the value added observed in the data. However, it is straightforward to estimate the value added in a more general context than Equation (1). To address concerns about misspecification, we therefore conduct an extensive analysis based on a log function $b_i \log(q_{i,t-1})$, and a fully flexible function $g_i(q_{i,t-1})$ as proposed by BvB. This robustness analysis reveals that the empirical results remain unchanged (see Section V.F).

III Methodology

A Motivation for the Nonparametric Approach

We now describe our novel nonparametric approach for estimating the cross-sectional distribution $\phi(m)$, where $m \in \{a, b, va, va^{lr}\}$ encompasses all four measures presented above. Our nonparametric approach imposes minimal structure on the true density ϕ and thus departs from a standard parametric/Bayesian approach which requires a full specification of the shape of ϕ. The choice between a parametric and nonparametric approach involves the usual trade-off between efficiency and misspecification. If the structure imposed by the parametric approach is correct, the estimated distribution is more precise. However, it can be heavily biased if the imposed structure is wrong.

The analysis of skill, scale, and value added favors a nonparametric approach because the risk of misspecification is large. Whereas theory predicts that performance should cluster around zero, it offers no such guidance for each measure m. In principle, we can gain parametric flexibility by using normal mixture models (e.g., Chen, Cliff, and Zhao (2017), Harvey and Liu (2018)). However, determining the correct number of mixtures is difficult because the parameters are estimated with significant noise (Yan and Cheng (2019)), the numerical optimization of the likelihood is non-standard (van der Vaart (1998; p. 74)), and the statistical inference is technically involved (Chen (2017)).[9] Misspecification risk further increases because we jointly study four measures

[9]For example, the classical theory of the log likelihood test statistic does not hold for testing the number of components in the mixture (e.g., Ghosh and Sen (1985)). Here, the inference is even more complicated because we do not observe the true measures, but only the estimated ones.

(a_i, b_i, va_i, va_i^{lr}). Therefore, a parametric/Bayesian approach involves the daunting task of correctly specifying a multivariate distribution whose marginals are potentially mixtures of distributions.

In addition to its robustness to misspecification, the nonparametric approach brings several benefits. First, its implementation is simple and fast. Intuitively, it is akin to computing a histogram using as inputs the estimated measures for all funds. In contrast, parametric/Bayesian approaches require sophisticated and computer-intensive Gibbs sampling and Expectation Maximization (EM) methods (e.g., Chen, Cliff, and Zhao (2017), Harvey and Liu (2018), Jones and Shanken (2005)).

Second, it provides a unified framework for estimating the density function $\phi(m)$, along with the moments (mean, variance, skewness, kurtosis), the proportion estimator inferred from the cumulative distribution function (cdf) $\Phi(x) = P[m_i \leq x] = \int_{-\infty}^{x} \phi(u)du$, and the quantile $Q(p) = \Phi^{-1}(p)$, where p denotes the probability level.

Third, it comes with a full-fledged inferential theory. We derive the asymptotic distribution of each estimator as the numbers of funds n and return observations T grow large (simultaneous double asymptotics with $n, T \to \infty$). We can therefore conduct proper statistical inference guided by econometric theory.

B Estimation Procedure

B.1 Estimation of the Different Measures

Our nonparametric estimation of the density $\phi(m)$ consists of three main steps. To begin, we estimate skill and scale for each fund i using the following time-series regression:

$$r_{i,t} = \alpha_{i,t} + \beta_i' f_t + \varepsilon_{i,t} = a_i - b_i q_{i,t-1} + \beta_i' f_t + \varepsilon_{i,t}, \tag{4}$$

where $r_{i,t}$ is the fund gross excess return (before fees) over the risk-free rate, f_t is a K-vector of benchmark excess returns, and $\varepsilon_{i,t}$ is the error term. We interpret Equation (4) as a random coefficient model (e.g., Hsiao (2003)) in which the coefficients a_i, b_i, and β_i are random realizations from a continuum of funds. Under this sampling scheme, we can invoke cross-sectional limits to infer the density of each measure m.[10,11]

[10] Gagliardini, Ossola, and Scaillet (2016) use a similar sampling scheme to develop testable applications of the arbitrage pricing theory in a large cross-section of assets (see also Gagliardini, Ossola, and Scaillet (2020) for a review of the literature).

[11] We can also apply our approach to estimate the cross-sectional distribution of the fund beta for each risk factor k ($k = 1, ..., K$), denoted by $\phi(\beta_k)$. As explained below, the common practice of estimating $\phi(\beta_k)$ using the estimated betas is biased because of the error-in-variable (EIV) problem.

The vector of coefficients $\hat{\gamma}_i = \left(\hat{a}_i, \hat{b}_i, \hat{\beta}_i'\right)'$ is computed as

$$\hat{\gamma}_i = \hat{Q}_{x,i}^{-1} \frac{1}{T_i} \sum_{t=1}^{T} I_{i,t} x_{i,t} r_{i,t}, \tag{5}$$

where $I_{i,t}$ is an indicator variable equal to one if $r_{i,t}$ is observable (and zero otherwise), $T_i = \sum_{t=1}^{T} I_{i,t}$ is the number of return observations for fund i, $x_{i,t} = (1, -q_{i,t-1}, f_t')'$ is the vector of explanatory variables, and $\hat{Q}_{x,i} = \frac{1}{T_i} \sum_{t=1}^{T} I_{i,t} x_{i,t} x_{i,t}'$ is the estimated matrix of the second moments of $x_{i,t}$. Using the estimated coefficients along with the size and squared size time-series averages, $\overline{q}_i = \frac{1}{T_i} \sum_{t=1}^{T} I_{i,t} q_{i,t-1}$, $\overline{q}_{i,2} = \frac{1}{T_i} \sum_{t=1}^{T} I_{i,t} q_{i,t-1}^2$, we can then infer each of the four measures as

$$\begin{aligned}
\text{Fd alpha} &: \hat{m}_i = \hat{a}_i, \\
\text{Size coefficient} &: \hat{m}_i = \hat{b}_i, \\
\underset{\text{(average)}}{\text{Value added}} &: \hat{m}_i = \widehat{va}_i = \hat{a}_i \overline{q}_i - \hat{b}_i \overline{q}_{i,2}, \\
\underset{\text{(long run)}}{\text{Value added}} &: \hat{m}_i = \widehat{va}_i^{lr} = \hat{a}_i \overline{q}_i - \hat{b}_i \overline{q}_i^2.
\end{aligned} \tag{6}$$

Our econometric framework formally accounts for the unbalanced nature of the panel of mutual fund returns by means of the observability indicators $I_{i,t}$. Given that the number of observations is small for some funds, the inversion of the matrix $\hat{Q}_{x,i}$ can be numerically unstable and yield unreliable estimates of m_i. To address this issue, we follow Gagliardini, Ossola, and Scaillet (2016) and introduce a formal fund selection rule $\mathbf{1}_i^{\chi}$ equal to one if the following two conditions are met (and zero otherwise):

$$\mathbf{1}_i^{\chi} = \mathbf{1}\left\{CN_i \leq \chi_{1,T}, \tau_{i,T} \leq \chi_{2,T}\right\}, \tag{7}$$

where $CN_i = \sqrt{eig_{max}\left(\hat{Q}_{x,i}\right) / eig_{\min}\left(\hat{Q}_{x,i}\right)}$ is the condition number of the matrix $\hat{Q}_{x,i}$ defined as the ratio of the largest to smallest eigenvalues eig_{max} and eig_{min}, $\tau_{i,T} = T/T_i$ is the inverse of the relative sample size T_i/T, and $\chi_{1,T}$, $\chi_{2,T}$ denote the two threshold values. The first condition $\{CN_i \leq \chi_{1,T}\}$ excludes funds for which the time series regression is poorly conditioned, i.e., a large value of CN_i indicates multicollinearity problems (Belsley, Kuh, and Welsch (2004), Greene (2008)). The second condition $\{\tau_{i,T} \leq \chi_{2,T}\}$ excludes funds for which the sample size is too small. Both thresholds $\chi_{1,T}$ and $\chi_{2,T}$ increase with the sample size T—with more return observations, we estimate the fund coefficients with greater accuracy which allows for a less stringent selection rule. Applying this formal selection rule, we obtain a total number of funds equal to $n_\chi = \sum_{i=1}^{n} \mathbf{1}_i^\chi$.

B.2 Kernel Density Estimation

In the next step, we estimate the skill density function using a standard nonparametric approach based on kernel smoothing.[12] We compute the estimated density $\hat{\phi}$ at a given point m as

$$\hat{\phi}(m) = \frac{1}{n_\chi h} \sum_{i=1}^{n} \mathbf{1}_i^\chi K\left(\frac{\hat{m}_i - m}{h}\right), \tag{8}$$

where h is the vanishing smoothing bandwidth—similar to the length of histogram bars, h determines how many observations around point m we use for estimation. The function K is a symmetric kernel function that integrates to one. Because the choice of K is not a crucial aspect of nonparametric analysis, we use the standard Gaussian kernel $K(u) = \frac{1}{\sqrt{2\pi}} \exp(-\frac{u^2}{2})$ for our empirical analysis (see Silverman (1986)).[13]

The following proposition examines the asymptotic properties of $\hat{\phi}(m)$ as the number of funds n and the number of periods T grow large for a vanishing bandwidth h.

Proposition III.1 *As $n, T \to \infty$ and $h \to 0$ such that $nh \to \infty$ and $\sqrt{nh}(h^2 T + (1/T)^{\frac{3}{2}}) \to 0$, we have*

$$\sqrt{nh}\left(\hat{\phi}(m) - \phi(m) - bs(m)\right) \Rightarrow N\left(0, K_1 \phi(m)\right), \tag{9}$$

where $\Rightarrow$ denotes convergence in distribution, and the bias term $bs(m)$ is the sum of two components,

$$bs_1(m) = \frac{1}{2} h^2 K_2 \phi^{(2)}(m), \tag{10}$$

$$bs_2(m) = \frac{1}{2T} \psi^{(2)}(m), \tag{11}$$

where $K_1 = \int K(u)^2 du$, $K_2 = \int u^2 K(u) du$, $\phi^{(2)}(m)$ is the second derivative of the density $\phi(m)$ and $\psi^{(2)}(m)$ is the second derivative of the function $\psi(m) = \omega(m)\phi(m)$ with $\omega(m) = E[S_i | \, m_i = m]$. The term S_i is the asymptotic variance of the estimated

[12] See, for instance, Ait-Sahalia (1996), and Ait-Sahalia and Lo (1998) for applications of kernel density estimation in finance.

[13] Similar to Equation (8), Okui and Yanagi (2020) consider a kernel estimator for the density of the mean and autocorrelation of random variables (see also Jochmans and Weidner (2018) and Okui and Yanagi (2019) for the analysis of the cumulative distribution function). However, their distributional results differ from our regression-based results aimed at measuring skill, scale, and value added.

centered measure $\sqrt{T}(\hat{m}_i - m_i)$ *equal to* $\underset{T\to\infty}{plim}(\frac{\tau_{i,T}^2}{T}\sum_{t,s=1}^{T} I_{i,t}I_{i,s}u_{i,t}u_{i,s})$. *For each measure, the term* $u_{i,t}$ *is given by*

$$
\begin{aligned}
\textit{Fd alpha} \quad &: \quad u_{i,t} = e_1' Q_{x,i}^{-1} x_{i,t}\varepsilon_{i,t}, \\
\textit{Size coefficient} \quad &: \quad u_{i,t} = e_2' Q_{x,i}^{-1} x_{i,t}\varepsilon_{i,t}, \\
\underset{(average)}{\textit{Value added}} \quad &: \quad u_{i,t} = E[q_{i,t-1}]e_1' Q_{x,i}^{-1} x_{i,t}\varepsilon_{i,t} + a_i(q_{i,t-1} - E[q_{i,t-1}]) \\
&\qquad - E[q_{i,t-1}^2]e_2' Q_{x,i}^{-1} x_{i,t}\varepsilon_{i,t} - b_i(q_{i,t-1}^2 - E[q_{i,t-1}^2]), \\
\underset{(long\ run)}{\textit{Value added}} \quad &: \quad u_{i,t} = E[q_{i,t-1}]e_1' Q_{x,i}^{-1} x_{i,t}\varepsilon_{i,t} + a_i(q_{i,t-1} - E[q_{i,t-1}]) \\
&\qquad - E[q_{i,t-1}]^2 e_2' Q_{x,i}^{-1} x_{i,t}\varepsilon_{i,t} - b_i 2E[q_{i,t-1}](q_{i,t-1} - E[q_{i,t-1}]), \quad (12)
\end{aligned}
$$

where e_1 (e_2) *is a vector with one in the first (second) position and zeros elsewhere and* $Q_{x,i} = E[x_{i,t}x_{i,t}']$. *Under a Gaussian kernel, the two constants* K_1 *and* K_2 *are equal to* $\frac{1}{2\sqrt{\pi}}$ *and* 1, *respectively.*

Proof. See the appendix (Section I.A). ■

Proposition III.1 yields several important insights. First, it shows that the estimated density function $\hat{\phi}(m)$ is asymptotically normally distributed, which facilitates the construction of confidence intervals. As shown in Equation (9), the width of this interval depends on the variance term $K_1\phi(m)$ which is higher in the peak of the density.

Second, $\hat{\phi}(m)$ is a biased estimator of the true density. Therefore, we can improve the estimation by adjusting for the bias term $bs(m)$. Equations (10)-(11) reveal that $bs(m)$ has two distinct components. The first component $bs_1(m)$ is the smoothing bias, which is standard in nonparametric density estimation (e.g., Silverman (1986), Wand and Jones (1995)). The second component $bs_2(m)$, which is referred to as the error-in-variable (EIV) bias, is non-standard in nonparametric statistics—it arises because we estimate ϕ using the estimated measures instead of the true ones (i.e., $\hat{m}_i$ instead of m_i).

Finally, Proposition III.1 provides guidelines for the choice of the bandwidth. We show in the appendix (Section I.B) that the choice of the optimal bandwidth h^*—the one that minimizes the Asymptotic Mean Integrated Squared Error (AMISE) of $\hat{\phi}(m)$—depends on the relationship between T and n: (i) if T is small relative to n ($n^{2/5}/T \to \infty$), h^* is proportional to $(nT)^{-\frac{1}{3}}$; (ii) if T is large relative to n ($n^{2/5}/T \to 0$), h^* is proportional to $n^{-\frac{1}{5}}$.[14] Our Monte-Carlo analysis in the appendix (Section IV) reveals

[14]The AMISE is defined as the integrated sum of the leading terms of the asymptotic variance and squared bias of the estimated density $\hat{\phi}(m)$. By minimising the AMISE, we explicitly control for the

that given our actual sample size, the two bandwidth choices produce similar results with a slight advantage to the first case. Motivated by these results, we use the following bandwidth in our baseline specification:

$$h^* = \left(\frac{K_2}{K_1} \int \phi^{(2)}(m) \psi^{(2)}(m) dm \right)^{-\frac{1}{3}} (n/T)^{-\frac{1}{3}}. \tag{13}$$

B.3 Bias Adjustment

Our final step is to adjust the density estimator $\hat{\phi}(m)$ for the bias. To do so, we apply a Gaussian reference model to compute the two bias terms and the optimal bandwidth in Equations (10)-(11), and (13).[15] Under this model, the fund measure m_i and the log of the asymptotic variance $s_i = \log(S_i)$ follow a bivariate normal distribution where $m_i \sim N(\mu_m, \sigma_m^2)$, $s_i \sim N(\mu_s, \sigma_s^2)$, and $corr(m_i, s_i) = \rho$.

Applying a simple Gaussian reference model has several appealing properties. First, the computation of the bias and the bandwidth is straightforward because they are all available in closed form. Second, the bias terms are precisely estimated because they only depend on the five parameters of the normal distribution $\theta = (\mu_m, \sigma_m, \mu_s, \sigma_s, \rho)'$. Third, the analysis of the closed-form expressions allows us to shed light on (i) the determinants of the bias, and (ii) the conditions under which the reference model provides a close approximation of the true bias.

These benefits are not shared by a fully nonparametric approach in which the bias terms are inferred from Equations (10)-(11) via a nonparametric estimation of the second-order derivatives $\phi^{(2)}$ and $\psi^{(2)}$. Estimating these derivative terms is notoriously difficult and generally leads to large estimation errors (e.g., Wand and Jones (1995; Ch. 2)).[16] Similarly, the standard bootstrap usually underestimates the bias in curve estimation problems (Hall (1990), Hall and Kang (2001)). The design of resampling techniques suitable for our unbalanced setting with an EIV problem is a difficult and still open question.

The following proposition derives the closed-form expressions for the two bias components and the optimal bandwidth under the Gaussian reference model as the number

trade-off between the bias and the variance over the entire support of the density. Therefore, we avoid overfitting the data (undersmoothing) because we do not choose a bandwidth that is too small.

[15] A Gaussian reference model underlies the celebrated Silverman rule for the optimal choice of the bandwidth in standard non-parametric density estimation without the EIV problem. This rule gives $h^* = 1.06\sigma n^{-\frac{1}{5}}$, where σ is the standard deviation of the observations (Silverman (1986)).

[16] 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.

of funds n and the number of periods T grow large for a vanishing bandwidth h.

Proposition III.2 *As $n, T \to \infty$ and $h \to 0$ such that $nh \to \infty$ and $\sqrt{nh}(h^2T + (1/T)^{\frac{3}{2}}) \to 0$, the two bias components under the reference model are equal to*

$$bs_1^r(m) = \left[\frac{1}{2}K_2h^2\frac{1}{\sigma_m^2}(\bar{m}_1^2 - 1)\right]\frac{1}{\sigma_m}\varphi(\bar{m}_1), \quad (14)$$

$$bs_2^r(m) = \left[\frac{1}{2T}\exp(\mu_s + \frac{1}{2}\sigma_s^2)\frac{1}{\sigma_m^2}(\bar{m}_2^2 - 1)\right]\frac{1}{\sigma_m}\varphi(\bar{m}_2), \quad (15)$$

where $\bar{m}_1 = \frac{m - \mu_m}{\sigma_m}$, $\bar{m}_2 = \frac{m - \mu_m - \rho\sigma_m\sigma_s}{\sigma_m}$, and $\varphi(x) = \frac{1}{\sqrt{2\pi}}\exp(-\frac{1}{2}x^2)$ is the density of the standard normal distribution. In addition, the optimal bandwidth h^* is given by

$$h^* = \left[\frac{K_2}{K_1 2\sqrt{\pi}}\frac{3}{4\sigma_m^5}\left(\frac{\rho^4\sigma_s^4}{12} - \rho^2\sigma_s^2 + 1\right)\exp\left(\mu_s + \frac{1}{2}\sigma_s^2(1 - \frac{\rho^2}{2})\right)\right]^{-\frac{1}{3}}(n/T)^{-\frac{1}{3}}. \quad (16)$$

Proof. See the appendix (Section I.C) ■

Equations (14)-(15) imply that the smoothing bias is negligible, whereas the EIV bias is not. As the total number of funds n increases, h^* shrinks towards zero, which reduces the magnitude of $bs_1^r(m)$. With a population of several thousand funds, the smoothing term becomes negligible for all values of m. In contrast, $bs_2^r(m)$ depends on the number of observations T because it arises from the noise gap between $\hat{m}_i$ and m_i. Therefore, the EIV bias remains significant even if the fund population is large. Intuitively, $bs_2^r(m)$ removes probability mass from the tails of the unadjusted density $\hat{\phi}(m)$ as a result of the estimation noise introduced by $\hat{m}_i$. It also leads to an asymmetric adjustment when m_i and s_i are correlated ($\rho \neq 0$). The appendix (Section III) contains a detailed analysis of the bias and a comparative static analysis of $bs_2^r(m)$ for different values of the variances of $\hat{m}_i$ and m_i, and the correlation ρ.

Using the results in Proposition III.2, we can compute the bias-adjusted density $\tilde{\phi}(m)$. We estimate the parameter vector θ using the estimated quantities $\hat{m}_i$ and $\hat{s}_i$ ($i = 1, ..., n_\chi$). To compute $\hat{s}_i = \log(\hat{S}_i)$, we use the standard variance estimator of Newey and West (1987):

$$\hat{S}_i = \frac{\tau_{i,T}^2}{T}\sum_{t=1}^{T} I_{i,t}\hat{u}_{i,t}^2 + 2\sum_{l=1}^{L}\left(1 - \frac{l}{L+1}\right)\left[\frac{\tau_{i,T}^2}{T}\sum_{t=1}^{T-l} I_{i,t}I_{i,t+l}\hat{u}_{i,t}\hat{u}_{i,t+l}\right], \quad (17)$$

where $\hat{u}_{i,t}$ is obtained by plugging the estimated quantities for the chosen measure in Equation (12), and L is the number of lags to capture potential serial correlation. Then, we plug the elements of the estimated vector $\hat{\theta}$ into Equations (14)-(16) to compute the bias terms $\widehat{bs}_1^r(m)$, $\widehat{bs}_2^r(m)$, and the optimal bandwidth h^*. Finally, we remove the bias terms from the raw density in Equation (8) to obtain the bias-adjusted density estimator

$$\tilde{\phi}(m) = \hat{\phi}(m) - \widehat{bs}_1^r(m) - \widehat{bs}_2^r(m). \quad (18)$$

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 this 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 $bs_2^r(m)$.[17] Second, our extensive Monte-Carlo analysis calibrated on the data reveals that the bias-adjusted density captures the true density remarkably well (see the appendix (Section IV)).[18]

In the appendix, we also derive the asymptotic properties of the other characteristics of the distribution (moments, proportion, and quantile). Similar to the density, we show that the different estimators are normally distributed and suffer from the EIV bias (except the mean because of its intrinsic linear nature). To compute the bias-adjusted characteristics, we can use the analytical expressions derived in the appendix (Section II). Alternatively, we can compute them via a numerical integration of the bias-adjusted density $\tilde{\phi}(m)$. For instance, the estimated proportion of funds with a negative m_i is given by the cdf estimate $\tilde{\pi}^- = \int_{-\infty}^{0} \tilde{\phi}(u)du$. Whereas both approaches (analytical and numerical) are asymptotically equivalent, the Monte-Carlo analysis reveals that the numerical integration generally produces a lower Mean Squared Error (MSE).[19] Motivated by these results, we use this approach in the empirical section of the paper.

