Skill and Value Creation in the Mutual Fund Industry

Laurent Barras, Patrick Gagliardini, and Olivier Scaillet*

This version, May 9, 2019

JEL Classification: G11, G12, C14, C33, C58.
Keywords: Mutual fund skill, non-parametric density estimation, large panel

*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, Jonathan Berk, Jean-Daniel Fermanian, Laurent Fresard, Alexander Kempf, Gaëlle Le Fol, Peter Limbach, Florian Weigert, and participants at the 18th Cologne Colloquium on Financial Markets, ISNPS 2018, McGill/HEC 2018 Spring Workshop, the 2018 SFI Research Days, the 2018 Paris December Conference, the 2018 EC2 Meeting, as well as seminar participants at the Universities of Geneva, McGill, McMaster, 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 simple, nonparametric approach for estimating the entire distribution of skill. Our approach avoids the challenge of correctly specifying the distribution, and accommodates the need to study both the investment and trading dimensions of skill. Our results show that most funds are skilled at detecting profitable trades, but unskilled at overriding capacity constraints. Aggregating both skill dimensions, we find overwhelming evidence that mutual funds produce significant value added. In addition, the active industry is (i) not concentrated because few funds are skilled on all dimensions, (ii) close to optimally sized as funds internalize the impact of capacity constraints, and (iii) in a strong bargaining position vis-a-vis the investors.
I Introduction

Over the past 50 years, the academic literature on mutual funds has largely focused on performance. For instance, Carhart (1997), Elton et al. (1993), and Jensen (1968) find that the aggregate alpha net of fees and trading costs is negative, while recent studies find the same result for the majority of funds (e.g., Barras, Scaillet, and Wermers (2010), Harvey and Liu (2018a)). Far less attention has been devoted to the analysis of mutual fund skill.\footnote{Notable exceptions include Berk and Green (2004), Berk and van Binsbergen (2015), Grinblatt and Titman (1989), Jones and Shanken (2005), Pastor, Stambaugh, and Taylor (2015), and Wermers (2000).} Whereas these two notions are often used interchangeably, they differ in important ways—a point forcefully made by Berk and van Binsbergen (2015; BvB hereafter). Skill is defined from the viewpoint of funds, i.e., it measures whether funds have investment abilities that allow them to create value. In contrast, performance is defined from the viewpoint of investors, i.e., it measures whether the value created by the funds, if any, is passed on to them.

Focusing on skill rather than performance is important for several reasons. First, the analysis of skill is informative about the prevalence of skilled funds in the population and the type of skill they exhibit. Second, it determines the extent to which funds earn economic rents from exploiting their investment abilities. Third, it allows us to study the structure of the mutual fund industry, i.e., whether it is concentrated, optimally sized, and in a strong bargaining position vis-a-vis their investors. Finally, it sheds light on the social value of active management. When skilled funds trade, they improve price efficiency and the real allocation of resources in the economy.

In this paper, we develop a novel approach for estimating the cross-sectional distribution of mutual fund skill. A unique feature of our analysis is to study several measures together. We begin by estimating the skill of each fund along two distinct dimensions: (i) its ability to detect profitable trades, and (ii) its ability to trade efficiently as its size increases. Therefore, our framework incorporates the widely held view that active funds are subject to capacity constraints (e.g., Berk and Green (2004; BG hereafter)). We then aggregate these two skill dimensions into a single measure that determines the total dollar value created by the fund (BvB).

Our approach is related to recent studies that use standard Bayesian/parametric approaches to infer the distribution of fund performance (e.g., Chen, Cliff, and Zhao (2017), Harvey and Liu (2018a), Jones and Shanken (2005)). However, it departs on one key aspect: it is nonparametric and thus does not require a full parametric specification of the distribution to be estimated. In the context of skill, we argue that this flexibility...
is essential. First, it avoids the challenge of correctly specifying the skill distribution. Theory offers little guidance to specify the mean or dispersion of skill across funds. In addition, explorative data analysis cannot guide specification either—as explained below, a simple histogram of the estimated skill measures is plagued by estimation errors and thus biased. Second, it accommodates the need to analyze multiple skill dimensions. Such analysis is challenging with standard approaches. A joint specification for all measures is subject to the curse of dimensionality, whereas a separate specification is vulnerable to inconsistencies because the skill measures are intertwined.

In addition to its flexibility, our novel approach is simple, widely applicable, and supported by econometric theory. It only involves simple manipulations of the fund skill measures estimated using time-series regressions. It provides a unified framework that is applicable to all the characterizations of the skill distribution, including its moments and quantiles. Finally, it allows for a solid statistical inference determined by the asymptotic properties of the different estimators.