[17] The two terms $bs_2(m)$ and $bs_2^r(m)$ only differ if ϕ is a mixture of distributions whose components have means extremely far away from one another. In this case, we have a trough instead of a peak around the mean.

[18] 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.

[19] Following Equation (18), we use a tilde to denote all the estimated quantities that are bias-adjusted.

IV Data Description

A Mutual Fund Data and Benchmark Model

We conduct our analysis on the entire population of open-end actively managed US equity funds. 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 2018. This allows us to construct the time-series of the gross return and size for the population and different groups (small/large cap, low/high turnover). Similar to BvB, we adjust fund size for inflation by expressing all numbers in January 1, 2000 dollars.

To estimate the regression for each fund in Equation (4), 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 S&P500 (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 S&P500, and between the Russell 3000 Value and Russell 3000 Growth.[20]

The motivation for using this model is that it correctly assigns a zero alpha to the S&P500 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 year over the period 1980-2005. Therefore, small cap funds that use this index as a benchmark are likely to be classified as unskilled under the Carhart model.

To apply the fund selection rules in Equation (6), we follow Gagliardini, Ossola, and Scaillet (2016) and select funds for which the condition number of the matrix of regressors $\hat{Q}_{x,i}$ is below 15 and the number of monthly observations is above 60 ($CN_i \leq 15$ and $\tau_{i,T} \leq 8.8$). These selection criteria produce a final universe of 2,106 funds. The appendix (Section V) provides more detail on the construction of the mutual fund dataset.

[20] 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.

B Summary Statistics

Table I reports summary statistics for our mutual fund sample. At the start of each month, we construct an equal-weighted portfolio of all funds in the population and for each investment category (small/large cap and low/high turnover). In Panel A, we report the first four moments of the portfolio gross excess returns. In the entire population, the portfolio achieves a risk-return tradeoff similar to that of the aggregate stock market (9.0% and 15.3% per year). It also exhibits a negative skewness (-0.74) and a positive kurtosis (5.33). The results are similar across groups, except for the small cap portfolio which produces higher mean and volatility.

In Panel B, we report the estimated portfolio betas for the four factors. We find that small cap funds are heavily exposed to the size factor (0.79), which is also the case for high turnover funds (0.46). Finally, Panel C reports additional characteristics which include the average number of funds in the portfolio and the time-series average of the median fund size, fees, and turnover. Consistent with intuition, small cap funds manage a smaller asset base—the median size is equal to $143M versus $273M for large cap funds. We also find that high turnover funds trade very aggressively. The median turnover reaches 131% per year versus 37% for low turnover funds.

Please insert Table I here

V Empirical Results

A Skill and Scale

A.1 Overview of the Distributions

We begin our empirical analysis by examining the cross-sectional distributions of skill and scale. For each fund, we estimate a_i and b_i (Equation (12)) and use our nonparametric approach to infer the densities $\phi(a)$ and $\phi(b)$. To describe the properties of each distribution, we compute the bias-adjusted estimates of the moments (mean, variance, skewness, kurtosis), the proportions of funds with negative and positive measures denoted by $\tilde{\pi}^-$ and $\tilde{\pi}^+$, and the quantiles at 5% and 95% denoted by $\tilde{Q}_{0.05}$ and $\tilde{Q}_{0.95}$. We also compute the standard deviation of each estimator (see the appendix (Section II) for the details). We show the summary statistics for $\tilde{\phi}(a)$ in Panel A of Table II. In addition, Panel B reports the median characteristics of funds with highly significant $\hat{a}_i$ (we choose a significance level equal to 0.05).

Panel A provides ample evidence that mutual funds are skilled—on average, the first dollar (fd) alpha equals 3.2% per year and is positive for 87.1% of the funds in the population. This result resonates with the numerical analysis of Berk and Green (2004) who find a similar proportion (80%) based on a calibration of their model. Therefore, the vast majority of funds invest in undervalued stocks based on their private information or their ability to correct the mispricing caused by noise traders.

A small number of funds are "charlatans", i.e., funds that have no profitable investment ideas ($a_i < 0$). It is a priori surprising that charlatans exist because funds always have the option to invest passively (such that $a_i = 0$). Berk and van Binsbergen (2019) suggest that such funds trade actively in order to mislead investors about their skill. Consistent with this view, Panel B shows that funds with negative and significant $\hat{a}_i$ trade aggressively (the median turnover equals 98% per year).

Please insert Table II here

Next, we conduct the same analysis for the scale distribution $\tilde{\phi}(b)$ in Table III. We find that 86.2% of the funds experience diseconomies of scale. The magnitude of the size coefficient is typically large—on average, a one standard deviation increase in size reduces the gross alpha by 1.5% per year. Equivalently, a \$100M increase in size lowers the gross alpha by at least 0.2% per year.[21] Overall, the results are largely consistent with models that emphasize the importance of capacity constraints for mutual funds (e.g., Berk and Green (2004), Pastor and Stambaugh (2012)).

At the same time, a minority of funds benefit from economies of scale ($b_i < 0$)—a view that is inconsistent with the theoretical predictions of the Berk and Green model. A plausible explanation for this result is simply statistical luck. We see that the negative portion of $\tilde{\phi}(b)$ remains close to zero ($\tilde{Q}_{0.05} = -0.6\%$). In addition, Panel B shows that (i) only 17 funds have negative and significant $\hat{b}_i$, and (ii) these funds are all classified as "false discoveries" by the False Discovery Rate (FDR) approach, i.e., funds with significant $\hat{b}_i$, whereas their true coefficient b_i equals zero (see Barras, Scaillet, and Wermers (2010)). An alternative explanation is that our linear model is incorrect for these particular funds. To address this concern, we consider several alternative specifications presented in Section V.F.

Please insert Table III here

[21] To obtain this lower bound, we use the Jensen inequality: $E[\frac{b_i}{\sigma_{qi}}] > \frac{E[b_i]}{E[\sigma_{qi}]} = \frac{\bar{b}}{\bar{\sigma}_q}$, where $\bar{b}$ is the average size coefficient, and $\bar{\sigma}_q$ is the average volatility of fund size (i.e., time-series volatility averaged across funds). With $\bar{b} = 1.5\%$ and $\bar{\sigma}_q = \$650$M, we obtain $\frac{\bar{b}}{\bar{\sigma}_q}100 = 0.2\%$.

A.2 Cross-Sectional Variation in Skill and Scale

An important insight from Tables II and III is the substantial variation of skill and scale. Some funds have stellar investment skills—5% of them exhibit a fd alpha above 9.0% per year, which is 2.8 times larger than the average. Similarly, funds largely differ in scale as the cross-sectional standard deviation of the size coefficient is larger than the average. This cross-sectional variation is potentially explained by different investment styles. To examine this issue, we estimate the distributions of skill and scale among funds with different levels of liquidity (small/large cap) and turnover (low/high turnover)—two key determinants of the fund investment strategy (Pastor, Stambaugh, and Taylor (2020)).

Table IV confirms that skill and scale vary across all four investment categories. The average values of the fd alpha and size coefficient vary between 1.9% and 4.9% per year, and between 1.0% and 2.0% per year. However, we still observe substantial cross-fund variation within each category. This implies that skill and scale are determined by fund-specific characteristics that go beyond the investment style.

Please insert Table IV here

The large cross-sectional variation in scale is at odd with the panel regression approach that imposes a constant size coefficient across all funds ($b_i = b$). To formally test the validity of this panel approach, we consider the linear specification $\alpha_{i,t} = a_i - bq_{i,t-1}$ used in previous work (e.g., Chen et al. (2004), Pastor, Stambaugh, and Taylor (2015), Yan (2008), Zhu (2018)). We also examine the log specification $\alpha_{i,t} = a_i - b\log(q_{i,t-1})$ based on the assumption that a relative (instead of absolute) size change has the same effect on all funds (e.g., Yan (2008), Zhu (2018)).[22] Then, we test the null hypothesis of homogeneous coefficients $H_0 : b_i = b$ $(i = 1, ..., n)$ using the procedure proposed by Pesaran and Yagamata (2008) (see the description in the appendix (Section VI.A)). For each specification (size, log size), we find that H_0 is rejected with probability one both in the population and within each category. This massive rejection is not surprising given the huge heterogeneity in the size coefficient observed in the data (Tables III and IV). It also implies that forming broad categories of funds is not sufficient to support the panel regression approach.[23]

[22] In the logarithmic specification, the intercept corresponds to the net alpha when $q_{i,t-1}$ is equal to 1, and hence depends on the measurement unit—different units (e.g, \$1 or \$1M) imply different interpretations of a "first dollar" alpha. The invariance to how we denominate size is an advantage of the linear specification.

[23] The large variation in scale across funds also helps to understand the statistically weak and conflicting evidence on the effect of size using panel information (e.g., Chen et al. (2004), Pastor, Stambaugh, and Taylor (2015), Zhu (2018)). When b_i varies across funds, the standard deviation of the estimated

A.3 Correlation between Skill and Scale

Our cross-sectional analysis also reveals that skill and scale are strongly correlated. In the entire population, the cross-sectional correlation between $\hat{a}_i$ and $\hat{b}_i$ is equal to 0.82. Put differently, great investment ideas are difficult to scale up.[24] Part of this correlation is explained by the fund's investment style—as discussed in Section II.C, a_i and b_i are correlated because they both depend on the characteristics of the fund strategy such as liquidity and turnover.

To illustrate, Figure 1 (Panels A and B) shows that small cap funds have both higher fd alphas and size coefficients than large cap funds. These results are consistent with the difference in liquidity between the two groups. Illiquidity tends to increase the mispricing of small cap stocks (higher a_i)—as noted by Hong, Lim, and Stein (2000), these stocks are largely untouched by mutual funds. At the same time, illiquidity increases the cost of trading small cap stocks (higher b_i).

We document a similar pattern for high versus low turnover funds (Panels C and D). By rebalancing their portfolio more often, high turnover funds are able to exploit a larger number of investment opportunities (higher a_i). However, they also incur higher trading costs (higher b_i), possibly as a result of excessive trading (e.g., Dow and Gorton (1997)). Overall, these results imply that the ability of funds to generate large profits depends on the trade-off between skill and scale—a point we examine in more detail in Sections V.D and V.E.[25]

Please insert Figure 1 here

B Average Value Added

B.1 Overview of the Distribution

We begin our analysis of the value added with the standard measure of BvB which captures the average value over the sample period. For each fund, we estimate va_i as

pooled $\hat{b}$ is inflated and its t-statistic decreases (Pesaran and Yagamata (2008)). Therefore, $\hat{b}$ may not be statistically significant even if most funds have a positive b_i (as shown in Table III).

[24] This correlation is cross-sectional and is therefore not driven by the fund-level correlation between $\hat{a}_i$ and $\hat{b}_i$. To verify this point, we run a simple simulation where we impose that the cross-sectional correlation between a_i and b_i is equal to zero (see the appendix (Section IV.C) for details). Consistent with theory, we find that the correlation between $\hat{a}_i$ and $\hat{b}_i$ is positive at the fund level, but the correlation between $\hat{a}_i$ and $\hat{b}_i$ across funds equals zero.

[25] The strong correlation between a_i and b_i also 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.

a function of a_i and b_i (Equation (12)) and use our nonparametric approach to infer the cross-sectional density $\phi(va)$. Panel A of Table V reports the summary statistics for $\tilde{\phi}(va)$, and Panel B shows the median characteristics of funds with highly significant $\widehat{va}_i$ (at a significance level of 0.05).

Panel A reveals that the majority of funds extract value from capital markets by exploiting their investment skills. We find that 60.9% of the funds produce a positive value added which, on average, equals \$1.4M per year. In contrast, BvB and Zhu (2018) find that the majority of funds destroy value. This difference is important for the debate on the size of the finance industry (e.g., Cochrane (2013), Greenwood and Scharfstein (2013)). Funds that destroy value are unambiguously too large—if their proportion in the population is large, it raises the bar for any theory that tries to rationalize the actual size of the fund industry.

A distinguishing feature of our estimator $\tilde{\pi}^+$ is that it controls for the EIV bias. Figure 2 shows that the bias-adjusted distribution $\tilde{\phi}(va_i)$ departs markedly from the unadjusted distribution $\hat{\phi}(va_i)$, which is distorted by (i) the estimation noise contained in $\widehat{va}_i$ and (ii) its correlation with the true value va_i (as captured by ρ in Equation (15)). As a result, failing to adjust for the EIV bias materially distorts the estimated values. We find that the unadjusted proportion of positive-value fund drops to 46.2% (similar to the values reported in BvB and Zhu (2018)). In addition, extreme observations heavily influence the unadjusted quantiles, i.e., their spread is 1.3 times larger than the one reported in Table V (\$27.1M versus \$35.6M).[26]

We can decompose the minority of funds with a negative value added into two groups. The first one includes charlatans ($a_i < 0$), and the second one includes skilled funds that grow too large to maintain revenues below costs ($a_i < b_i q_{i,t-1}$). Contrary to charlatans, these funds can potentially create value in the long run if they scale down their size. Panel B shows that only 38.8% of the funds with highly significant negative $\widehat{va}_i$ have a negative estimated fd alpha. In other words, value destruction is primarily caused by funds that grow too large, and not by charlatans.

Please insert Table V and Figure 2 here

B.2 The Dynamics of Fund Size

The standard measure va_i captures the average value added over the entire sample period. Therefore, it averages over times when the fund size can be very small or large.

[26] More generally, we must interpret boxplots provided by statistical packages with caution if each datapoint is an estimate instead of a true observation because the EIV bias plagues them inevitably.

These size fluctuations arise naturally if investors must learn about skill and scale—as they update their priors about a_i and b_i, the amount they are willing to invest in the fund changes.

To examine the dynamics of size across funds, we split the set of size observations for each fund into 10 periods (1=start of the period, 10=end of the period). For each time period τ ($\tau = 1, ..., 10$), we then compute the difference $\Delta q_{i,\tau} = \sum_{t\in\tau} q_{i,t-1} - \overline{q}_i$, where $\overline{q}_i$ is the average size.[27] Figure 3 plots the median value of $\Delta q_{i,\tau}$ for each time period. Consistent with intuition, the size is generally substantially below its average when the fund is young. In period 1, the median size gap equals -\$195M which represents -83.3% of the average fund size. Then, the size reaches its maximum value in period 8 before falling back close to $\overline{q}_i$—in the last period, the median size gap is a mere -\$6.5M (-6.7% in relative terms). Motivated by these results, we now focus on our new measure which measures the value added at the average fund size. By taking the average as the reference point, this measure provides an estimate of the value added in the long run (i.e., at the end of the sample period).

Please insert Figure 3 here

C Long-Run Value Added

C.1 Overview of the Distribution

To study the long-run value added, we estimate va_i^{lr} for each fund as a function of a_i and b_i (Equation (12)), and use our nonparametric approach to infer the cross-sectional density $\phi(va^{lr})$. Panel A of Table VI reports the summary statistics for $\tilde{\phi}(va^{lr})$, and Panel B shows the median characteristics of funds with highly significant $\widehat{va}_i^{lr}$ (at a significance level equal to 0.05).

The difference between the two measures of value added is economically large. On average, the long-run value added equals \$7.8M per year—a gap of \$6.4M vis-a-vis the standard measure va_i. This positive gap is consistent with theory. With diseconomies of scale ($b_i > 0$), the function $va_i(q_{i,t-1}) = (a_i - b_i q_{i,t-1})q_{i,t-1}$ is concave. As a result, the Jensen's inequality implies that $E[va_i(q_{i,t-1})] < va_i(E[q_{i,t-1}]) \Rightarrow va_i < va_i^{lr}$. In the context of models with investors' learning (e.g., Pastor and Stambaugh (2012)), the difference between va_i^{lr} and va_i captures the impact of the learning process—we have

[27] For this analysis, we only include funds with $\hat{a}_i > 0$ and $\hat{b}_i > 0$. This selection keeps the sample unchanged as we compare the actual size with its optimal value whose computation requires $\hat{a}_i > 0$ and $\hat{b}_i > 0$ (see Section E.1). It also removes charlatans whose size is expected to fall down to zero over time (see Section E.2).

$va_i^{lr} - va_i = b_i V[q_{i,t-1}]$, where the variance term captures the fluctuations in size as investors learn about a_i and b_i. Combined with strong diseconomies of scale (Table III), these fluctuations can create a large wedge between the two measures.

In addition, the proportion of funds that create value is also higher in the long run (71.7% versus 60.9%). This difference arises because some large funds see a reduction of their size over the sample period. As a result, their value added is negative on average ($va_i < 0$), but positive at the end of the sample period ($va_i^{lr} > 0$). Consistent with this interpretation, Panel B shows that we classify most funds with highly significant negative $\widehat{va}_i^{lr}$ as charlatans (i.e., 87.5% have negative fd alphas)—in contradiction to the result obtained with the average value added $\widehat{va}_i$.

Please insert Table VI here

C.2 The Social Value of the Mutual Fund Industry

Overall, we find that the large majority of funds create value. Through their trading activity, they make asset prices more efficient and help to improve the allocation of rare resources. Therefore, mutual funds perform a socially valuable function, regardless of how funds and investors share the value added.

To elaborate, mutual funds improve the efficiency of secondary markets, which may reduce the cost of capital in primary markets (Cochrane (2013), Pedersen (2018)) and produce real effects on the economy (Bond, Edmans, and Goldstein (2012)). For instance, managers may learn from equity prices and improve the firm's investment decisions. Active funds can also improve the efficiency of the primary markets by directly investing in private firms and participating to initial public offerings (e.g., Cumming and MacIntosh (2007), Reuter (2006)).

The value creation by the active industry as a whole is not inconsistent with the famous arithmetic of Sharpe (1991). This rule states that if passive investors (i) do not trade and (ii) hold the entire market, the aggregate return of the active industry cannot be superior to the market return. In reality, passive investors (i) trade regularly for many reasons (e.g., changes in market composition, lifecycle investing, index rebalancing), and (ii) do not hold the entire market as they focus on specific indices such as the SP500 (Pedersen (2018)). Therefore, even if we assume that mutual funds represent the only active investors, they can still beat the market and popular indices by trading based on information and providing liquidity to passive investors.[28]

[28] To illustrate, several passive investors track the Russell 2000 index which exhibits an annual turnover close to 50% per year. As passive investors are forced to rebalance their portfolios each year, Petajisto

D Skill, Scale, and Value Added

The fund's value added depends on its skill and scale (as per Equations (2)-(3)). Therefore, the value added distribution reflects the large heterogeneity in a_i and b_i. For one, the cross-sectional standard deviation of va_i^{lr} is 2.7 times larger than the mean (Table VI). The value added is also impacted by the trade-off between skill and scale. Because a_i and b_i are strongly correlated, the most profitable funds are not necessarily the ones with the best investment ideas. As these funds typically incur high trading costs, they could be dominated by funds that are able to scale up less profitable investment ideas.

To examine this issue, we sort $\hat{a}_i$ and $\hat{b}_i$ for each fund into deciles to create a scoring system from 1 to 10 (1=lowest, 10=highest). We then analyze the levels of skill and scale among the 89 funds with highly significant positive $\widehat{va}_i$. The median skill and scale scores for these funds are equal to 6 and 4. In addition, only 8 funds achieve the best skill score (10), and only 9 funds obtain the best scale score (1). We obtain similar results across funds with highly significant positive $\widehat{va}_i^{lr}$. Therefore, the most valuable funds are those able to strike a balance between skill and scale.

Our previous analysis of the investment categories in Table IV shows that they all imply specific skill-scale combinations. To examine which combination provides the highest value, Panel A of Table VII examines the distribution of the average value added within each category. Small cap group generally create more value than large cap funds over the sample period—for one, the difference in the cross-sectional average equals \$3.3M per year (\$3.1M versus \$-0.2M). We obtain qualitatively similar results when comparing low turnover funds with high turnover funds. Interestingly, small cap funds and low turnover funds both create more value even though they rely on a very different skill-scale combination: (i) high investment skills for small cap funds, (ii) high scaling abilities for low turnover funds.

In Panel B, we repeat the analysis for the long-run value added. Similar to the entire population (Table VI), we observe a large gap between the two measures of value added. For instance, the proportion of large cap funds that create value increases from 52.1% to 69.4% once we take a long-run perspective. The difference between va_i^{lr} and va_i is fairly homogeneous across investment categories. Therefore, we still observe that small cap funds and low turnover funds create more value in the long run.

Please insert Table VII here

(2011) find that they pay a cost between 37 and 78 bps to active investors which include mutual funds (Da, Gao, and Jagannathan (2011)).

E Equilibrium Considerations

E.1 Actual versus Optimal Value Added

Our analysis so far shows that most funds create value—especially once we focus on the end of the sample period. Therefore, we can rule out that most funds grow unambiguously too large. However, it does not imply that the size of the fund industry is consistent with a rational model for fund investment. To examine this issue, we examine the equilibrium predictions of the structural model of Berk and Green (2004).

In this model, we have (i) a set of skilled funds in scarce supply, and (ii) a large number of rational investors that compete for performance. As a result, funds are in a strong bargaining position and maximize profits π_i under the constraint that investors break even (i.e., fees $f_{e,i}$ must be equal to the gross alpha α_i). Therefore, a key prediction of the Berk and Green model is that funds maximize the value added, i.e., we have $\pi_i = f_{e,i}q_i = \alpha_i q_i = va_i$.

If both a_i and b_i are positive, we can replace α_i with $a_i - b_i q_i$ and use the first order condition $\frac{\partial va_i}{\partial q_i} = 0$ to solve for the optimal value added:

$$va_i^* = a_i q_i^* - b_i \left(q_i^*\right)^2 = \frac{a_i^2}{4b_i}. \tag{19}$$

where $q_i^* = \frac{a_i}{2b_i}$ denotes the optimal active size. Equation (19) shows that va_i^* is a simple function of skill and scale. Therefore, we can apply our nonparametric approach to statistically compare the optimal value added with its actual value.

For the average value added, we use the following expressions for the estimated measure $\hat{m}_i$ and its error term $u_{i,t}$:

$$\begin{aligned} \hat{m}_i &= \widehat{va}_i^* - \widehat{va}_i = \frac{\hat{a}_i^2}{4\hat{b}_i} - (\hat{a}_i \overline{q}_i - \hat{b}_i \overline{q}_{i,2}), \\ u_{i,t} &= u_{i,t}^* - u_{i,t}^{avg}, \end{aligned} \tag{20}$$

where $u_{i,t}^* = \frac{2a_i}{4b_i} e_1' Q_{x,i}^{-1} x_{i,t}\varepsilon_{i,t} - \frac{a_i^2}{4b_i^2} e_2' Q_{x,i}^{-1} x_{i,t}\varepsilon_{i,t}$, and $u_{i,t}^{avg}$ is the error term of $\widehat{va}_i$ obtained from Equation (12). For the long-run value added, we use the following expressions:

$$\begin{aligned} \hat{m}_i &= \widehat{va}_i^* - \widehat{va}_i^{lr} = \frac{\hat{a}_i^2}{4\hat{b}_i} - (\hat{a}_i \overline{q}_i - \hat{b}_i \overline{q}_i^2), \\ u_{i,t} &= u_{i,t}^* - u_{i,t}^{lr}, \end{aligned} \tag{21}$$

where $u_{i,t}^{lr}$ is the error term of $\widehat{va}_i^{lr}$ obtained from Equation (12). Our analysis is based on the sample of funds for which $\widehat{va}_i^*$ exists (Equation (19) requires that $\hat{a}_i > 0$ and $\hat{b}_i > 0$). We summarize the results in Table VIII which reports the average and standard deviation of the three cross-sectional distributions $\tilde{\phi}(va^*)$, $\tilde{\phi}(va^* - va)$, and $\tilde{\phi}(va^* - va^{lr})$.

The average value added over the sample period is far from the optimal level. In the population, the average difference between $\widehat{va}_i^*$ and $\widehat{va}_i$ reaches \$15.6M per year. Taken at face value, this result implies that funds do a poor job at optimally exploiting their investment skills. However, it could be the case that funds improve their optimization over time. A rationale for this argument is investors' learning—whereas previous work shows that investors respond to past returns (e.g., Chevalier and Ellison (1997)), Pastor and Stambaugh (2012) argue that the learning process about skill and scale is slow. In this context, our new measure of value added is informative because it allows for an adjustment period in the fund allocation.

The results obtained with the long-run value added reverse the conclusions about optimality. The difference between $\widehat{va}_i^*$ and $\widehat{va}_i^{lr}$ shrinks in magnitude to an average of \$3.2M per year, which means that funds extract around 80% of the optimal profits at the average size. In addition, the strong correlation between $\widehat{va}_i^{lr}$ and $\widehat{va}_i^*$ (0.97) confirms that funds with higher value potential do create more value. In short, the evidence suggests that the value created by the fund industry at the end of the sample period is relatively close to the predictions of the Berk and Green model.

Please insert Table VIII here

Similar to Figure 3, we examine the gap between the fund size and its optimal level in each period τ $(\tau = 1, ..., 10)$: $\Delta q_{i,\tau}^* = \sum_{t \in \tau} q_{i,t-1} - \hat{q}_i^*$, where $\hat{q}_i^* = \frac{\hat{a}_i}{2\hat{b}_i}$ denotes the estimated optimal active size. Figure 4 shows that the median value of $\Delta q_{i,\tau}^*$ is close to the optimal value in period 3. The gap then increases substantially before narrowing down at a level that represents 40.4% of $\hat{q}_i^*$. In theory, a positive gap between $q_{i,t-1}$ and $\hat{q}_i^*$ does not imply that funds fail to optimize the value added—as shown by Berk and Green (2004), they can still maximize the value added as long as they passively invest the excess funds $q_{i,t-1} - q_i^*$. The small difference between va_i^* and va_i^{lr} in Table VIII suggests that funds follow this strategy, albeit not perfectly. One possible reason is that funds are unsure of their own skill and scale and must learn about a_i and b_i alongside with investors.

Please insert Figure 4 here

Accounting for the strong correlation between skill and scale is essential to improve the ability of the Berk and Green model to explain the data. To elaborate, Zhu (2018) finds that the average difference between $\widehat{va}_i^*$ and $\widehat{va}_i$ is enormous at $26M per year. However, this paper uses a log size panel specification ($b_i = b$) to compute the optimal value. This homogeneous specification yields a positive bias for $\widehat{va}_i^*$ that grows larger when skill and scale are correlated.[29] Similarly, Roussanov, Ruan, and Wei (2020) use a panel specification to compare the actual and optimal sizes, and find a large negative (positive) size gap for less (more) skilled funds. Accounting for the strong correlation between skill and scale flattens the relationship between skill and optimal size and thus improves the fit of the Berk and Green model to the data.

E.2 Implications for Price Efficiency

Price efficiency is inversely related to the aggregate amount actively invested by skilled funds. Consistent with the Berk and Green model, our estimation results suggest that funds are in a strong bargaining position vis-a-vis investors. When maximizing profits, these funds internalize the impact of size on the gross alpha and thus ration the amount of active management. This implies that asset prices are less efficient than in a fully competitive market for asset management.

At the same time, skill and scale are strongly correlated, i.e., great investment ideas with high scalability are difficult to find. As a result, the mutual fund industry is not heavily concentrated. Table VI shows that the top 5% of the funds only capture 29.1% of the total value added at the average size.[30,31] This proportion would increase to 60.6% in a hypothetical world where a_i and b_i are uncorrelated (see the appendix (Section IV.D) for details of this counterfactual analysis). This lack of concentration implies that individual funds do not internalize the impact of size on aggregate mispricing, which increases both the amount of active management and price efficiency.

[29] Zhu (2018) uses a panel regression specification where $\alpha_{i,t} = a_i - b\log(q_{i,t-1})$ such that $va_i^*(b) = b\exp\left(\frac{a_i}{b} - 1\right)$. Using Jensen's inequality, we have that $E[va_i^*] = E\left[b_i \exp\left(\frac{a_i}{b_i} - 1\right)\right] < E[va_i^*(b)] = E\left[E[b_i]\exp\left(\frac{a_i}{E[b_i]} - 1\right)\right]$, with $E[b_i]$ corresponding to the pooled constant coefficient b.

[30] We denote the total value added for the entire population (all) and the top 5% (top) by $vaall^{lr} = nE[va_i^{lr}]$ and $vatop^{lr} = n \cdot 0.05E(va_i^{lr} \mid va_i^{lr} > Q_{0.95})$. To compute the industry concentration $\frac{vatop^{lr}}{vaall^{lr}}$, we replace $E[va_i^{lr}]$ with the estimated mean, and estimate $E(va_i^{lr} \mid va_i^{lr} > Q_{0.95})$ via a numerical integration of $\tilde{\phi}(va^{lr})$ to obtain 29.1%.

[31] This result contrasts with the high level of asset concentration observed among fund families (Berk, van Binsbergen, and Liu (2017)). One plausible explanation is that there are cost benefits in forming large families such as shared resources and improved trading commissions (Chen et al. (2004)).

E.3 From Value Added to Performance

An important question for investors is whether they benefit from the value created by mutual funds. To address this issue, we measure the net alpha received by investors: $\alpha_i^n = E[\alpha_{i,t}] - E[f_{e,i,t}] = a_i - b_i E[q_{i,t-1}] - E[f_{e,i,t}]$, where $E[f_{e,i,t}]$ denotes the average fees. To infer its cross-sectional density $\phi(\alpha^n)$, we apply our nonparametric approach using the following expressions for $\hat{m}_i$ and $u_{i,t}$:

$$\begin{aligned} \hat{m}_i &= \hat{\alpha}_i^n = \hat{a}_i - \hat{b}_i \bar{q}_i - \bar{f}_{e,i}, \\ u_{i,t} &= e_1' Q_{x,i}^{-1} x_{i,t}\varepsilon_{i,t} - E[q_{i,t-1}] e_2' Q_{x,i}^{-1} x_{i,t}\varepsilon_{i,t} \\ &\quad -b_i(q_{i,t-1} - E[q_{i,t-1}]) - (f_{e,i,t} - E[f_{e,i,t}]), \end{aligned} \tag{22}$$

where $f_{e,i,t}$ is the monthly fund fees, and $\bar{f}_{e,i} = \frac{1}{T_i}\sum_{t=1}^{T} I_{i,t} f_{e,i,t}$ is the sample average.

Of course, we are not the first to estimate the entire net alpha distribution—recent studies use standard parametric approaches to infer $\phi(\alpha^n)$ (e.g., Chen, Cliff, and Zhao (2017), Harvey and Liu (2018)). However, our nonparametric approach potentially brings several advantages discussed in Section III.A. In particular, it is robust to misspecification, simple to apply, and based on a full-fledged asymptotic theory.

Consistent with the literature on performance, Table IX provides limited evidence that investors benefit from the value created by mutual funds. Yet, the performance evaluation is more positive than in previous studies—for one, $\tilde{\pi}^+$ equals 38.2% versus 0.6% for the FDR approach (Barras, Scaillet, and Wermers (2010)). This difference arises because the FDR approach has a limited power to discriminate funds with net alphas close to zero (around 60% of the funds have $|\hat{\alpha}_i^n|$ <1.0% per year). In addition, we use the index-based model of Cremers, Petajisto, and Zitzewitz (2012) which produces higher net alphas than the traditional benchmark models (e.g., $\tilde{\pi}^+$ drops to 26.0% with the model of Carhart (1997)).

A common explanation for poor performance is that low-skilled funds exploit unsophisticated investors (e.g., Gruber (1996)). Consistent with this view, we find that the worst performing funds—those with $\hat{\alpha}_i^n$ significantly below the bottom quartile $\tilde{Q}_{0.25}$—are primarily classified as charlatans (i.e., 82.1% of them have negative estimated fd alphas), and charge high fees (the median fees equal 1.40% per year).[32] However, this explanation cannot fully account for the large proportion of negative-alpha funds because

[32] Over time, investors detect these funds because the median size gap $\Delta q_{i,\tau} = \sum_{t\in\tau} q_{i,t-1} - \bar{q}_i$ is highly negative in the last period (-40% in relative terms). However, the disinvestment takes a long time, consistent with slow learning—the majority of these funds have operated for more than 10 years (the median T equals 152).

charlatans only represent 12.9% of the population (Table II).

Please insert Table IX here

Overall, we find a sizable number of funds with both positive value added ($va_i^{lr} > 0$) and negative alphas ($\alpha_i^n < 0$). One explanation for these two findings in the context of the Berk and Green model is that investors have more optimistic views about skill than the funds themselves. We illustrate this point theoretically in Figure 5 using a simple example where (i) a skilled fund knows its skill a_i and sets fees such it operates at the optimal active size q_i^* ($f_{e,i} = \alpha_i(a_i, q_i^*)$), and (ii) investors believe that skill is equal to $a_i^1 > a_i$ and invest q_i^1 until they reach their perceived break-even point defined as $f_{e,i} = \alpha_i(a_i^1, q_i^1)$. Whereas the fund invests the difference $q_i^1 - q_i^*$ passively to keep the value added at its optimal value (i.e., $va(q_i^1) = \frac{a_i q_i^* - b_i (q_i^*)^2}{q_i^1} q_i^1 = va_i^*$), the alpha is negative (i.e., $f_{e,i} > \alpha_i(a_i, q_i^1)$).

Please insert Figure 5 here

An alternative explanation is that investors face information frictions (search costs) which prevent them from evaluating the entire fund population (Roussanov, Ruan, and Wei (2020)). In this setting, individual funds—including those with positive value added—may find it optimal to incur marketing expenses to attract investors with high search costs. These investors are then charged high fees and receive negative alphas as they do not switch to cheaper funds.

If these mechanisms are at play, we should expect a reduction in the high proportion of underperforming funds. The net alphas of existing funds should increase as investors sharpen their views about skill and scale and reallocate their funds. In addition, the recent advances in technology should reduce the information frictions faced by investors and give them access to a larger sample of funds to choose from.

F Additional Results

F.1 Alternative Asset Pricing Models

Our estimation of skill and scale depends on the choice of the asset pricing model. We therefore repeat our analysis using the four-factor model of Carhart (1997) and the five-factor model of Fama and French (2015) (see the appendix (Section VI.B)). Whereas the distributions of skill and scale remain largely unchanged, we observe two noticeable differences. First, the average fd alpha among small cap funds drops from 4.9% to 3.4%

per year under the Carhart model, consistent with the analysis of Cremers, Petajisto, and Zitzewitz (2012). Second, the proportion of funds with a positive fd alpha decreases from 87.1% to 75.9% with the Fama-French model. This reduction arises because some funds tilt their portfolios toward profitability- and investment-based strategies.

F.2 Fund Size and the Small Sample Bias

The estimated size coefficient of each fund potentially suffers from a small-sample bias. This issue arises because of the positive correlation between the return residual $\varepsilon_{i,t}$ and the innovation in size $\varepsilon_{q_i,t}$, i.e., $\varepsilon_{i,t} = \phi_i \varepsilon_{q_i,t} + \upsilon_{i,t}$. Whereas this bias vanishes asymptotically, it may have a significant impact for funds with short return time-series (small T_i).[33] To address this issue, we use the approach of Amihud and Hurvich (2004) and add a proxy for $\varepsilon_{q_i,t}$ to the set of regressors in Equation (4) (see the appendix (Section VI.C) for details). We find that the average size coefficient decreases from 1.5% to 1.4% per year. This is consistent with theory which predicts that the small-sample bias is positive, i.e., $E[\hat{b}_i - b_i] = -\phi E[\hat{\rho}_{q_i} - \rho_{q_i}] >0$, where ρ_{q_i} is the autocorrelation of fund size (Stambaugh (1999)).

F.3 Alternative Specifications for the Value Added

The analysis of the value added depends on the linear specification for the gross alpha (Equation (1)). To address this issue, we estimate the distributions of the value added under two alternative specifications for the gross alpha (see the appendix (Section VI.D) for details). First, we use a log specification in which $\alpha_{i,t} = a_i - b_i \log(q_{i,t-1})$. Second, we follow the estimation procedure of BvB and use a fully flexible function $g_i(q_{i,t-1})$, i.e., $\alpha_{i,t} = a_i - b_i g_i(q_{i,t-1})$. As long as this function is uncorrelated with the factors f_t, we can compute the two measures of the value added $\widehat{va}_i$ and $\widehat{va}_i^{lr}$ without having to specify its form. We find that the results are nearly indistinguishable from those reported in Table IV to VI. This strong similarity suggests that the flexibility of allowing for fund specific coefficients (instead of a panel specification) mitigates the impact of misspecification.

VI Conclusion

In this paper, we apply a new approach to study skill, scale, and value added in the mutual fund industry. For each of these measures, we provide an estimation of the

[33] In our analysis, we face a small-sample bias because we use a fund-by-fund time-series regression to estimate a_i and b_i (Equation (4)). This bias is therefore different from the incidental parameter bias that arises in a panel regression setting (Hjalmarsson (2010)).

entire distribution across funds. Our approach is nonparametric and thus avoids the challenge of correctly specifying each distribution—a great challenge when examining jointly skill, scale, and value added. In addition to its flexibility, our approach is simple to implement, applicable to the different characterizations of each distribution (e.g., moments, quantiles), and supported by econometric theory.

Our empirical analysis brings several insights. First, we find that most funds are skilled and thus able to extract value from capital markets. Second, the value created is close to its optimal value once we allow for an adjustment period. A natural interpretation of this result is that investors need time to learn about skill and scale and adjust their fund allocation. Third, we find that the shape of the value added distribution depends on skill and scale which both vary substantially across funds. Skill and scale are also strongly correlated as great investment ideas are difficult to scale up. As a result, funds with the highest value added follow investment strategies that strike a balance between skill and scale. Overall, these results contribute to the debate on the size of finance industry. They suggest that the size of active management is largely consistent with a rational model in which funds have bargaining power over investors. They also suggest that active funds make prices more efficient and contribute to the allocation of resources in the economy. As a result, the fund industry has a social value, regardless of how funds and investors share the value added.

Whereas our paper focuses on mutual funds, our nonparametric approach has potentially wide applications in finance and economics. It provides a new tool for measuring heterogeneity in structural models (Bonhomme and Shaikh (2017)). We can use it to estimate the cross-sectional distribution of any coefficient of interest in a random coefficient model. This is, for instance, the case in asset pricing for capturing the heterogeneity across stocks (risk exposure, commonality in liquidity), or in corporate finance for capturing the heterogeneity across firms (investment and financing decisions), and, more recently, in household finance for capturing the heterogeneity in time preference and risk aversion across households (see Calvet et al. (2019)).

References

[1] Ait-Sahalia, Y., 1996, Nonparametric Pricing of Interest Rate Derivative Securities, Econometrica, 1996, 527-560.

[2] Ait-Sahalia, Y., and A. Lo, 1998, Nonparametric Estimation of State-Price Densities Implicit in Financial Asset Prices, Journal of Finance, 53, 499-547.

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

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

[5] 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.

[6] Belsley, D. A., E. Kuh, and R.E. Welch, 2004, Regression Diagnostics—Identifying Influential Data and Sources of Collinearity. New York: Wiley.

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

[8] Berk, J. B., and J. van Binsbergen, 2019, Regulation of Charlatans in High-Skill Professions, Working Paper.

[9] Berk, J. B., J. van Binsbergen, and B. Liu, 2017, Matching Capital and Labor, Journal of Finance 72, 2467-2504.

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

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

[12] Bond, P., A. Edmans, and I . Goldstein, 2012, The Real Effects of Financial Markets, Annual Review of Financial Economics 4, 339-360.

[13] Bonhomme S., and A. Shaikh, 2017, Keeping the Econ in Econometrics: (Micro-) Econometrics in the Journal of Political Economy, Journal of Political Economy 125, 1846-1853.

[14] Calvet, L., J. Campbell, F. Gomes, and P. Sodini, 2019, The Cross-Section of Household Preferences, Working Paper.

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

[16] Chen, J., 2017, On Finite Mixture Models, Statistical Theory and Related Fields 1, 15-27.

[17] Chen, Y., M. Cliff, and H. Zhao, 2017, Hedge Funds: The Good, the Bad, and the Lucky. Journal of Financial and Quantitative Analysis 52, 1081–1109.

[18] 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.

[19] Chevalier, J., and G. Ellison, 1997, Risk Taking by Mutual Funds as a Response to Incentives, Journal of Political Economy 105, 1167–1200.

[20] Cici, G., S. Jaspersen, and A. Kempf, 2017, Speed of Information Diffusion within Fund Families, Review of Asset Pricing Studies 7, 145-170.

[21] Cochrane, J. H., 2013, Finance: Function Matters, not Size, Journal of Economic Perspectives 27, 29-50.

[22] Cohen, R. B., 2002, Dimensional Fund Advisors, Harvard Business School Case 203-026.

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

[24] Cumming, D. J., and J. G. MacIntosh, 2007, Mutual Funds that Invest in Private Equity? An Analysis of Labour-Sponsored Investment Funds, Cambridge Journal of Economics 31, 445-487.

[25] Da, Z., P. Gao, and R. Jagannathan, 2011, Impatient Trading, Liquidity Provision, and Stock Selection by Mutual Funds, Review of Financial Studies 24, 675-720.

[26] Dow, J., and G. Gorton, 1997, Noise Trading, Delegated Portfolio Management, and Economic Welfare, Journal of Political Economy 105, 1024-1050.

[27] Elton, E. J., M. J. Gruber, S. Das, and M. Hlavka, 1993, Efficiency with Costly Information: A Reinterpretation of Evidence from Managed Portfolios, Review of Financial Studies 6, 1-22.

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

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

[30] Gagliardini, P., E. Ossola, and O. Scaillet, 2020, Estimation of Large Dimensional Conditional Factor Models in Finance, in S. Durlauf, L. Hansen, J. Heckman, and R. Matzkin, editors, Handbook of Econometrics 7, Forthcoming.

[31] Ghosh, J., and P. Sen, 1985, On the Asymptotic Performance of the Log Likelihood Ratio Statistic for the Mixture Model and Related Results, in Le Cam, L. and R. Olshen, editors, Proc. Berkeley Conf. in honor of Jerzy Neyman and Jack Kiefer, 789-806.

[32] Greene, W., 2008, Econometric Analysis. Upper Saddle River: Prentice Hall.

[33] Greenwood, R., and D. Scharfstein, 2013, The Growth of Modern Finance, Journal of Economic Perspectives 27, 3-28

[34] Gruber, M. J., 1996, Another Puzzle: The Growth in Actively Managed Mutual Funds, Journal of Finance 51, 783-810.

[35] Habib, M., and D. B. Johnsen, 2016, The Quality-Assuring Role of Mutual Fund Advisory Fees, International Review of Law and Economics 46, 1-19.

[36] Hall, P., 1990, Using the Bootstrap to Estimate Mean Squared Error and Select Smoothing Parameter in Nonparametric Problems, Journal of Multivariate Analysis 32, 177-203.

[37] Hall, P., and K.-H. Kang, 2001, Bootstrapping Nonparametric Density Estimators with Empirically Chosen Bandwidths, Annals of Statistics 29, 1443-1468.

[38] Harvey, C., and Y. Liu, 2018, Detecting Repeatable Performance, Review of Financial Studies 31, 2499–2552.

[39] Hjalmarsson, E., 2010, Predicting Global Stock Returns, Journal of Financial and Quantitative Analysis 45, 49-80.

[40] Hong, H., T. Lim, and J. Stein, 2000, Bad News Travels Slowly: Size, Analyst Coverage, and the Profitability of Momentum Strategies, Journal of Finance 55, 265–295.

[41] Hsiao, C., 2003, Analysis of Panel Data, Cambridge University Press.

[42] Jagannathan, R., G. Skoulakis, and Z. Wang, 2010, The Analysis of the Cross-Section of Security Returns, Handbook of Financial Econometrics Vol. 2, 73-134.

[43] Jensen, M. C., 1968, The Performance of Mutual Funds in the Period 1945-1964, Journal of Finance 23, 389-416.

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

[45] Jones, C. S., and J. Shanken, 2005, Mutual Fund Performance with Learning across Funds, Journal of Financial Economics 78, 507-552.

[46] Kacperczyk, M., S. van Nieuwerburgh, and L. Veldkamp, 2014, Time-Varying Fund Manager Skill, Journal of Finance 69, 1455-1484.

[47] Kan, R., C. Robotti, and J. Shanken, 2013, Pricing Model Performance and the Two-Pass Cross-Sectional Regression Methodology, Journal of Finance 68, 2617-2649.

[48] Kosowski, R., A. Timmermann, R. Wermers, and H. White, 2006, Can Mutual Fund "Stars" Really Pick Stocks? New Evidence from a Bootstrap Analysis, Journal of Finance 61, 2551-2595.

[49] Moon, H. R., and P. C. B. Phillips, 2000, Estimation of Autoregressive Roots Near Unity Using Panel Data, Econometric Theory 16, 927–998.

[50] Newey, W. K., and K. D. West, 1987, A Simple, Positive Semi-Definite, Heteroscedasticity and Autocorrelation Consistent Covariance Matrix, Econometrica 55, 703-708.

[51] Neyman, J., and E.L. Scott, 1948, Consistent Estimates Based on Partially Consistent Observations, Econometrica 16, 1-32.

[52] Nickel, S., Biases in Dynamic Models with Fixed Effects, 1981, Econometrica 49, 1417-1426.

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

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

[55] Pastor, L., and R. F. Stambaugh, 2002, Mutual Fund Performance and Seemingly Unrelated Assets, Journal of Financial Economics 63, 315-349.

[56] Pastor, L., and R. F. Stambaugh, 2012, On the Size of the Active Management Industry, Journal of Political Economy 120, 740-781.

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

[58] Pastor, L., R. F. Stambaugh, and L. Taylor, 2020, Fund Tradeoffs, Journal of Financial Economics, Forthcoming.

[59] Pedersen, L. H., 2018, Sharpening the Arithmetic of Active Management, Financial Analysts Journal 74, 21-36.

[60] Perold, A. F., and R. S. Salomon, Jr., 1991, The Right Amount of Assets under Management, Financial Analysts Journal 47, 31-39.

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

[62] Petajisto, A., 2011, The Index Premium and Its Hidden Cost for Index Funds, Journal of Empirical Finance 18, 271-288.

[63] Reuter, J., 2006, Are IPO Allocations for Sale? Evidence from Mutual Funds, Journal of Finance 61, 2289-2324.

[64] Roussanov, N., H. Ruan, and Y. Wei, 2020, Marketing Mutual Funds, Review of Financial Studies, Forthcoming.

[65] Shanken, J., 1992, On the Estimation of Beta-Pricing Models, Review of Financial Studies 5, 1-33.

[66] Sharpe, W. F., 1991, The Arithmetic of Active Management, Financial Analysts Journal 47, 7-9.

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

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

[69] van der Vaart, A., 1998, Asymptotic Statistics, Cambridge Series in Statistical and Probabilistic Mathematics, Cambridge University Press.

[70] Wand, M. P., and M. C., Jones, 1995, Kernel Smoothing, Chapman and Hall, London.

[71] Wermers, R., 2000, Mutual Fund Performance: An Empirical Decomposition into Stock-Picking Talent, Style, Transaction Costs, and Expenses, Journal of Finance 55, 1655-1695.

[72] Yan, X. S., 2008, Liquidity, Investment Style, and the Relation between Size and Fund Performance, Journal of Financial and Quantiative Analysis 43, 741-767.

[73] Yan, C., and T. Cheng, 2019, In Search of the Optimal Number of Fund Subgroups, Journal of Empirical Finance 50, 78-92.

[74] Zhu, M., 2018, Informative Fund Size, Managerial Skill, and Investor Rationality, Journal of Financial Economics 130, 114-134.

Table I
Summary Statistics for the Equal-Weighted Portfolio of Funds

Panel A reports the first four moments of the portfolio gross excess return for all funds in the population, small/large cap funds, and low/high turnover funds (i.e., bottom and top terciles of funds sorted on turnover). Panel B reports the estimated portfolio betas on the market, size, value, and momentum factors using the index-based version of the four-factor model proposed by Cremers, Petajisto, and Zitzewitz (2012). Panel C reports the average number of funds in the portfolio, as well as the time-series average of the median fund size, fees, and turnover. We compute each portfolio as an equal-weighted average of all existing funds at the start of each month. All statistics are computed using monthly data between January 1975 and December 2018.

Panel A: Portfolio Gross Excess Return

	Mean (Ann.)	Std. Dev. (Ann.)	Skewness	Kurtosis
All Funds	9.0	15.3	-0.74	5.33
Fund Groups				
Small Cap	11.0	18.4	-0.65	5.04
Large Cap	8.6	14.7	-0.66	5.13
Low Turnover	8.5	14.4	-0.73	5.65
High Turnover	10.1	16.5	-0.67	5.00

Panel B: Portfolio Betas

	Market	Size	Value	Momentum
All Funds	0.94	0.36	-0.08	0.02
Investment Categories				
Small Cap	0.98	0.79	-0.13	0.03
Large Cap	0.95	0.17	-0.05	0.02
Low Turnover	0.91	0.29	0.04	-0.03
High Turnover	0.95	0.46	-0.23	0.07

Panel C: Portfolio Characteristics

	Number of Funds	Median Size ($M)	Median Fees (Ann.)	MedianTurnover (Ann.)
All Funds	941	218	1.10	64
Investment Categories				
Small Cap	184	143	1.27	68
Large Cap	404	273	1.00	62
Low Turnover	251	300	1.01	37
High Turnover	270	197	1.20	131

Table II
Skill

Panel A contains the summary statistics of the skill distribution (first dollar (fd) alpha) for all funds in the population. It reports the first four moments, the proportions of funds with a negative and positive fd alpha, 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 shows the characteristics of the funds with highly significant estimated fd alphas (at a significance level of 5%). It reports the number of significant funds, the proportion of "false discoveries" (funds with significant fd alphas by statistical luck alone), the median number of observations, and the median size, fees, and turnover.

Panel A: Distribution of the First Dollar Alpha

	Moments				Proportions (%)		Quantiles (Ann.)	
	Mean (Ann.)	Std. Dev. (Ann.)	Skewness	Kurtosis	Negative	Positive	5%	95%
All Funds	3.2 (0.1)	3.7 (0.2)	0.7 (0.4)	11.7 (1.5)	12.9 (0.7)	87.1 (0.7)	-1.5 (0.1)	9 (0.2)

Panel B: Characteristics of Funds with Significant First Dollar Alpha

	Significance		Proportion (%)		Median Fund Characteristics			
	Number of Funds	False Discoveries (%)	Estimated fd Alpha<0	Estimated fd Alpha>0	Number of Obs.	Size ($M.)	Fees (Ann.)	Turnover (Ann.)
Negative and Significant	33	79.2	100.0	0.0	176	141	1.12	98
Positive and Significant	601	4.3	0.0	100.0	242	398	1.23	71

Table III
Scale

Panel A contains the summary statistics of the scale distribution (size coefficient) for all funds in the population. It reports the first four moments, the proportions of funds with a negative and positive size 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 shows the characteristics of the funds with highly significant estimated size coefficients (at a significance level of 5%). It reports the number of significant funds, the proportion of "false discoveries" (funds with significant size coefficients by statistical luck alone), the proportions of funds with negative and positive estimated first dollar (fd) alphas, the median number of observations, and the median size, fees, and turnover.

Panel A: Distribution of the Size Coefficient

	Moments				Proportions (%)		Quantiles (Ann.)	
	Mean (Ann.)	Std. Dev. (Ann.)	Skewness	Kurtosis	Negative	Positive	5%	95%
All Funds	1.5 (0.1)	1.6 (0.1)	0.9 (0.4)	12.1 (1.9)	13.8 (0.7)	86.2 (0.7)	-0.6 (0.1)	4 (0.1)

Panel B: Characteristics of Funds with Significant Size Coefficient

	Significance		Proportion (%)		Median Fund Characteristics			
	Number of Funds	False Discoveries (%)	Estimated fd Alpha<0	Estimated fd Alpha>0	Number of Obs.	Size ($M.)	Fees (Ann.)	Turnover (Ann.)
Negative and Significant	17	100.0	100.0	0.0	209	146	1.12	66
Positive and Significant	491	6.2	0.2	99.8	224	324	1.23	78

Table IV
Skill and Scale across Investment Categories

Panel A shows the summary statistics of the skill distribution (first dollar (fd) alpha) across small/large cap funds and low/high turnover funds (i.e., bottom and top terciles of funds sorted on turnover). It reports the first four moments, the proportions of funds with a negative and positive fd alpha, and the quantiles at 5% and 95%. Panel B reports the same summary statistics for the scale distribution (size coefficient). 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 A: Distribution of the First Dollar Alpha

	Moments				Proportions (%)		Quantiles (Ann.)	
	Mean (Ann.)	Std. Dev. (Ann.)	Skewness	Kurtosis	Negative	Positive	5%	95%
Small Cap	4.9 (0.2)	4.4 (0.3)	1.2 (0.4)	9.7 (1.9)	8.8 (1.2)	91.2 (1.2)	-1.1 (0.3)	11.8 (0.3)
Large Cap	1.9 (0.1)	2.4 (0.2)	1.1 (1.1)	16.9 (13)	20 (1.3)	80 (1.3)	-1.7 (0.1)	5.9 (0.2)
Low Turnover	2.8 (0.1)	4.1 (0.2)	0.9 (0.7)	15.2 (5.2)	19.3 (1)	80.7 (1)	-2.6 (0.2)	9.4 (0.2)
High Turnover	3.6 (0.2)	4.7 (0.2)	1.3 (0.3)	10.8 (1.2)	17.7 (0.9)	82.3 (0.9)	-2.7 (0.2)	11 (0.2)

Panel B: Distribution of the Size Coefficient

	Moments				Proportions (%)		Quantiles (Ann.)	
	Mean (Ann.)	Std. Dev. (Ann.)	Skewness	Kurtosis	Negative	Positive	5%	95%
Small Cap	2 (0.1)	1.9 (0.1)	0.4 (0.5)	6.5 (2)	12.5 (1.3)	87.5 (1.3)	-0.8 (0.1)	5.1 (0.2)
Large Cap	1 (0.1)	1.2 (0.1)	1.4 (1.1)	19 (15.1)	18.7 (1.2)	81.3 (1.2)	-0.7 (0.1)	2.9 (0.1)
Low Turnover	1.2 (0.1)	1.6 (0.1)	-0.2 (0.6)	11.7 (3)	20.3 (1)	79.7 (1)	-1.1 (0.1)	3.7 (0.1)
High Turnover	1.7 (0.1)	2.2 (0.1)	0.2 (0.3)	5.8 (1.1)	20.4 (1)	79.6 (1)	-1.5 (0.1)	5.3 (0.1)

Table V
Average Value Added

Panel A contains the summary statistics of the distribution of the average value added over the sample period for all funds in the population. It reports the first four moments, the proportions of funds with a negative and positive value added, 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 shows the characteristics of the funds with highly significant estimated value added (at a significance level of 5%). It reports the number of significant funds, the proportion of "false discoveries" (funds with significant value added by statistical luck alone), the proportions of funds with negative and positive estimated first dollar (fd) alphas, the median number of observations, and the median size, fees, and turnover.

Panel A: Distribution of the Value Added

	Moments				Proportions (%)		Quantiles (Ann.)	
	Mean (Ann.)	Std. Dev. (Ann.)	Skewness	Kurtosis	Negative	Positive	5%	95%
All Funds	1.4 (0.3)	13.4 (1.1)	5.1 (1.2)	64.6 (7.2)	39.1 (1)	60.9 (1)	-7.1 (0.3)	20.1 (0.4)

Panel B: Characteristics of Funds with Significant Value Added

	Significance		Proportion (%)		Median Fund Characteristics			
	Number of Funds	False Discoveries (%)	Estimated fd Alpha<0	Estimated fd Alpha>0	Number of Obs.	Size ($M.)	Fees (Ann.)	Turnover (Ann.)
Negative and Significant	85	55.3	38.8	61.2	170	169	1.11	83
Positive and Significant	89	52.8	4.5	95.5	268	478	1.23	53

Table VI
Long-Run Value Added

Panel A contains the summary statistics of the distribution of the long-run value added (measured at the average fund size) for all funds in the population. It reports the first four moments, the proportions of funds with a negative and positive value added, 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 shows the characteristics of the funds with highly significant estimated value added (at a significance level of 5%). It reports the number of significant funds, the proportion of "false discoveries" (funds with significant value added by statistical luck alone), the proportions of funds with negative and positive estimated first dollar (fd) alphas, the median number of observations, and the median size, fees, and turnover.

Panel A: Distribution of the Value Added

	Moments				Proportions (%)		Quantiles (Ann.)	
	Mean (Ann.)	Std. Dev. (Ann.)	Skewness	Kurtosis	Negative	Positive	5%	95%
All Funds	7.8 (0.5)	21.1 (1.1)	3.1 (0.4)	26.1 (3.2)	28.3 (0.9)	71.7 (0.9)	-16.9 (0.5)	37.8 (0.6)

Panel B: Characteristics of Funds with Significant Value Added

	Significance		Proportion (%)		Median Fund Characteristics			
	Number of Funds	False Discoveries (%)	Estimated fd Alpha<0	Estimated fd Alpha>0	Number of Obs.	Size ($M.)	Fees (Ann.)	Turnover (Ann.)
Negative and Significant	40	92.5	87.5	12.5	169	89	1.26	92
Positive and Significant	269	13.8	0.7	99.3	257	518	1.29	62

Table VII
Value Added across Investment Categories

Panel A shows the summary statistics of the distribution of the average value added over the sample period across small/large cap funds and low/high turnover funds (i.e., bottom and top terciles of funds sorted on turnover). It reports the first four moments, the proportions of funds with a negative and positive value added, and the quantiles at 5% and 95%. Panel B reports the same summary statistics for the distribution of the long-run value added (measured at the average fund size). 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 A: Distribution of the Average Value Added

	Moments				Proportions (%)		Quantiles (Ann.)	
	Mean (Ann.)	Std. Dev. (Ann.)	Skewness	Kurtosis	Negative	Positive	5%	95%
Small Cap	3.1 (0.5)	11.6 (1.4)	3.9 (0.7)	35 (4.3)	34.7 (1.9)	65.3 (1.9)	-6.1 (0.4)	20.1 (0.7)
Large Cap	-0.2 (0.4)	12.3 (1.9)	5.8 (2.1)	98.4 (17.1)	47.9 (1.6)	52.1 (1.6)	-9.7 (0.4)	14 (0.5)
Low Turnover	4.3 (0.6)	21.9 (1.4)	2.4 (0.8)	28.4 (2.1)	30.5 (1.2)	69.5 (1.2)	-9.5 (0.5)	43.1 (0.9)
High Turnover	0.7 (0.4)	13.6 (1.6)	5.6 (1.7)	92.2 (19.4)	43.9 (1.2)	56.1 (1.2)	-10.2 (0.4)	18.6 (0.4)

Panel B: Distribution of the Long-Run Value Added

	Moments				Proportions (%)		Quantiles (Ann.)	
	Mean (Ann.)	Std. Dev. (Ann.)	Skewness	Kurtosis	Negative	Positive	5%	95%
Small Cap	8.2 (0.7)	16.3 (1.4)	2.9 (0.5)	20 (2.8)	17.6 (1.7)	82.4 (1.7)	-10 (0.7)	29.7 (0.8)
Large Cap	5 (0.5)	15.7 (1.1)	2.5 (0.8)	22.1 (7.3)	30.6 (1.4)	69.4 (1.4)	-12.6 (0.5)	30.4 (0.9)
Low Turnover	10.3 (0.7)	26.4 (1.1)	0.7 (0.6)	11.5 (2.8)	20.3 (1.1)	79.7 (1.1)	-13.8 (0.6)	63 (1.5)
High Turnover	8 (0.6)	23.4 (1.3)	2.7 (0.4)	20.9 (2.3)	29.2 (1.1)	70.8 (1.1)	-19.6 (0.6)	42.4 (0.9)

Table VIII
Optimal Versus Actual Value Added

The table shows the summary statistics for the difference between the optimal and actual value added for all funds in the population and across small/large cap funds, and low/high turnover funds (i.e., bottom and top terciles of funds sorted on turnover). It reports the mean and standard deviation of the distributions of (i) the optimal value added predicted by the Berk and Green model, (ii) the difference between the optimal value added and the average value added over the sample period, and (iii) the difference between the optimal value added and the long-run value added (measured at the average fund size). 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.

	Optimal Value Added Moments		vs Average Value Added Moments		vs Long-Run Value Added Moments	
	Mean (Ann.)	Std. Dev. (Ann.)	Mean (Ann.)	Std. Dev. (Ann.)	Mean (Ann.)	Std. Dev. (Ann.)
All Funds	15.6 (0.7)	32.2 (2)	14.2 (0.5)	23.3 (1.3)	3.2 (0.3)	14.3 (2)
Investment Categories						
Small Cap	11.4 (0.7)	16.1 (1.5)	9.8 (0.5)	10.5 (0.9)	2.5 (0.3)	5.4 (1.1)
Large Cap	16 (1.1)	35.4 (3.2)	17.2 (1)	31 (2.9)	3.8 (0.7)	21.1 (3.4)
Low Turnover	20.5 (0.9)	34.6 (2)	18.4 (0.7)	26.7 (1.5)	5.6 (0.5)	20.6 (3)
High Turnover	14.2 (0.7)	25.8 (1.9)	16.9 (0.6)	22.7 (1.2)	4 (0.3)	10.6 (2)

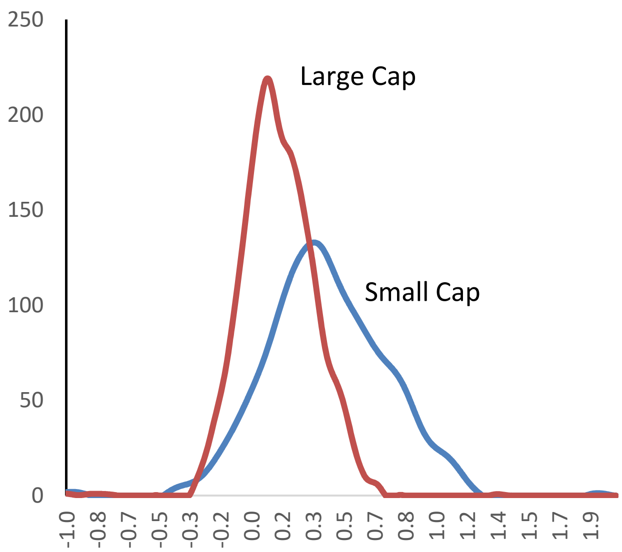
Table IX
Mutual Fund Performance

The table shows the summary statistics of the performance distribution (net alpha) for all funds in the population and across small/large cap funds, and low/high turnover funds (i.e., bottom and top terciles of funds sorted on turnover). It reports the first four moments, the proportions of funds with a negative and positive net alpha, 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.

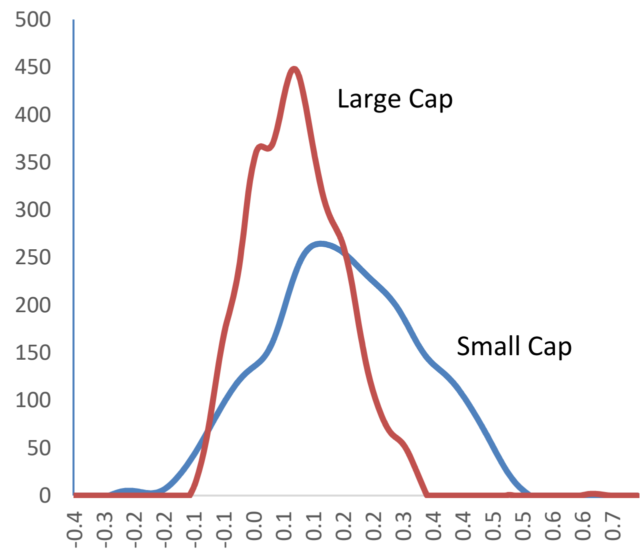
	Moments				Proportions (%)		Quantiles (Ann.)	
	Mean (Ann.)	Std. Dev. (Ann.)	Skewness	Kurtosis	Negative	Positive	5%	95%
All Funds	-0.4 (0.1)	1.4 (0.1)	-0.3 (0.3)	5.9 (1.4)	61.8 (1)	38.2 (1)	-2.5 (0.1)	1.8 (0.1)
Investment Categories								
Small Cap	0.4 (0.1)	1.8 (0.1)	0.4 (0.4)	4.3 (1.8)	40.5 (2)	59.5 (2)	-2.2 (0.2)	3.4 (0.2)
Large Cap	-0.8 (0.1)	1.1 (0.1)	-0.6 (0.5)	8 (2.4)	79.5 (1.3)	20.5 (1.3)	-2.4 (0.1)	0.9 (0.1)
Low Turnover	-0.2 (0.1)	1.5 (0.1)	-0.3 (0.5)	9.5 (2.6)	57 (1.2)	43 (1.2)	-2.5 (0.1)	2.2 (0.1)
High Turnover	-0.2 (0.1)	1.9 (0.1)	0.4 (0.3)	6 (1)	55.5 (1.2)	44.5 (1.2)	-3 (0.1)	2.8 (0.1)

Figure 1
Distributions of Skill and Scale across Investment Categories

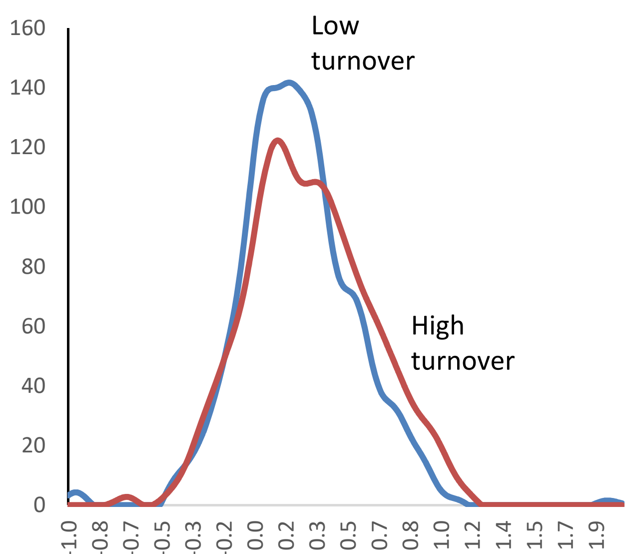
Panels A and B plot the cross-sectional densities of the first dollar (fd) alpha and size coefficient across small/large cap funds. Panel C and D plot the cross-sectional densities of the fd alpha and size coefficient, across low/high turnover funds (i.e., bottom and top terciles of funds sorted on turnover). We adjust all the estimated densities for the error-in-variable (EIV) bias using our nonparametric approach.



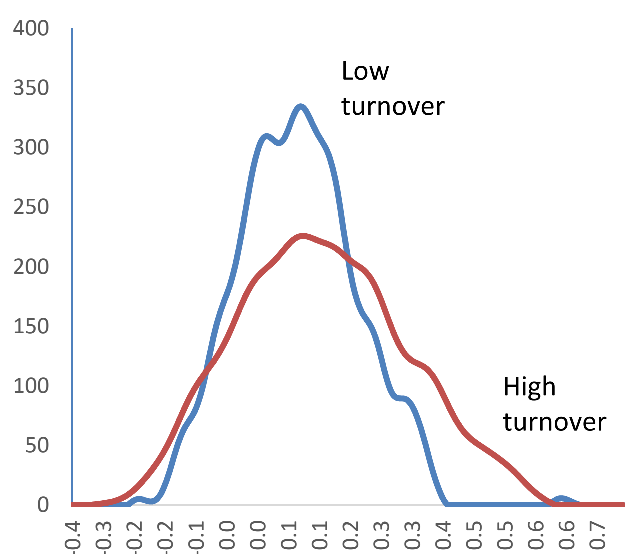
Panel A: Size Groups – Fd Alpha



Panel B: Size Groups – Size Coefficient



Panel C: Turnover Groups – Fd Alpha



Panel D: Turnover Groups – Size Coefficient

Figure 2
Impact of the Error-in-Variable Bias:
Distribution Comparison for the Value Added

This figure plots the cross-sectional density of the value added over the lifecycle for all funds in the population with and without the adjustment for the error-in-variable (EIV) bias.

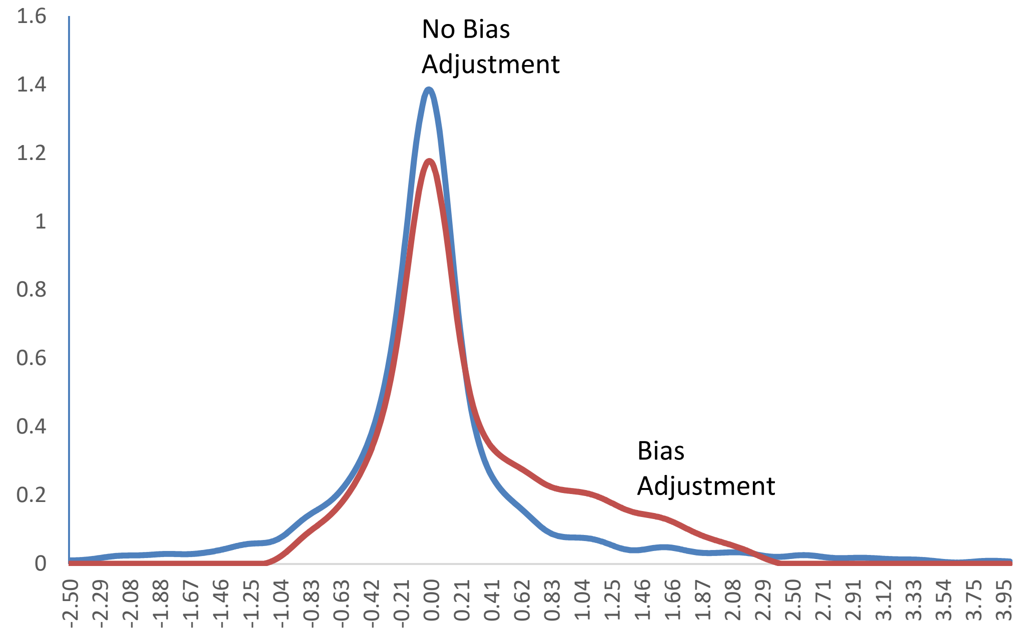
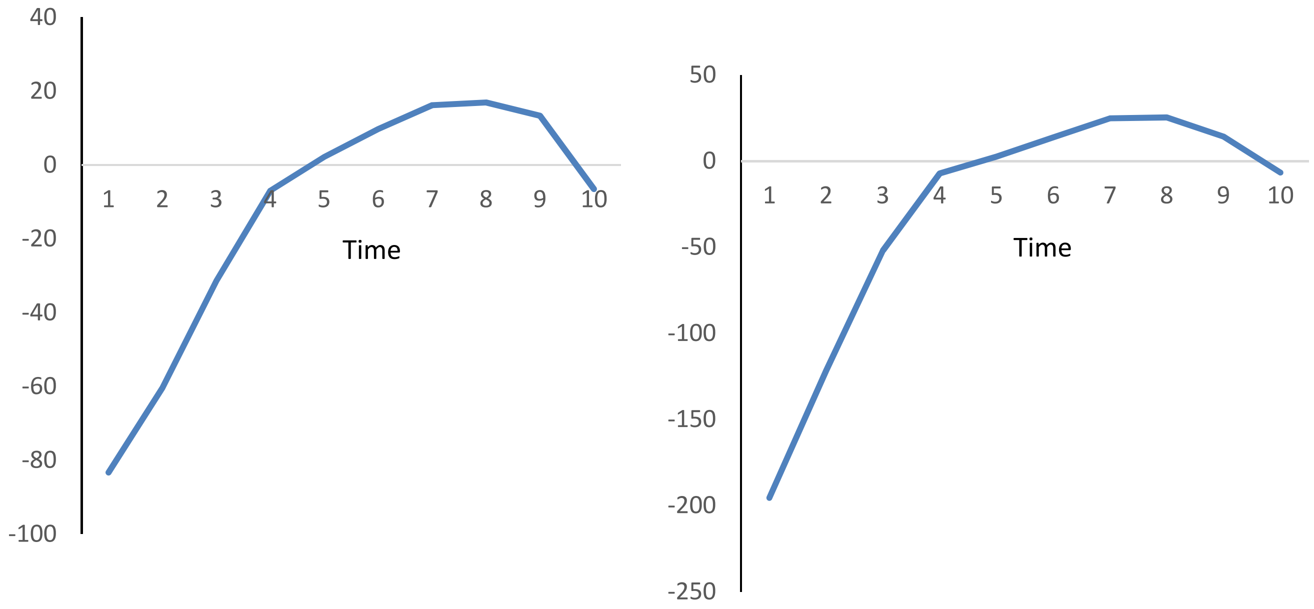


Figure 3
Variation of Fund Size over Time

This figure plots the cross-sectional median difference between the average size of the fund over 10 subperiods and its full-sample average. This analysis is based on the sample of funds with positive estimated first dollar alphas and size coefficients.

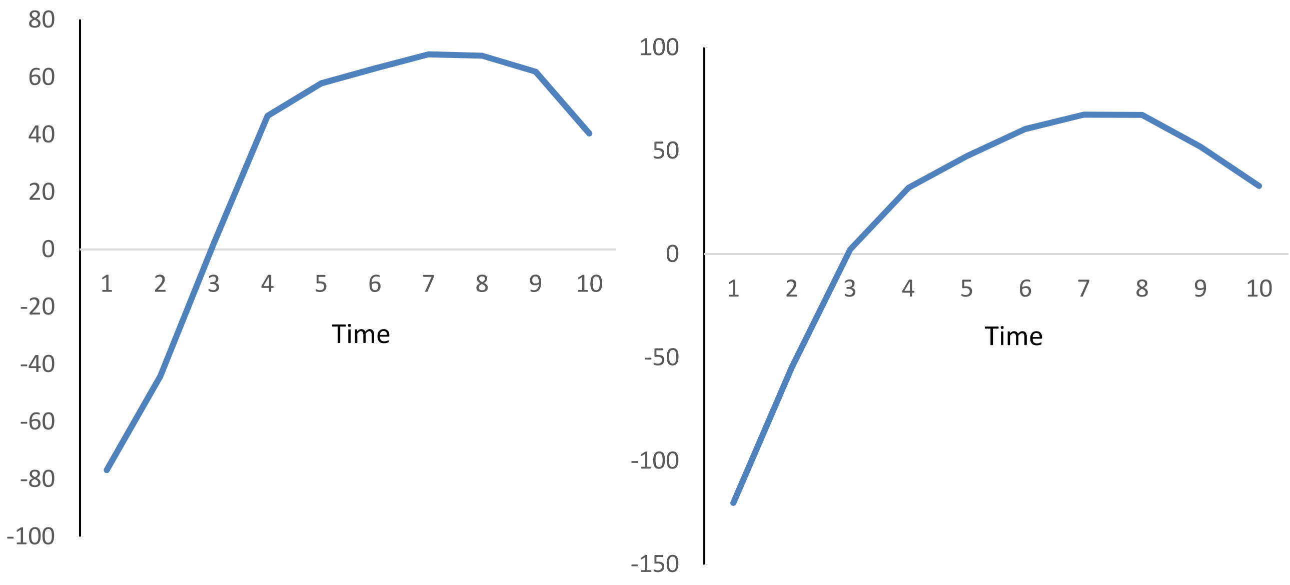


Panel A: Relative Difference

Panel B: Dollar Difference ($M)

Figure 4
Optimal versus Actual Fund Size over Time

This figure plots the cross-sectional median difference between the average size of the fund over 10 subperiods and the optimal active size predicted by the Berk and Green model. This analysis is based on the sample of funds with positive estimated first dollar alphas and size coefficients.

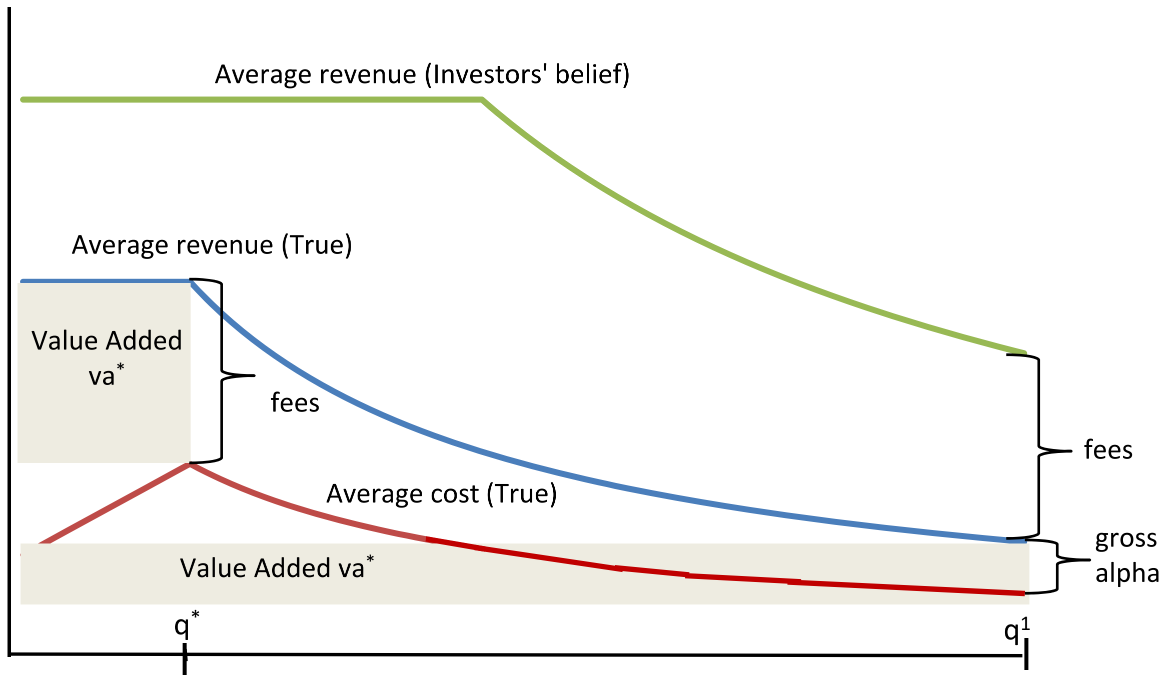


Panel A: Relative Difference

Panel B: Dollar Difference ($M)

Figure 5
Example of a Fund with Positive Value Added and Negative Alpha

This figure plots the average revenue and cost functions of a fund under the Berk and Green model. The fund knows its average revenue and cost functions (denoted by true) and sets fees such that the fund size is equal to the optimal actve size (q^*). Investors have optimistic beliefs about skill (fd alpha) and, given the level of fees, are willing to invest q^1. To keep the value added unchanged at its optimal level (va^*), the fund invests the difference between q^1 and q^* passively (which explains the kinked shapes of the true average revenue and cost functions). Whereas this strategy maximizes the value added, it still produces a negative alpha, i.e., the difference between the gross alpha and fees is negative.



INTERNET APPENDIX
Skill, Scale, and Value Creation in the Mutual Fund Industry

Laurent Barras, Patrick Gagliardini, and Olivier Scaillet*

This version, July 8, 2020

*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 contains the new econometric theory for the kernel density estimator used in the paper. We derive its asymptotic properties and optimal bandwidth, and show how to use a Gaussian reference model for practical purposes. Section II examines the estimators of the other distribution characteristics (moments, proportion, and quantile). Section III provides a detailed analysis of the Error-in-Variable (EIV) bias. Section IV describes our extensive Monte-Carlo analysis, and presents simulation results when we assume that skill and scale are uncorrelated. Section V describes the construction of the data set and different investment categories. Finally, Section VI gathers additional empirical results on (i) the test on the panel specification, (ii) alternative asset pricing models, (iii) the small sample bias, and (iv) alternative specifications for the gross alpha.

I Estimator of the Kernel Density

A Asymptotic Properties

Proof of Proposition III.1. In this section, we provide a proof of the first proposition in the paper which determines the asymptotic properties of the kernel density $\hat{\phi}(m)$ for the different measures $m \in \{a_i, b_i, va_i, va_i^{lr}\}$. We begin our presentation with the first dollar (fd) alpha ($m_i = a_i$). 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. From the OLS estimation of Equation (4), we have:

$$\begin{aligned} \hat{m}_i &= e_1'\hat{Q}_{x,i}^{-1}\frac{1}{T_i}\sum_t I_{i,t}x_t r_{i,t} = m_i + e_1'\hat{Q}_{x,i}^{-1}\frac{1}{T_i}\sum_t I_{i,t}x_t\varepsilon_{i,t} \\ &= m_i + \frac{1}{\sqrt{T}}\tau_{i,T}e_1'\hat{Q}_{x,i}^{-1}\left(\frac{1}{\sqrt{T}}\sum_t I_{i,t}x_t\varepsilon_{i,t}\right) \equiv m_i + \frac{1}{\sqrt{T}}\hat{\eta}_{i,T}. \end{aligned} \quad \text{(A1)}$$

Moreover, let us write

$$\hat{\eta}_{i,T} = \eta_{i,T} + \frac{1}{\sqrt{T}}\hat{v}_{i,T}, \quad \text{(A2)}$$

where $\eta_{i,T} = \tau_i\frac{1}{\sqrt{T}}\sum_t I_{i,t}u_{i,t}$, $u_{i,t} = e_1'Q_{x,i}^{-1}x_t\varepsilon_{i,t}$, $\hat{v}_{i,T} = (\tau_{i,T} - \tau_i)\sum_t I_{i,t}e_1'\hat{Q}_{x,i}^{-1}x_t\varepsilon_{i,t} + \tau_i\sum_t I_{i,t}e_1'(\hat{Q}_{x,i}^{-1} - Q_{x,i}^{-1})x_t\varepsilon_{i,t}$, $\tau_i = \text{plim}_{T\to\infty}\tau_{i,T}$, 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:

$$\begin{aligned}
I_1 &= \frac{1}{h} E\left[K\left(\frac{m_i - m}{h}\right)\right] - \phi(m), \\
I_2 &= \frac{1}{h} E\left[K\left(\frac{m_i + \eta_{i,T}/\sqrt{T} - m}{h}\right)\right] - \frac{1}{h} E\left[K\left(\frac{m_i - m}{h}\right)\right], \\
I_3 &= \frac{1}{nh} \sum_i \left\{K\left(\frac{m_i + \eta_{i,T}/\sqrt{T} - m}{h}\right) - E\left[K\left(\frac{m_i + \eta_{i,T}/\sqrt{T} - m}{h}\right)\right]\right\}, \\
I_4 &= \frac{1}{nh} \sum_i \left[\mathbf{1}_i^{\chi} K\left(\frac{m_i + \eta_{i,T}/\sqrt{T} + \hat{v}_{i,T}/T - m}{h}\right)\right] \\
&\quad - \frac{1}{nh} \sum_i \left[K\left(\frac{m_i + \eta_{i,T}/\sqrt{T} - m}{h}\right)\right]. \qquad \text{(A3)}
\end{aligned}$$

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

$$I_2 = \sum_{j=1}^{\infty} \frac{1}{j! T^{j/2} h^{j+1}} E\left[K^{(j)}\left(\frac{m_i - m}{h}\right)(\eta_{i,T})^j\right]. \qquad \text{(A4)}$$

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

$$\begin{aligned}
\frac{1}{h^{j+1}} E\left[K^{(j)}\left(\frac{m_i - m}{h}\right)(\eta_{i,T})^j\right] &= \frac{1}{h^{j+1}} \int K^{(j)}\left(\frac{u - m}{h}\right) \psi_{T,j}(u) du \\
&= (-1)^j \frac{1}{h} \int K\left(\frac{u - m}{h}\right) \psi_{T,j}^{(j)}(u) du \\
&= (-1)^j \int K(u)\, \psi_{T,j}^{(j)}(m + hu) du, \qquad \text{(A5)}
\end{aligned}$$

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 x_t as given constants. Then:

$$V[I_3] = \frac{1}{nh^2} V\left[K\left(\frac{m_i + \eta_{i,T}/\sqrt{T} - m}{h}\right)\right]. \tag{A6}$$

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

$$\begin{aligned}\frac{1}{h}E\left[K\left(\frac{m_i + \eta_{i,T}/\sqrt{T} - m}{h}\right)^2\right] &= \int K(u)^2 du \frac{1}{h}E\left[\bar{K}\left(\frac{m_i + \eta_{i,T}/\sqrt{T} - m}{h}\right)\right] \\ &= \phi(m)\int K(u)^2 du + o(1), \end{aligned} \tag{A7}$$

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

$$V[I_3] = \frac{1}{nh}\phi(m)\int K(u)^2 du + o\left(\frac{1}{nh}\right). \tag{A8}$$

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.[1] QED

We can apply the same arguments for all the other measures used in the paper: (i) the size coefficient ($m_i = b_i$), (ii) the lifecycle value added ($m_i = va_i$), (iii) the value added at average size ($m_i = va_i^{lr}$), (iv) the optimal value added ($va_i^* = m_i$), (v) the difference between the optimal and actual value added ($m_i = va_i^* - va_i$ and $m_i = va_i^* - va_i^{lr}$), and (vi) the net alpha ($m_i = \alpha_i^n$). The only required change is to use the appropriate definition for $\hat{m}_i$ and $u_{i,t}$ given in the paper.

B Choice of the Optimal Bandwidth

The estimation of the skill density $\hat{\phi}(m)$ requires the choice of the bandwidth h. To guide this choice, we solve for the optimal bandwidth h^* that minimizes the Asymptotic Mean Integrated Squared Error (AMISE) of the density. From the above arguments, we get the asymptotic expansion of the bias of the estimator $\hat{\phi}(m)$ with leading terms:

[1] 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.

$bs(m) = bs_1(m) + bs_2(m)$, where $bs_1(m)$ denotes the smoothing bias and $bs_2(m)$ denotes the EIV bias:

$$bs_1(m) = \frac{1}{2}h^2 K_2 \phi^{(2)}(m), \tag{A9}$$

$$bs_2(m) = \frac{1}{2T}\psi^{(2)}(m). \tag{A10}$$

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

$$\begin{aligned} AMISE(h) &= \int [\sigma^2(u) + bs(u)^2]du \\ &= \frac{1}{nh}K_1 + \frac{h^4 K_2^2}{4}\int [\phi^{(2)}(u)]^2 du \\ &\quad + \frac{h^2 K_2}{2T}\int \phi^{(2)}(u)\psi^{(2)}(u)du + \frac{1}{4T^2}\int \left[\psi^{(2)}(u)\right]^2 du, \end{aligned} \tag{A11}$$

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:

$$\begin{aligned} & -\frac{1}{nh^2} + c_1 h^3 + c_2\frac{h}{T} = 0 \\ \Longleftrightarrow \quad & 1 = c_1 n h^5 + c_2 \frac{nh^3}{T}. \end{aligned} \tag{A12}$$

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

We now investigate the form of the optimal bandwidth h^* as a function of the relative increase of n and T. Specifically, we have:

(i) The optimal bandwidth is such that (i) nh^3/T tends to a nonzero constant and (ii) nh^5 tends to zero. Then, we have

$$h^* \sim c_2^{-1/3}(n/T)^{-1/3}. \tag{A13}$$

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.

(ii) The optimal bandwidth is such that (i) nh^3/T tends to zero and (ii) nh^5 tends to a nonzero constant. Then, we have

$$h^* \sim c_1^{-1/5} n^{-1/5}, \tag{A14}$$

which is 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.[2]

We now consider the asymptotic distribution when $h = h^*$ is the optimal bandwidth. We can check that $\sqrt{nh^*}(h^{*3}+h^{*2}/T+1/T^{3/2}) = o(1)$ if $n/T^4 \to 0$. Then, we can replace the bias component $bs(m)$ by its asymptotic approximation to get:

$$\sqrt{nh^*}\left(\hat{\phi}(m) - \phi(m) - \frac{1}{2}\phi^{(2)}(m)K_2 h^{*2} - \frac{1}{2T}\psi^{(2)}(m)\right) \Rightarrow N\left(0, \phi(m)K_1\right). \qquad \text{(A15)}$$

Under case (i), we have $n^{2/5}/T \to \infty$ and $Th^{*2} \to 0$. As a result, the smoothing bias is negligible and the dominant component is the EIV bias of order $O(1/T)$. Under case (ii), we have $n^{2/5}/T \to 0$ and $Th^{*2} \to \infty$. As a result, the EIV bias is negligible and the dominant component is the smoothing bias of order $O(h^2)$. To determine which choice of bandwidth produces a stronger performance in finite samples, we use a Monte-Carlo simulation that replicates the salient features of the data. This analysis, which is discussed below, reveals that bandwidth (i) is preferable.

C The Gaussian Reference Model

Proof of Proposition III.2. We now prove the second propostion in the paper which provides closed form expressions for the optimal bandwidth h^* and the two bias components $bs_1(m)$ and $bs_2(m)$. 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.[3] 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:

$$c_1 = 2\sqrt{\pi}\int [\phi^{(2)}(u)]^2 du, \qquad \text{(A16)}$$

$$c_2 = 2\sqrt{\pi}\int \phi^{(2)}(u)\psi^{(2)}(u) du = 2\sqrt{\pi}\int \phi^{(4)}(u)\psi(u) du, \qquad \text{(A17)}$$

where we use twice partial integration for c_2.

[2] 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 (A12) 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).

[3] 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.

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

$$\phi^{(1)}(m) = -\frac{1}{\sigma_m}\left(\frac{m-\mu_m}{\sigma_m}\right)\phi(m), \tag{A18}$$

$$\phi^{(2)}(m) = \frac{1}{\sigma_m^2}\left(\left(\frac{m-\mu_m}{\sigma_m}\right)^2 - 1\right)\phi(m). \tag{A19}$$

Therefore, the first integral is equal to

$$\begin{aligned}\int[\phi^{(2)}(u)]^2du &= \frac{1}{\sigma_m^5}\int(z^2-1)^2\frac{1}{2\pi}\exp(-z^2)dz = \frac{1}{2\sqrt{\pi}\sigma_m^5}\int(v^2/2-1)^2\varphi(v)dv \\ &= \frac{3}{8\sqrt{\pi}\sigma_m^5},\end{aligned} \tag{A20}$$

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

$$\begin{aligned}\int\phi^{(4)}(m)\psi(m)dm &= \frac{\exp\left(\mu_s + \frac{1}{2}\sigma_s^2(1-\rho^2)\right)}{\sigma_m^5}\int\varphi^{(4)}(z)\exp(\rho\sigma_s z)\varphi(z)dz \\ &= \frac{\exp\left(\mu_s + \frac{1}{2}\sigma_s^2(1-\rho^2)\right)}{2\sqrt{\pi}\sigma_m^5}\int(v^4/4 - 3v^2 + 3)\exp(\lambda v)\varphi(v)dv,\end{aligned} \tag{A21}$$

where

$$\begin{aligned}\psi(m) &= E[\exp(s_i)|m_i = m]\phi(m) \\ &= \exp\left(\mu_s + \rho\sigma_s\left(\frac{m-\mu_m}{\sigma_m}\right) + \frac{1}{2}\sigma_s^2(1-\rho^2)\right)\phi(m),\end{aligned} \tag{A22}$$

$\lambda = \rho\sigma_s/\sqrt{2}$ 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 (A21), 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

$$\int\phi^{(4)}(m)\psi(m)dm = \frac{3\exp\left(\mu_s + \frac{1}{2}\sigma_s^2\left(1-\frac{\rho^2}{2}\right)\right)}{8\sqrt{\pi}\sigma_m^5}(\rho^4\sigma_s^4/12 - \rho^2\sigma_s^2 + 1). \tag{A23}$$

Using these results, we obtain the optimal bandwidth h^*. Under case (i), we have

$$
\begin{aligned}
h^* &\sim c_2^{-1/3}(n/T)^{-1/3} \\
&= \left[\frac{3(\rho^4\sigma_s^4/12-\rho^2\sigma_s^2+1)}{4\sigma_m^5}\exp\left(\mu_s+\frac{1}{2}\sigma_s^2\left(1-\frac{\rho^2}{2}\right)\right)\right]^{-\frac{1}{3}}(n/T)^{-1/3}, \quad \text{(A24)}
\end{aligned}
$$

where $c_2 \geq 0$ when either $\rho^2\sigma_s^2 \leq 6-2\sqrt{6}$, or $\rho^2\sigma_s^2 \geq 6+2\sqrt{6}$. Under case (ii), we have

$$
h^* \sim c_1^{-1/5}n^{-1/5} = \left(\frac{3}{4\sigma_m^5}\right)^{-1/5} n^{-1/5}. \quad \text{(A25)}
$$

Finally, we can use the Gaussian reference model to obtain closed form expressions of the smoothing bias and the EIV bias. Differentiating $\psi(m)$ in Equation (A22) twice, we obtain[4]

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

Using Equations (A19) and (A26), we can replace $\phi^{(2)}(m)$ and $\psi^{(2)}(m)$ in Equations (A9) and (A10) to obtain the two bias terms under the reference model:

$$
\begin{aligned}
bs_1^r(m) &= \frac{1}{2}h^2K_2\phi^{(2)}(m) = \left[\frac{1}{2}h^2\frac{1}{\sigma_m^2}(\bar{m}_1^2-1)\right]\frac{1}{\sigma_m}\varphi(\bar{m}_1), \quad \text{(A27)} \\
bs_2^r(m) &= \frac{1}{2T}\psi^{(2)}(m) = \left[\frac{1}{2T}\exp(\mu_s+\frac{1}{2}\sigma_s^2)\frac{1}{\sigma_m^2}(\bar{m}_2^2-1)\right]\frac{1}{\sigma_m}\varphi(\bar{m}_2), \quad \text{(A28)}
\end{aligned}
$$

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

[4] 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.

II Estimators of the Distribution Characteristics

A Moments

We now turn to the analysis of the estimated moments of the distribution. 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 linear factor model in Equation (4) and in an unbalanced panel.

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

$$\sqrt{n}\left(\frac{1}{n}\sum_i g(\hat{m}_i)\mathbf{1}_i^{\chi} - E\left[g(m_i)\right] - \mathcal{B}_T\right) \Rightarrow N\left(0, V\left[g(m_i)\right]\right), \tag{A29}$$

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 (A1) yields the mean value expansion

$$g(\hat{m}_i) = g(m_i) + g^{(1)}(\bar{m}_i)\frac{1}{\sqrt{T}}\hat{\eta}_{i,T} + g^{(2)}(\bar{m}_i)\frac{1}{2T}\hat{\eta}_{i,T}^2, \tag{A30}$$

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

$$\begin{aligned}
& \sqrt{n}\left(\frac{1}{n}\sum_i g(\hat{m}_i)\mathbf{1}_i^{\chi} - E\left[g(m_i)\right] - \mathcal{B}_T\right) \\
= \; & \frac{1}{\sqrt{n}}\sum_i \left(g(m_i) - E\left[g(m_i)\right]\right) - \frac{1}{\sqrt{n}}\sum_i g(m_i)(1-\mathbf{1}_i^{\chi}) + \frac{1}{\sqrt{nT}}\sum_i \mathbf{1}_i^{\chi} g^{(1)}(\bar{m}_i)\hat{\eta}_{i,T} \\
& + \frac{1}{2T}\frac{1}{\sqrt{n}}\sum_i \left(\mathbf{1}_i^{\chi} g^{(2)}(\bar{m}_i)\hat{\eta}_{i,T}^2 - E\left[g^{(2)}(m_i)S_i\right]\right) \\
\equiv \; & I_{21} + I_{22} + I_{23} + I_{24}.
\end{aligned} \tag{A31}$$

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 $n = o(T^3)$.[5] Therefore, the asymptotic distribution in Equation (A29) depends on the first term $I_{21} \Rightarrow N(0, V[g(m_i)])$ from the standard CLT. QED

The distribution results in Equation (A29) reveal that we have an asymptotic bias

[5] 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.

$\mathcal{B}_T$ of order $1/T$ which comes from the estimation error of $\hat{m}_i$ (EIV contribution). In addition, the estimator is precisely estimated because it converges at a rate proportional to the square root of n.

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

$$\mathcal{B}_T(SD) = (2SD)^{-1}\mathcal{B}_T(E[m_i^2]), \tag{A32}$$

where the asymptotic bias of the second moment is given by

$$\mathcal{B}_T(E[m_i^2]) = \frac{1}{2T}E[2S_i]. \tag{A33}$$

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

$$\mathcal{B}_T(Sk) = (\nabla_3 Sk)\mathcal{B}_T(E[m_i^3]) + (\nabla_2 Sk)\mathcal{B}_T(E[m_i^2]), \tag{A34}$$

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

$$\begin{aligned}
\mathcal{B}_T(E[m_i^3]) &= \frac{1}{2T}E[6m_iS_i], \\
\nabla_3 Sk &= V[m_i]^{-3/2}, \\
\nabla_2 Sk &= -3E[m_i]V[m_i]^{-3/2} + E[m_i^3](\frac{-3}{2})V[m_i]^{-5/2} \\
&\quad + \left\{-3E[m_i^2]E[m_i] + 2E[m_i]^3\right\}(\frac{-3}{2})V[m_i]^{-5/2}.
\end{aligned} \tag{A35}$$

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

$$\mathcal{B}_T(Ku) = (\nabla_4 Ku)\mathcal{B}_T(E[m_i^4]) + (\nabla_3 Ku)\mathcal{B}_T(E[m_i^3]) + (\nabla_2 Ku)\mathcal{B}_T(E[m_i^2]), \quad \text{(A36)}$$

where the different terms are given by

$$\begin{aligned}
\mathcal{B}_T(E[m_i^4]) &= \frac{1}{2T}E[12m_i^2 S_i], \\
\nabla_4 Ku &= V[m_i]^{-2}, \\
\nabla_3 Ku &= -4E[m_i]V[m_i]^{-2}, \\
\nabla_2 Ku &= 6E[m_i]^2 V[m_i]^{-2} + \left\{E[m_i^4] - 4E[m_i^3]E[m_i]\right\}(-2)V[m_i]^{-3} \\
&\quad + \left\{6E[m_i^2]E[m_i]^2 - 3E[m_i]^4\right\}(-2)V[m_i]^{-3}. \qquad \text{(A37)}
\end{aligned}$$

Alternatively, we can use a numerical integration to correct for the EIV bias. From the definitions of the standard deviation, skewness, and kurtosis, we obtain:

$$SD = V^{\frac{1}{2}} = \left(\int \phi(u)(u-M)^2 du\right)^{\frac{1}{2}}, \quad \text{(A38)}$$

$$Sk = \frac{\int \phi(u)(u-M)^3 du}{V^{\frac{3}{2}}}, \quad \text{(A39)}$$

$$Ku = \frac{\int \phi(u)(u-M)^4 du}{V^2}, \quad \text{(A40)}$$

where we replace $\phi(u)$ with the bias-adjusted density estimator $\tilde{\phi}(u)$ in the above expressions to obtain the bias-adjusted estimators $\widetilde{SD}$, $\widetilde{Sk}$, and $\widetilde{Ku}$. Asymptotically, both approaches (analytical or numerical) are equivalent.

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 $V\left[\tilde{M}\right] = V$. Therefore, we simply need a consistent estimator of the variance of the distribution V.

(ii) For the estimated volatility, we have:

$$V\left[\widetilde{SD}\right] = E\left[\left((2SD)^{-1}\Psi_2\right)^2\right], \quad \text{(A41)}$$

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:

$$V\left[\widetilde{Sk}\right] = E\left[\left(SD^{-3}\Psi_3 - \frac{3}{2}SD^{-2}Sk\Psi_2 - 3SD^{-1}\Psi_1\right)^2\right], \quad \text{(A42)}$$

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:

$$V\left[\widetilde{Ku}\right] = E\left[\left(SD^{-4}\Psi_4 - 2SD^{-2}Ku\Psi_2 - SD^{-1}Sk\Psi_1\right)^2\right], \quad \text{(A43)}$$

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

B Proportion and Quantile

Finally, we focus on the proportion estimator inferred from the cumulative distribution function (cdf) and the associated quantile. 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. 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 $n = o(T^3)$,

$$\sqrt{n}\left(\hat{\Phi}(m) - \Phi(m) - \mathcal{B}_T(m)\right) \quad \Rightarrow \quad N\left(0, V_\Phi\right), \quad \text{(A44)}$$

$$\sqrt{n}\left(\hat{Q}(p) - Q(p) + \frac{\mathcal{B}_T(Q(p))}{\phi(Q(p))}\right) \quad \Rightarrow \quad N\left(0, V_Q\right), \quad \text{(A45)}$$

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

Proof of Proposition A.2. The proof builds on our previous analysis. From Equation (A1), 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), Gordy (2003), and Gagliardini and Gourieroux (2011), we obtain:

$$P\left[m_i + \frac{1}{\sqrt{T}}\hat{\eta}_{i,T} \leq m\right] = \Phi(m) - \frac{1}{\sqrt{T}}\phi(m)E[\hat{\eta}_{i,T}|m_i = m] + \frac{1}{2T}\frac{d}{dm}(\phi(m)E[\hat{\eta}_{i,T}^2|m_i = m]) + o(1/T). \quad \text{(A46)}$$

From Equation (A44), 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.[6] With our bivariate Gaussian reference model, the term $\psi^{(1)}(m)$ in the bias is equal to

$$\begin{aligned}\psi^{(1)}(m) &= \exp\left(\mu_s + \rho\sigma_s\left(\frac{m-\mu_m}{\sigma_m}\right) + \frac{1}{2}\sigma_s^2(1-\rho^2)\right)\phi(m)\left(\frac{\sigma_s\rho}{\sigma_m} - \frac{m-\mu_m}{\sigma_m^2}\right) \\ &= \exp\left(\mu_s + \frac{1}{2}\sigma_s^2\right)\frac{-1}{\sigma_m}\left(\frac{m-\mu_m-\rho\sigma_s\sigma_m}{\sigma_m}\right) \\ &\quad \times\frac{1}{\sigma_m}\varphi\left(\frac{m-\mu_m-\rho\sigma_s\sigma_m}{\sigma_m}\right). \end{aligned} \quad \text{(A47)}$$

Similar to the numerical approach used for the moments, we can also compute the bias-adjusted proportion and quantile via a numerical integration of the density, i.e., we have

$$\Phi(m) = \int_{-\infty}^{m}\phi(u)du, \quad \text{(A48)}$$

$$\int_{-\infty}^{Q(p)}\phi(u)du = p, \quad \text{(A49)}$$

where we replace $\phi(u)$ with the bias-adjusted density estimator $\tilde{\phi}(u)$ to obtain the bias-adjusted estimators of $\Phi(m)$ and $Q(p)$ (for the quantile, we use an iterative procedure until Equation (A49) holds). Asymptotically, both approaches (analytical or numerical) are equivalent.

[6] 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.

To estimate the asymptotic variances V_Φ and V_Q, we then use the bias-corrected estimated proportion and quantile to compute $\Phi(m)(1-\Phi(m))$ and $\frac{p(1-p)}{\phi(Q(p))^2}$, where ϕ is the normal density obtained from the Gaussian reference model.

III Analysis of the EIV Bias

A The Shape of the EIV Bias

In this section, we provide additional information on the EIV bias obtained with the Gaussian reference model. As shown in Equation (15), $bs_2^r(m)$ depends on the number of observations T because it arises from the gap between $\hat{m}_i$ and m_i. Therefore, the EIV bias remains significant even if the fund population is large. The EIV bias adjustment exhibit two main features. First, it removes probability mass from the tails of the unadjusted density $\hat{\phi}(m)$. Therefore, the bias-adjusted density $\tilde{\phi}(m)$ is more peaked at the center. Formally, the sign of $bs_2^r(m)$ depends on the term $(\bar{m}_2^2-1)$ which is negative in the center and positive in the tails. The intuition behind this result is that the estimated measure is a noisy version of the true measure ($\hat{m}_i = m_i +$ estimation noise), which inflates the probability of observing extreme levels for m_i (i.e., $\hat{\phi}(m)$ is too flat).

Second, the EIV bias adjustment can be asymmetric. Therefore, the bias-adjusted density $\tilde{\phi}(m)$ can be significantly more skewed than the unadjusted density $\hat{\phi}(m)$. The degree of asymmetry depends on the correlation ρ between the true measure and the estimation variance. As noted by Jones and Shanken (2005), ρ depends on the concentration of the fund holdings. If, for instance, skilled funds only hold a few stocks, the idiosyncratic variance increases and the correlation between the fd alpha a_i and its estimation variance S_i is positive. In this case, the adjustment induces positive skewness because the probability mass is not transfered around the mean, but to its right ($\mu_m + \rho\sigma_m\sigma_s > \mu_m$). The intuition behind this result is that funds with high or low estimated fd alphas are more likely to have a higher estimation variance and thus a true skill measure (i.e., a large $|\hat{m}_i|$ means a higher s_i and a positive m_i).

B Comparative Static Anaysis

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 fd alpha. 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 0.26% 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 0.23. 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 0.23 to 0.46. 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 increases $\sigma_{\hat{m}}^2$ implies an increase in probability transfer from 15% to 26%. Finally, Panel C shows 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 (fd alpha and size coefficient) obtained with our nonparametric approach. We consider a hypothetical population of n_χ funds with a number of observed returns equal to T_χ, where $n_\chi = \sum_{i=1} \mathbf{1}_i^\chi$ and $T_\chi = \frac{1}{n_\chi} \sum_{i=1} \mathbf{1}_i^\chi T_i$ ($n_\chi = 1,000$, 2,500, 5,000, and 10,000; $T_\chi = 100$, 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 in Equation (4),

$$r_{i,t} = a_i - b_i q_{i,t-1} + \beta_i' f_t + \varepsilon_{i,t}, \tag{A50}$$

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

$$lq_{i,t} = \theta_{lq_i} + \rho_{lq} lq_{i,t-1} + \varepsilon_{lq_i,t}, \tag{A51}$$

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 $\varepsilon_{lqi,t}$ are drawn from a bivariate normal: $\varepsilon_{i,t} \sim N(0, \sigma^2_{\varepsilon_i})$, $\varepsilon_{lq_i,t} \sim N(0, \sigma^2_{\varepsilon_{lq_i}})$, where $\sigma^2_{\varepsilon_{lq_i}} = (1-\rho^2_{lq})\sigma^2_{lq_i}$ 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 $corr(\varepsilon_{i,t}, \varepsilon_{lqi,t})$ equal to ρ_ε.

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 $\{\hat{a}_i,\ \hat{b}_i,\ \hat{\beta}'_i,\ \hat{\mu}_{q_i},\ \hat{\sigma}^2_{q_i}\}$. This approach allows us to maintain the correlation structure between the different parameters, in particular between skill, scale, and size: $\mu_{lq_i} = \mu_{q_i}(a_i, b_i)$, $\sigma^2_{lq_i} = \sigma^2_{lq_i}(a_i, b_i)$. The remaining parameters are calibrated using the median values in the data, which yields: $\rho_{lq} = 0.97$, $\rho_\varepsilon = 0.20$, and $\sigma^2_{\varepsilon_i} = 0.0167^2$. We also have $n = 1.7n_\chi$ and $T = 2.4T_\chi$.

To reproduce the salient features of the skill distributions, we rescale the estimated values of $\hat{a}_i$ and $\hat{b}_i$ to match the cross-sectional volatility reported in Tables II and III (3.5% and 1.6% per year for a_i and b_i). The true distributions of the fd alpha and size coefficient 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 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 $\varepsilon_{lq_i,1}(s)$ from the bivariate normal. Third, we construct the fund gross return and log size at time 1 as

$$\begin{aligned} r_{i,1}(s) &= a_i - b_i q_{i,0}(s) + \beta'_i f_1(s) + \varepsilon_{i,1}(s), \\ lq_{i,1}(s) &= \theta_{lq,i} + \rho_{lq} lq_{i,0}(s) + \varepsilon_{lq_i,1}(s), \end{aligned} \tag{A52}$$

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

$$MISE = \int [\sigma^2(m) + bs(m)^2] dm, \tag{A53}$$

where the bias and variance functions are given by

$$bs(m) = \frac{1}{S}\sum_{s=1}^{S} \tilde{\phi}(m; s) - \phi(m), \tag{A54}$$

$$\sigma^2(m) = \frac{1}{S}\sum_{s=1}^{S}\left(\tilde{\phi}(m; s) - \frac{1}{S}\sum_{s=1}^{S}\tilde{\phi}(m; s)\right)^2. \tag{A55}$$

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

$$MSE(\tilde{\varphi}) = \sigma^2(\tilde{\varphi}) + bs^2(\tilde{\varphi}), \tag{A56}$$

where the bias and the variance terms are given by

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

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

B Main Results

In Table AI, we report the MISE and its two components (integrated squared bias and variance) for the fd alpha distribution. In Panels A and B, we compute the MISE of the bias-adjusted density $\tilde{\phi}(m)$ for each optimal bandwidth h^* (case (i) or (ii)). For comparison, 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 is critical for improving the estimation of the true distribution $\phi(m)$. To illustrate, we consider the scenario where $n_\chi = 2,500$ and $T_\chi = 250$, which is representative of our actual sample. We find that the MISE of $\hat{\phi}(m)$ is nearly two times larger than the level observed for $\tilde{\varphi}(m)$ with bandwidth (i) (4.96 vs 9.06). Second, our nonparametric approach yields a stronger performance when bandwidth (i) is used. In all scenarios, it produces a lower MISE than the one obtained with bandwidth (ii).

In Table AII, we examine the performance of the moment, and proportion estimators for the fd alpha. Panels 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 bandwidth (i)). Panels B reports the same statistics for the bias-adjusted estimators obtained with the analytical approach based on linearization. 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 $n_{\chi} = 2,500$ and $T_{\chi} = 250$, the bias for the volatility and the unadjusted proportion is equal to 1.08% per year and -5.52%, respectively. In contrast, our non-parametric approach largely 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 size coefficient. Tables AIII and AIV report the MISE of the estimated density and the MSE of the moment, quantile, and proportion estimators. Similar to the fd alpha, we find that $\tilde{\varphi}(m)$ largely outperforms $\hat{\phi}(m)$. If $n_{\chi} = 2,500$ and $T_{\chi} = 250$, 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 size coefficient by 6.8% (in absolute terms).

Please insert Tables AI to AIV here

To sum up, the Monte Carlo analysis yields three main insights. First, the EIV bias has a significant impact on the different estimators and thus cannot be ignored. Second, the choice of bandwidth (i) over bandwidth (ii) produces a lower MISE for the bias-adjusted density. Third, the numerical approach generally outperforms the analytical approach. Motivated by these results, we therefore use bandwidth (i) and the numerical approach for the empirical analysis reported in the paper.

C Simulations with Uncorrelated Skill and Scale

The asymptotic distribution of the OLS estimators $\hat{a}_i$ and $\hat{b}_i$ implies that they are correlated at the fund level, i.e., we have

$$\sqrt{T}\begin{bmatrix} \hat{a}_i - a_i \\ \hat{b}_i - b_i \end{bmatrix} \xrightarrow{d} N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \frac{E[q_{i,t-1}^2]}{V[q_{i,t-1}]}\sigma_{\varepsilon_i}^2 & \frac{E[q_{i,t-1}]}{V[q_{i,t-1}]}\sigma_{\varepsilon_i}^2 \\ \frac{E[q_{i,t-1}]}{V[q_{i,t-1}]}\sigma_{\varepsilon_i}^2 & \frac{1}{V[q_{i,t-1}]}\sigma_{\varepsilon_i}^2 \end{bmatrix}\right), \tag{A59}$$

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 skill and scale a_i and b_i of each fund i ($i = 1, ..., 2, 106$) independently from the estimated vectors observed in our sample (rescaled to match the cross-sectional volatility in Tables II and III). 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 $\sigma_{q_i}^2$ is proportional to μ_{q_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 $\mu_{q,i} + \frac{1}{2}\sigma_{q_i}^2 = \log\left(\frac{a_i}{2b_i}\right) \Rightarrow \left(1 + \frac{1}{2}k\right)\mu_{q,i} = \log\left(\frac{a_i}{2b_i}\right) \Rightarrow \mu_{q,i} = \log\left(\frac{a_i}{2b_i}\right) / \left(1 + \frac{1}{2}k\right)$. With a log-normally distributed size, we can then compute the parameters of the asymptotic distribution in Equation (A59) as $E[q_{i,t-1}^2] = e^{2\mu_{q,i} + 2k\mu_{q,i}}$ and $V[q_{i,t-1}] = E[q_{i,t-1}^2] - (E[q_{i,t-1}])^2$. Otherwise, if a_i and b_i are negative, we measure $E[q_{i,t-1}]$, $E[q_{i,t-1}^2]$, 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. We compute the average fund-level correlation (FLC) between $\hat{a}_i$ and $\hat{b}_i$ as

$$FLC(\hat{a}_i, \hat{b}_i) = \frac{1}{n}\sum_i \left(\frac{1}{S}\sum_s (\hat{a}_i(s) - a_i)\left(\hat{b}_i(s) - b_i\right)\right), \tag{A60}$$

and the average cross-sectional correlation (CSC) as

$$CSC(\hat{a}_i, \hat{b}_i) = \frac{1}{S}\sum_s \left(\frac{1}{n}\sum_i (\hat{a}_i(s) - \bar{a}(s))\left(\hat{b}_i(s) - \bar{b}(s)\right)\right), \tag{A61}$$

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)

D Industry Concentration with Uncorrelated Skill and Scale

We can use the same setup to measure the concentration of the industry when skill and scale are uncorrelated. For each iteration ($s = 1, ..., 500$), we draw the true skill and scale $a_i(s)$ and $b_i(s)$ of each fund i ($i = 1, ..., 2, 106$) independently, and use the same rule to compute $E[q_{i,t-1}](s)$. Then, we compute the total value added at average size for the entire population (all) and for the top 5% of the funds (top):

$$\begin{aligned} vaall^{lr}(s) &= \sum_i va_i^{lr}(s) = \sum_i a_i(s)E[q_{i,t-1}](s) - b_i(s)(E[q_{i,t-1}](s))^2, \\ vatop^{lr}(s) &= \sum_{i \in I(s)} va_i^{lr}(s) = \sum_{i \in I(s)} a_i(s)E[q_{i,t-1}](s) - b_i(s)(E[q_{i,t-1}](s))^2, \end{aligned} \quad \text{(A62)}$$

where $I(s)$ is the set of funds such that $va_i^{lr}(s) > Q_{0.95}(s)$. Finally, we can measure the concentration of the industry (IC) as

$$IC = \frac{1}{S}\sum_s \frac{vatop^{lr}(s)}{vaall^{lr}(s)}. \quad \text{(A63)}$$

We find that the concentration increases substantially once skill and scale are uncorrelated, i.e., IC is equal to 61% (versus 29.1% in our sample with correlated skill and scale).

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 2018 (528 observations). We measure the monthly gross return as the sum of the fund 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 measure the 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.[7]

We apply a set of filters before conducting the empirical analysis. First, we remove all funds which are classifed as passive or closed for more than 25% of the observations using (i) the indexfund indicator (letter B, D, or E), (ii) the ETFindicator (letter F or N), (iii) and the closedfund indicator (letter N). Our final set of funds includes all 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 25% of the observations (see 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 obtain our final universe of funds, we apply the selection rules in Equation (6). 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 ($CN_i \leq 15$ and $\tau_{i,T} \leq 8.8$). These selection criteria produce a final universe of 2,290 funds. To apply our nonparametric approach, we compute the asymptotic variance of the skill measure using a lag of three months ($L = 3$). To mitigate the impact of outliers on the 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

We form two groups of funds based on their investment styles (small cap, large cap). At the start of each month, we classify each fund in the different 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 2,290 equity funds. In addition, it shows the

[7]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.

mapping between the 32 styles and the growth/value and small/large cap dimensions. A value of: (i) 1 refers to a growth (small cap) fund, (ii) 2 refers to a "neutral" fund, and (iii) 3 refers to a value (large cap) fund. A fund is included in a given group if its style corresponds to that of the group for a sufficiently long period such that the two selection rules in Equation (6) are satisfied ($CN_i \leq$15 and $\tau_{i,T} \leq 8.8$).

Please insert Table AV here

We follow a similar procedure to form four groups of funds based on their characteristics (turnover and expense ratios). At the start of each month, we form terciles of funds based on their turnover and expense ratios. To measure the monthly turnover, we follow Pastor, Stambaugh, and Taylor (2017) and use the most recently observed ratio of min(buys,sells) on fund size. We eliminate observations for the annual expense ratio that are below 0.3% or above 10%. We also eliminate observations for the annual turnover that are below 30% or above 1000%.

VI Additional Results

A Test for the Panel Regression Approach

There is an extensive literature on mutual fund diseconomies of scale which uses a panel regression approach ($b_i = b$). To see if the panel specification is consistent with the data, 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: $b_i \neq b$ for a non-zero fraction of pairwise slopes for $i \neq j$. We denote by r_i and q_i the T_i-vector of fund excess gross returns and fund sizes, by F_i the $T_i \times K$ matrix of the factors, and by $Z_i = (\iota_{T_i}, F_i)$ the $T_i \times (K+1)$ matrix built from T_i-vector of ones ι_{T_i} and F_i.

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:

$$\hat{S} = \sum_i (\hat{b}_i - \hat{b}_{WFE})^2 \frac{q_i' M_i q_i}{\hat{\sigma}_i^2}, \tag{A64}$$

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 size coefficient of each fund, $\hat{b}_{WFE} = \left(\sum_i \frac{q_i'M_iq_i}{\hat{\sigma}_i^2}\right)^{-1} \left(\sum_i \frac{q_i'M_ir_i}{\hat{\sigma}_i^2}\right)$ 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 the test statistic

$$\hat{\Delta} = \sqrt{n} \left(\frac{\frac{1}{n}\hat{S} - 1}{\sqrt{2}} \right) \tag{A65}$$

is asymptotically distributed as a standard Gaussian random variable when $n, T \to \infty$ such that $\sqrt{n}/T^2_{\min} \to 0$ with $T^2_{\min} = \min_{1 \leqslant i \leqslant n} T_i$. Therefore, we can build the chi-square test statistic $\hat{\Delta}^2$ which is asymptotically distributed as a chi-square random variable χ^2_1 with one degree of freedom.[8]

We examine two specifications for the panel regression: (i) the linear specification $\alpha_{i,t} = a_i - b q_{i,t-1}$, and (ii) the log specification $\alpha_{i,t} = a_i - b\log(q_{i,t-1})$. We also conduct the test in the entire population and within each category (small/large cap and low/high turnover). Examining each category separately allows us to determine whether grouping stocks into well-defined investment styles absorbs the heterogeneity across funds. Our results reveal that the test of homogeneity is always strongly rejected, i.e., for each specification (size, log size) and each group of funds, we reject H_0 with probability one.

B Alternative Asset Pricing Models

Our analysis of mutual fund skill may depend on the choice of the asset pricing model. To address this issue, we estimate the skill and scale distributions (fd alpha and size coefficient) using the four-factor model of Carhart (1997) and the five-factor model of Fama and French (2015). We report the summary statistics for both distributions in Tables AVI and AVII.

(Carhart model). Overall, we find that the distributions of the two skill dimensions remain qualitatively unchanged. The fd alpha is equal to 2.6% per year on average, and is positive for 83.0% of the funds (vs 3.2% and 87.1% for the baseline results). The size coefficient is, on average, equal to 1.4% per year and 84.3% of the funds have a positive coefficient (vs 1.5% and 86.2%). We observe the main difference for the small cap group in which the annual fd alpha drops from 4.9% to 3.4% on average. This sharp reduction arises because the Carhart model assigns a negative alpha to the Russell 2000 index, which penalizes small cap funds.

(Five-factor model). Next, we focus on the five-factor model of Fama and French

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

(2015) which includes the market, size, value, profitability, and investment factors.[9] Using this model reduces the proportion of funds with a positive fd alpha from 87.1% to 75.9%. This reduction suggests that some funds achieve a positive fd alpha because they implement profitability- and investment-based strategies.

Please insert Tables AVI and AVII here

C Fund Size and the Small Sample Bias

It is well-known that the OLS estimation of the coefficients in Equation (4) is subject to small-sample bias (e.g., Stambaugh (1999)). While this bias disappears as the sample grows large, it could potentially be an issue for funds with short return histories. The small sample bias arises because the mutual fund error term $\varepsilon_{i,t}$ is contemporaneously correlated with the change in size $\varepsilon_{qi,t}$, where $\varepsilon_{qi,t}$ is the error term of the size regression: $q_{i,t} = \theta_{q_i} + \rho_{q_i} q_{i,t-1} + \varepsilon_{q_i,t}$. Projecting $\varepsilon_{q_i,t}$ onto the space spanned by the factors f_t, we obtain a new innovation vector denoted by $e_{q_i,t}$, i.e., $e_{q_i,t} = \varepsilon_{q_i,t} - \beta'_{qi} x_t$, where $x_t = (1, f'_t)'$. We can then rewrite the fund error term as $\varepsilon_{i,t} = +\phi_i e_{q_i,t} + \upsilon_{i,t}$ and insert this expression in Equation (4) to obtain

$$r_{i,t} = a_i - b_i q_{i,t-1} + \beta'_i f_t + \phi_i e_{q_i,t} + \upsilon_{i,t}. \tag{A66}$$

As noted by Amihud and Hurvich (2004), adding the regressor $e_{qi,t}$ eliminates the small-sample bias because the orthogonality holds, i.e., $E[\upsilon_i \,|X] = 0$, where $\upsilon_i = (\upsilon_{i,1}, ..., \upsilon_{i,T})'$ and $X = (X_1, ..., X_T)'$, and $X_t = (1, -q_{i,t-1}, f'_t, e_{q_i,t})'$.

Of course, we do not observe the true innovation term $e_{qi,t}$. To address this issue, we compute a proxy $e^c_{qi,t}$ using the procedure proposed by Amihud and Hurvich (2004). First, we estimate the AR(1) model for size: $q_{i,t} = \hat{\theta}_{q_i} + \hat{\rho}_{q_i} q_{i,t-1} + \hat{\varepsilon}_{q_i,t}$. Second, we compute $\varepsilon^c_{q_i,t}$ as

$$\varepsilon^c_{q_i,t} = q_{i,t} - (\hat{\theta}^c_{q_i} + \hat{\rho}^c_{q_i} q_{i,t-1}). \tag{A67}$$

The coefficients $\hat{\theta}^c_{q_i}$ and $\hat{\rho}^c_{q_i}$ are the second-order bias corrected coefficients:

$$\begin{aligned} \hat{\rho}^c_{q_i} &= \hat{\rho}_i + (1 + 3\hat{\rho}_{q_i})/T_i + 3(1 + 3\hat{\rho}^2_i)/T^2_i, \\ \hat{\theta}^c_{q_i} &= (1 - \hat{\rho}_i)\overline{q}_i, \end{aligned} \tag{A68}$$

where T_i denotes the total number of fund return observations and $\overline{q}_i$ is the average fund size. Third, we regress $\varepsilon^c_{qi,t}$ on the factors to obtain $e_{q_i,t} = \varepsilon^c_{q_i,t} - \hat{\beta}'_q x_t$. Fourth, we

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

insert $e^c_{q_i,t}$ in Equation (A66) to obtain bias-corrected estimators of a_i and b_i from the following OLS regression:

$$r_{i,t} = a_i - b_i q_{i,t-1} + \beta_i' f_t + \phi_i \varepsilon^c_{qi,t} + \upsilon_{i,t}. \tag{A69}$$

Finally, we can apply our nonparametric approach to obtain the bias-adjusted distributions of skill. The inclusion of the estimated variable $e^c_{q_i,t}$ does not change the properties of the nonparametric density estimator—in particular, the smoothing and EIV biases in Equations (A27)-(A28) remain unchanged because $\varepsilon^c_{q_i,t}$ only affects the higher order terms beyond T^{-1}.

We show the summary statistics for the skill and scale distributions in Table AVIII. We find that our main conclusions remain qualitatively unchanged. For one, the proportions of funds with positive fd alpha and size coefficient are equal to 80.7% and 79.8% (vs 87.1% and 86.2% for the baseline results). Therefore, these results suggest that the sample size is sufficiently large to mitigate the impact of the small sample bias. A noticeable change is the reduction in the average size coefficient from 1.5% to 1.4% per year, which is consistent with theory. As noted by Stambaugh (1999), we have $E[\hat{b}_i - b_i] = -E[\hat{\rho}_{q_i} - \rho_{q_i}]\phi_i$, which, in our case, is positive.

Please insert Table AVIII here

D Alternative Specifications for the Value Added

The analysis of the value added depends on the linear specification for the gross alpha (Equation (1)). To mitigate the concern that misspecification drives our empirical results, we examine two alternative specifications for the gross alpha. First, we use a log specification in which

$$\alpha_{i,t} = a_i - b_i \log(q_{i,t-1}) = a_i - b_i l q_{i,t-1}. \tag{A70}$$

Using the same estimation procedure as in Section III.A, we obtain the following estimates of the two measures of the value added:

$$\begin{aligned} \underset{\text{(average)}}{\text{Value added}} \quad &: \quad \hat{m}_i = \widehat{va}_i = \hat{a}_i \overline{q}_i - \hat{b}_i \overline{q}_{i,3}, \\ \underset{\text{(long run)}}{\text{Value added}} \quad &: \quad \hat{m}_i = \widehat{va}_i^{lr} = \hat{a}_i \overline{q}_i - \hat{b}_i \overline{lq}_i \overline{q}_i, \end{aligned} \tag{A71}$$

where $\overline{q}_{i,3} = \frac{1}{T_i}\sum_{t=1}^T I_{i,t} l q_{i,t-1} q_{i,t-1}$, and $\overline{lq}_i = \frac{1}{T_i}\sum_{t=1}^T I_{i,t} l q_{i,t-1}$. For each of these measures, the term $u_{i,t}$ is given by

$$
\begin{aligned}
\underset{\text{(average)}}{\text{Value added}} \quad : \quad & u_{i,t} = E[q_{i,t-1}]e_1' Q_{x,i}^{-1} x_{i,t}\varepsilon_{i,t} + a_i(q_{i,t-1} - E[q_{i,t-1}]) \\
& - E[lq_{i,t-1}q_{i,t-1}]e_2' Q_{x,i}^{-1} x_{i,t}\varepsilon_{i,t} - b_i(lq_{i,t-1}q_{i,t-1} - E[lq_{i,t-1}q_{i,t-1}]), \\
\underset{\text{(long run)}}{\text{Value added}} \quad : \quad & u_{i,t} = E[q_{i,t-1}]e_1' Q_{x,i}^{-1} x_{i,t}\varepsilon_{i,t} + a_i(q_{i,t-1} - E[q_{i,t-1}]) \\
& - E[lq_{i,t-1}]E[q_{i,t-1}]e_2' Q_{x,i}^{-1} x_{i,t}\varepsilon_{i,t} - b_i E[q_{i,t-1}](lq_{i,t-1} - E[lq_{i,t-1}]) \\
& - - b_i E[lq_{i,t-1}](q_{i,t-1} - E[q_{i,t-1}]), \qquad \text{(A72)}
\end{aligned}
$$

where e_1 (e_2) is a vector with one in the first (second) position and zeros elsewhere and $Q_{x,i} = E[x_{i,t}x_{i,t}']$, where $x_{i,t} = (1, -lq_{i,t-1}, f_t')'$.

Second, we use a general specification

$$\alpha_{i,t} = a_i - b_i g_i(q_{i,t-1}), \qquad \text{(A73)}$$

where $g_i(q_{i,t-1})$ is an (unspecified) general function which is assumed to be uncorrelated with the factors f_t.[10] To begin, we estimate the following unconditional time-series regression for each fund i:

$$r_{i,t} = \alpha_i^u + \beta_i^{u\prime} f_t + \varepsilon_{i,t}^u, \qquad \text{(A74)}$$

where $r_{i,t}$ is the fund gross excess return (before fees) over the riskfree rate, f_t is a K-vector of benchmark excess returns, and $\varepsilon_{i,t}^u$ is the error term. Using the estimated coefficients, we can compute the two measures of the value added without having to specify $g_i(q_{i,t-1})$:

$$
\begin{aligned}
\underset{\text{(average)}}{\text{Value added}} \quad : \quad & \hat{m}_i = \widehat{va}_i = \frac{1}{T}\sum_{t=1}^{T} I_{i,t}(\hat{\alpha}_i^u + \hat{\varepsilon}_{i,t}^u)q_{i,t-1} = \frac{1}{T}\sum_{t=1}^{T} I_{i,t}(r_{i,t} - \hat{\beta}_i^{u\prime} f_t)q_{i,t-1}, \\
\underset{\text{(long run)}}{\text{Value added}} \quad : \quad & \hat{m}_i = \widehat{va}_i^{lr} = \left(\frac{1}{T}\sum_{t=1}^{T} I_{i,t}(\hat{\alpha}_i^u + \hat{\varepsilon}_{i,t}^u)\right)\overline{q}_i \\
= \quad & \left(\frac{1}{T}\sum_{t=1}^{T} I_{i,t}(r_{i,t} - \hat{\beta}_i^{u\prime} f_t)\right)\overline{q}_i. \qquad \text{(A75)}
\end{aligned}
$$

[10] This estimation procedure is used by Berk and van Binsbergen (2015) to estimate the average value added (they do not look at the second measure—the long-run value added).

For these measures, the term $u_{i,t}$ is given by

$$\underset{\text{(average)}}{\text{Value added}} \;:\; u_{i,t} = r_{i,t}q_{i,t-1} - E[r_{i,t}q_{i,t-1}] - \sum_{k=1}^{K} E[f_{t,k}q_{i,t-1}]e'_{k+1}Q_{x,i}^{-1}x_{i,t}\varepsilon_{i,t}$$
$$- \sum_{k=1}^{K} \beta_{i,k}^{u}(f_{t,k}q_{i,t-1} - E[f_{t,k}q_{i,t-1}]),$$
$$\underset{\text{(long run)}}{\text{Value added}} \;:\; u_{i,t} = q_{i,t-1}(r_{i,t} - E[r_{i,t}]) + r_{i,t}(q_{i,t-1} - E[q_{i,t-1}])$$
$$- \sum_{k=1}^{K} E[f_{t,k}]E[q_{i,t-1}]e'_{k+1}Q_{x,i}^{-1}x_{i,t}\varepsilon_{i,t} - \sum_{k=1}^{K} \beta_{i,k}^{u}(q_{i,t-1}(f_{t,k} - E[f_{t,k}])$$
$$- \sum_{k=1}^{K} \beta_{i,k}^{u}(f_{t,k}(q_{i,t-1} - E[q_{i,t-1}]), \tag{A76}$$

where e_1 (e_{k+1}) is a vector with one in the first (k^{th}+1) position and zeros elsewhere and $Q_{x,i} = E[x_{i,t}x'_{i,t}]$ and $x_{i,t} = (1, f'_t)'$.

For both specifications, Tables AIX and AX produce the same results as those obtained with the baseline specification. For instance, the average value added is equal to \$1.5M and \$1.7M per year under the two alternative specifications (vs \$1.4M for the baseline specification). Similarly, the long-run value added is equal to \$7.9M and \$8.0M per year under the two alternative specifications (vs \$7.8M for the baseline specification).

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

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

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

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

[6] 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.

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

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

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

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

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

[12] Gordy, M., 2003, A Risk-Factor Model Foundation for Ratings Based Bank Capital Rules, Journal of Financial Intermediation 12, 199-232.

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

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

[15] Jones, C.S., and J. Shanken, 2005, Mutual Fund Performance with Learning across Funds, Journal of Financial Economics 78, 507-552.

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

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

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

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

[20] Pastor, L., R. F. Stambaugh, and L. Taylor, 2017, Do Funds Make More When They Trade More?, Journal of Finance 72, 1483-1528.

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

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

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

Table AI
Properties of the Estimated Density:
First-Dollar Alpha

Panel A and B show the Mean Integrated Squared Error (MISE) and its two components (integrated squared bias and variance) for the bias-adjusted density under the two choices of bandwidth (bandwidth (i) and (ii)) across different values for the number of available funds and the number of available monthly observations. For comparison, Panel C reports the same information for the bias-unadjusted density.

Panel A: Bias Adjustment (Bandwidth 1)

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 (Bandwidth 2)

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: First-Dollar Alpha

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 (bandwidth (i)) across different values for the number of available funds and the number of available 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:
First-Dollar Alpha (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:
First-Dollar Alpha (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:
Size Coefficient

Panel A and B show the Mean Integrated Squared Error (MISE) and its two components (integrated squared bias and variance) for the bias-adjusted density under the two choices of bandwidth (bandwidth (i) and (ii)) across different values for the number of available funds and the number of available monthly observations. For comparison, Panel C reports the same information for the bias-unadjusted density.

Panel A: Bias Adjustment (Bandwidth 1)

	MISE					Bias^2					Variance				
n\T	100	250	500	750	1000	100	250	500	750	1000	100	250	500	750	1000
1000	77.11	18.04	5.35	2.95	2.08	75.98	16.46	3.95	1.69	0.95	1.13	1.58	1.40	1.26	1.14
2500	75.05	17.01	4.63	2.15	1.55	74.51	16.29	4.03	1.64	1.05	0.53	0.72	0.61	0.52	0.50
5000	75.13	17.03	4.09	1.89	1.18	74.54	16.66	3.80	1.63	0.89	0.59	0.37	0.29	0.26	0.29
7500	74.82	16.78	4.07	1.79	1.04	74.12	16.57	3.87	1.60	0.85	0.70	0.21	0.20	0.19	0.19
10000	74.13	16.61	4.03	1.71	1.02	73.90	16.41	3.89	1.56	0.88	0.23	0.19	0.14	0.15	0.14

Panel B: BiasAdjustment (Bandwidth 2)

	MISE					Bias^2					Variance				
n\T	100	250	500	750	1000	100	250	500	750	1000	100	250	500	750	1000
1000	91.92	20.42	5.85	3.16	2.12	90.72	19.42	5.14	2.45	1.45	1.21	1.00	0.71	0.71	0.67
2500	87.96	19.27	5.24	2.43	1.68	87.38	18.76	4.93	2.17	1.41	0.58	0.51	0.31	0.26	0.27
5000	86.18	18.79	4.63	2.18	1.32	85.45	18.55	4.47	2.03	1.16	0.73	0.24	0.17	0.15	0.17
7500	84.71	18.29	4.57	2.05	1.17	83.90	18.16	4.44	1.94	1.05	0.80	0.13	0.13	0.12	0.12
10000	83.13	17.98	4.48	1.95	1.15	82.88	17.85	4.38	1.85	1.06	0.24	0.13	0.09	0.10	0.09

Panel C: No Bias Adjustment

	MISE					Bias^2					Variance				
n\T	100	250	500	750	1000	100	250	500	750	1000	100	250	500	750	1000
1000	97.17	31.16	10.69	5.84	3.60	96.81	30.78	10.20	5.32	3.08	0.36	0.38	0.48	0.52	0.51
2500	92.06	28.58	9.41	4.25	2.66	91.89	28.39	9.20	4.06	2.44	0.18	0.19	0.22	0.20	0.22
5000	89.26	27.34	8.39	3.80	2.14	88.94	27.25	8.27	3.68	2.01	0.32	0.09	0.12	0.12	0.14
7500	87.56	26.59	8.14	3.52	1.86	87.13	26.52	8.05	3.43	1.76	0.43	0.07	0.09	0.09	0.10
10000	85.48	26.25	7.87	3.36	1.77	85.42	26.20	7.80	3.29	1.70	0.06	0.05	0.07	0.08	0.07

Table AIV
Properties of the Estimated Moments and Proportion: Size 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 (bandwidth (i)) across different values for the number of available funds and the number of available 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:
Size 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: Size 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 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
Cross-Sectional Distributions of Skill and Scale
Four-Factor Model

Panel A shows the summary statistics of the skill distribution (first dollar (fd) alpha) for all funds in the population and across small/large cap funds and low/high turnover funds (i.e., bottom and top terciles of funds sorted on turnover) using the Carhart model. It reports the first four moments, the proportions of funds with a negative and positive fd alpha, and the quantiles at 5% and 95%. Panel B reports the same summary statistics for the scale distribution (size coefficient). All cross-sectional estimates are computed by integrating numerically the bias-adjusted density obtained with our nonparametric approach. Figures in parentheses denote the estimated standard deviation of each estimator.

Panel A: Distribution of the First Dollar Alpha

	Moments				Proportions (%)		Quantiles (Ann.)	
	Mean (Ann.)	Std. Dev. (Ann.)	Skewness	Kurtosis	Negative	Positive	5%	95%
All Funds	2.6 (0.1)	3.4 (0.2)	0.2 (0.5)	13.1 (1.4)	17 (0.8)	83 (0.8)	-1.8 (0.2)	7.8 (0.2)
Investment Categories								
Small Cap	3.4 (0.2)	4 (0.3)	1 (0.4)	7.7 (1.7)	14.6 (1.4)	85.4 (1.4)	-2.1 (0.4)	10.1 (0.3)
Large Cap	1.8 (0.1)	2.6 (0.2)	1 (1.1)	17.6 (12.2)	21.5 (1.3)	78.5 (1.3)	-1.8 (0.2)	5.9 (0.2)
Low Turnover	2.3 (0.2)	4.2 (0.3)	0.5 (0.9)	20.5 (4.8)	23.4 (1.1)	76.6 (1.1)	-3.1 (0.2)	8.9 (0.3)
High Turnover	2.9 (0.2)	4.6 (0.2)	1.2 (0.4)	11.3 (2.1)	22.4 (1)	77.6 (1)	-3.3 (0.2)	10.2 (0.3)

Panel B: Distribution of the Size Coefficient

	Moments				Proportions (%)		Quantiles (Ann.)	
	Mean (Ann.)	Std. Dev. (Ann.)	Skewness	Kurtosis	Negative	Positive	5%	95%
All Funds	1.4 (0.1)	1.5 (0.1)	0.3 (0.5)	11.8 (2.1)	15.7 (0.7)	84.3 (0.7)	-0.7 (0.1)	3.8 (0.1)
Investment Categories								
Small Cap	1.9 (0.1)	1.8 (0.1)	-0.2 (0.5)	5.8 (1.5)	14.1 (1.4)	85.9 (1.4)	-1.1 (0.2)	4.9 (0.2)
Large Cap	0.9 (0.1)	1.2 (0.1)	1.6 (1.3)	23.2 (18.8)	21.9 (1.3)	78.1 (1.3)	-0.7 (0.1)	2.9 (0.1)
Low Turnover	1.1 (0.1)	1.7 (0.1)	-0.3 (0.7)	15.9 (3.5)	24 (1.1)	76 (1.1)	-1.3 (0.1)	3.6 (0.1)
High Turnover	1.6 (0.1)	2.2 (0.1)	0.5 (0.4)	7.6 (1.4)	21.8 (1)	78.2 (1)	-1.4 (0.1)	5.3 (0.2)

Table AVII
Cross-Sectional Distributions of Skill and Scale: Five-Factor Model

Panel A shows the summary statistics of the skill distribution (first dollar (fd) alpha) for all funds in the population and across small/large cap funds and low/high turnover funds (i.e., bottom and top terciles of funds sorted on turnover) using the Fama-French four-factor model. It reports the first four moments, the proportions of funds with a negative and positive fd alpha, and the quantiles at 5% and 95%. Panel B reports the same summary statistics the scale distribution (size coefficient). All cross-sectional estimates are computed by integrating numerically the bias-adjusted density obtained with our nonparametric approach. Figures in parentheses denote the estimated standard deviation of each estimator.

Panel A: Distribution of the First Dollar Alpha

	Moments				Proportions (%)		Quantiles (Ann.)	
	Mean (Ann.)	Std. Dev. (Ann.)	Skewness	Kurtosis	Negative	Positive	5%	95%
All Funds	2.4 (0.1)	3.9 (0.2)	0.4 (0.4)	11.3 (1.4)	24.1 (0.9)	75.9 (0.9)	-2.6 (0.2)	8.5 (0.2)
Investment Categories								
Small Cap	3.4 (0.2)	4.5 (0.3)	0.9 (0.4)	7.4 (2.1)	19.6 (1.6)	80.4 (1.6)	-2.8 (0.4)	10.7 (0.4)
Large Cap	1.6 (0.1)	3 (0.2)	1.7 (0.6)	17 (7.7)	29.8 (1.4)	70.2 (1.4)	-2.5 (0.2)	6.5 (0.3)
Low Turnover	1.6 (0.2)	4.4 (0.3)	0.1 (0.8)	16 (4.2)	34.1 (1.2)	65.9 (1.2)	-4 (0.2)	8.5 (0.3)
High Turnover	3.6 (0.2)	5.2 (0.2)	0.8 (0.4)	9.6 (1.2)	19.4 (1)	80.6 (1)	-3.3 (0.3)	11.6 (0.3)

Panel B: Distribution of the Size 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)	0 (0.5)	12.9 (2.2)	20 (0.8)	80 (0.8)	-0.9 (0.1)	4 (0.1)
Investment Categories								
Small Cap	1.7 (0.1)	1.9 (0.2)	-0.6 (0.7)	9.6 (3.2)	16 (1.5)	84 (1.5)	-1 (0.2)	4.7 (0.2)
Large Cap	0.9 (0.1)	1.4 (0.1)	0.7 (1.3)	20.7 (9.7)	27 (1.4)	73 (1.4)	-1 (0.1)	3 (0.1)
Low Turnover	0.9 (0.1)	1.7 (0.1)	-0.7 (0.6)	13.2 (4.3)	29.1 (1.1)	70.9 (1.1)	-1.5 (0.1)	3.6 (0.1)
High Turnover	1.5 (0.1)	2.4 (0.1)	0.4 (0.4)	9.8 (1.2)	23.4 (1)	76.6 (1)	-2 (0.2)	5.3 (0.2)

Table AVIII
Cross-Sectional Distributions of Skill and Scale: Small-Sample Bias

Panel A shows the summary statistics of the skill distribution (first dollar (fd) alpha) for all funds in the population and across small/large cap funds and low/high turnover funds (i.e., bottom and top terciles of funds sorted on turnover) using the adjustment of Amihud and Hurvich (2004) to control for the small-sample bias. It reports the first four moments, the proportions of funds with a negative and positive fd alpha, and the quantiles at 5% and 95%. Panel B reports the same summary statistics the scale distribution (size coefficient). All cross-sectional estimates are computed by integrating numerically the bias-adjusted density obtained with our nonparametric approach. Figures in parentheses denote the estimated standard deviation of each estimator.

Panel A: Distribution of the First Dollar Alpha

	Moments				Proportions (%)		Quantiles (Ann.)	
	Mean (Ann.)	Std. Dev. (Ann.)	Skewness	Kurtosis	Negative	Positive	5%	95%
All Funds	2.9 (0.1)	4.3 (0.2)	0.7 (0.4)	10.9 (1.8)	19.3 (0.8)	80.7 (0.8)	-1.4 (0.2)	7.6 (0.2)
Investment Categories								
Small Cap	4.5 (0.2)	5 (0.3)	0.7 (0.5)	8.2 (1.8)	14.3 (1.4)	85.7 (1.4)	-1 (0.4)	10.2 (0.4)
Large Cap	1.8 (0.1)	3 (0.2)	1 (0.5)	10 (4.4)	24.1 (1.3)	75.9 (1.3)	-1.5 (0.2)	5.4 (0.2)
Low Turnover	2.3 (0.2)	4.9 (0.2)	0.1 (0.5)	10.2 (2.6)	26 (1.1)	74 (1.1)	-2.7 (0.3)	7.9 (0.3)
High Turnover	3.5 (0.2)	6.2 (0.3)	0.7 (0.4)	9.1 (0.8)	24.7 (1.1)	75.3 (1.1)	-2.9 (0.3)	10.2 (0.3)

Panel B: Distribution of the Size Coefficient

	Moments				Proportions (%)		Quantiles (Ann.)	
	Mean (Ann.)	Std. Dev. (Ann.)	Skewness	Kurtosis	Negative	Positive	5%	95%
All Funds	1.4 (0.1)	1.9 (0.1)	1.1 (0.3)	9.5 (1)	20.2 (0.8)	79.8 (0.8)	-0.6 (0.1)	3.5 (0.1)
Investment Categories								
Small Cap	1.8 (0.1)	2.2 (0.1)	0.6 (0.3)	5.7 (0.7)	19.6 (1.6)	80.4 (1.6)	-0.9 (0.2)	4.3 (0.2)
Large Cap	1 (0.1)	1.5 (0.1)	2 (0.4)	15.4 (4.6)	23.5 (1.3)	76.5 (1.3)	-0.6 (0.1)	2.7 (0.1)
Low Turnover	0.9 (0.1)	2 (0.1)	0 (0.3)	6.9 (1.3)	28.5 (1.1)	71.5 (1.1)	-1.3 (0.1)	3.3 (0.1)
High Turnover	1.7 (0.1)	2.8 (0.1)	0.4 (0.3)	5.8 (0.5)	26.6 (1.1)	73.4 (1.1)	-1.5 (0.1)	4.9 (0.2)

Table AIX
Cross-Sectional Distributions of Value Added: Log Specification

Panel A shows the summary statistics of the distribution of the average value added over the sample period for all funds in the population and across small/large cap funds and low/high turnover funds (i.e., bottom and top terciles of funds sorted on turnover) using the log specification for the gross alpha. It reports the first four moments, the proportions of funds with a negative and positive value added, and the quantiles at 5% and 95%. Panel B reports the same summary statistics for the distribution of the long-run value added (measured at the average fund size). All cross-sectional estimates are computed by integrating numerically the bias-adjusted density obtained with our nonparametric approach. Figures in parentheses denote the estimated standard deviation of each estimator.

	Moments				Proportions (%)		Quantiles (Ann.)	
	Mean (Ann.)	Std. Dev. (Ann.)	Skewness	Kurtosis	Negative	Positive	5%	95%
All Funds	1.5 (0.3)	12.3 (1)	4.2 (1.2)	63.1 (6.7)	37.9 (1)	62.1 (1)	-6.4 (0.3)	18.8 (0.5)
Investment Categories								
Small Cap	3.4 (0.5)	12.6 (1.6)	4.7 (1.6)	38.5 (4.8)	34.9 (1.9)	65.1 (1.9)	-6.4 (0.5)	21 (1.1)
Large Cap	-0.5 (0.3)	9.8 (1.1)	1.2 (2.3)	55.3 (10.8)	51.2 (1.6)	48.8 (1.6)	-10 (0.5)	10.4 (0.6)
Low Turnover	3.7 (0.5)	19.3 (1.5)	1.5 (1.3)	39 (3.5)	30.6 (1.2)	69.4 (1.2)	-8.5 (0.6)	34 (0.8)
High Turnover	0.8 (0.4)	13.3 (1.4)	3.5 (1.8)	71.9 (11.1)	43.4 (1.2)	56.6 (1.2)	-10 (0.5)	18.1 (0.6)

Panel B: Distribution of the Long-Run Value Added

	Moments				Proportions (%)		Quantiles (Ann.)	
	Mean (Ann.)	Std. Dev. (Ann.)	Skewness	Kurtosis	Negative	Positive	5%	95%
All Funds	7.9 (0.5)	21.2 (1.1)	3.1 (0.3)	25.9 (3.2)	28.2 (0.9)	71.8 (0.9)	-17.2 (0.6)	38.1 (0.7)
Investment Categories								
Small Cap	8.2 (0.7)	16.3 (1.4)	2.9 (0.5)	19.8 (2.8)	17.8 (1.7)	82.2 (1.7)	-10 (0.9)	29.7 (1)
Large Cap	5.1 (0.5)	15.8 (1.1)	2.5 (0.4)	21.9 (7.3)	30.7 (1.4)	69.3 (1.4)	-12.8 (0.7)	30.4 (0.9)
Low Turnover	10.4 (0.7)	28.2 (1.1)	0.6 (0.5)	10 (2.1)	20.5 (1.1)	79.5 (1.1)	-21.5 (1)	62.8 (1.4)
High Turnover	8 (0.6)	23.5 (1.3)	2.7 (0.2)	20.7 (2.3)	29.2 (1.1)	70.8 (1.1)	-19.6 (0.8)	42.6 (0.9)

Table AX
Cross-Sectional Distributions of Value Added:
Fully Flexible Specification

Panel A shows the summary statistics of the distribution of the average value added over the sample period for all funds in the population and across small/large cap funds and low/high turnover funds (i.e., bottom and top terciles of funds sorted on turnover) using a fully flexible specification for the gross alpha. It reports the first four moments, the proportions of funds with a negative and positive value added, and the quantiles at 5% and 95%. Panel B reports the same summary statistics for the distribution of the long-run value added (measured at the average fund size). All cross-sectional estimates are computed by integrating numerically the bias-adjusted density obtained with our nonparametric approach. Figures in parentheses denote the estimated standard deviation of each estimator.

Panel A: Distribution of the Average Value Added

	Moments				Proportions (%)		Quantiles (Ann.)	
	Mean (Ann.)	Std. Dev. (Ann.)	Skewness	Kurtosis	Negative	Positive	5%	95%
All Funds	1.7 (0.3)	13.4 (1.1)	5.4 (1.2)	64.4 (7)	38.3 (1)	61.7 (1)	-6.6 (0.3)	20.8 (0.5)
Investment Categories								
Small Cap	3.1 (0.5)	11.2 (1.4)	4 (1.2)	38 (4.6)	34.4 (1.9)	65.6 (1.9)	-5.9 (0.5)	19.6 (0.9)
Large Cap	-0.1 (0.4)	12.4 (1.9)	6 (2.9)	97.5 (17.1)	47.1 (1.6)	52.9 (1.6)	-9.2 (0.5)	14.5 (0.6)
Low Turnover	4.3 (0.6)	21.6 (1.5)	2.8 (0.9)	31.2 (2.1)	29.8 (1.2)	70.2 (1.2)	-8.8 (0.6)	40.9 (0.9)
High Turnover	0.8 (0.3)	12.1 (1.2)	3.5 (1.5)	68.1 (13.5)	43.1 (1.2)	56.9 (1.2)	-9.5 (0.4)	17.4 (0.5)

Panel B: Distribution of the Long-Run Value Added

	Moments				Proportions (%)		Quantiles (Ann.)	
	Mean (Ann.)	Std. Dev. (Ann.)	Skewness	Kurtosis	Negative	Positive	5%	95%
All Funds	8 (0.5)	21.3 (1.1)	3.1 (0.3)	25.3 (3.2)	28.6 (0.9)	71.4 (0.9)	-17.6 (0.6)	37.8 (0.7)
Investment Categories								
Small Cap	8.2 (0.7)	16.4 (1.4)	2.9 (0.6)	20.1 (2.7)	18.1 (1.7)	81.9 (1.7)	-9.7 (0.8)	28.9 (1)
Large Cap	5.7 (0.6)	18.2 (1.4)	2.6 (0.4)	23.5 (6.8)	32.9 (1.4)	67.1 (1.4)	-16.4 (0.8)	33.5 (0.9)
Low Turnover	10.4 (0.7)	28.5 (1.1)	1 (0.4)	10.9 (2.1)	22.3 (1.1)	77.7 (1.1)	-22.9 (1)	61.6 (1.4)
High Turnover	8.1 (0.6)	23.4 (1.3)	2.6 (0.2)	20.2 (2.2)	29.3 (1.1)	70.7 (1.1)	-19.8 (0.8)	42.1 (0.9)

Figure A1
Comparative Static Analysis of the EIV Bias (First Dollar Alpha)

This figure performs a comparative static analysis of the EIV bias function for the first dollar (fd) alpha. 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 fd alpha by 0.002/100. In Panel B, we plot the new EIV bias function after increasing the variance of the estimated fd alpha by 0.002/100. In Panel C, we plot the new EIV bias function after increasing the correlation between the true fd alpha and the estimation variance by 50% in relative terms.