The main estimation challenge is to adjust for the bias. Because the true skill measures are unobservable, we can only rely on the estimated measures to infer the skill distribution. This creates an Error-in-Variable (EIV) bias that is reminiscent of the well-known EIV bias in the two-pass regression (e.g., Jagannathan, Skoulakis, and Wang (2013), Kan, Robotti, and Shanken (2013), Shanken (1992)). To address this issue, we develop a simple procedure to adjust our nonparametric distribution. The bias adjustment is easy to interpret, available in closed form, and validated through an extensive Monte Carlo analysis. It is also essential to correctly measure skill—the unadjusted distribution overestimates the probability in the tails, does not capture the strong skill asymmetry, and wrongly signals that the majority of funds destroy value.

Our empirical analysis is based on all US active equity funds between January 1975 and December 2018 (2,291 funds). We measure the two skill dimensions using the economic model of BG in which the fund gross alpha $\alpha_{it}$ depends on its lagged size $q_{i,t-1}$:

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

The first dollar (fd) alpha $a_i$ captures the fund skill at detecting profitable trades, whereas the size coefficient $b_i$ captures the fund skill at mitigating capacity constraints. We then aggregate both dimensions by measuring two formulations of the value added. The lifecycle value added, defined as $va_{i,l} = E[\alpha_{i,t}q_{i,t-1}]$, determines the value created by the fund over its lifecycle—this is the formulation examined by BvB. The steady state value added, defined as $va_{i,ss} = E[\alpha_{i,t}] E[q_{i,t-1}]$, determines the value created by the fund once it reaches its average, steady state size $E[q_{i,t-1}]$.

Our results reveal several insights regarding the two skill dimensions. The ability of funds to detect profitable trades is both widespread and economically significant.
Controlling for the standard risk factors, we find that the fd alpha is positive for 86% of the funds, and reaches 3.1% per year on average. At the same time, only a handful of funds have the ability to override capacity constraints. The size coefficient is positive for around 85% of the funds, and causes an average 1.4% decrease in annual alpha following a one standard deviation increase in fund size. Therefore, these results provide overwhelming support to models that emphasize the importance of capacity constraints for the mutual fund industry (e.g., BG, Pastor and Stambaugh (2012)).

Our fund-level analysis uncovers a strong heterogeneity in skill. The cross-sectional volatility for both $a_i$ and $b_i$ is larger than the average—a finding that is inconsistent with the common practice of imposing constant values across funds to reduce estimation errors (e.g., Chen et al. (2004)). Both skill dimensions are also strongly correlated—the pairwise correlation between the estimated values $\hat{a}_i$ and $\hat{b}_i$ is equal to 0.82. While the BG model is silent on the drivers of skill, we find that part of the heterogeneity of $a_i$ and $b_i$ and their correlation is explained by the strategies followed by the funds. For instance, investing in small cap stocks involves illiquidity. As trading costs increases, it becomes more difficult to arbitrage any mispricing away. Consistent with this intuition, small cap funds are more skilled at detecting profitable trades than large cap funds (higher $a_i$), but are also more exposed to capacity constraints (higher $b_i$).

The majority of funds create a substantial value from exploiting their skill. The proportion of funds with a positive lifecycle and steady state value added is around 60% and 70%, respectively. On average, individual funds create 1.7 mio. per year over their lifecycle, but more than 7.3 mio. once they reach their steady state size. This sharp difference is driven by capacity constraints which lower the value added when size moves away from its average (e.g., $va_{i,l} - va_{i,ss} < 0$ when $b_i > 0$). We also find that the industry is not concentrated—despite the large heterogeneity for both $a_i$ and $b_i$, the top 5% of the funds only capture 23% of the industry-wide value added. This lack of concentration results from the strong correlation between $a_i$ and $b_i$. Funds have a limited ability to create value because it is difficult to be skilled along the two dimensions.

The extensive value created by mutual funds has several implications. First, it makes financial prices more efficient which, in turn, improves the allocation of real resources in the economy. Some funds may do so by participating to the primary market, while others may improve the information efficiency of the secondary market, which helps managers make better investment decisions (e.g., Bond, Edmans, and Goldstein (2012)). Because the industry is not concentrated, each fund contributes to this socially valuable function.

Second, it does not contradict the famous arithmetic of Sharpe (1991) which states that the active industry as a whole cannot beat the market. As noted by Pedersen (2018),
this rule breaks down because passive investors only hold a subset of the market and need to trade regularly as companies issue new shares or pay dividends. Therefore, it is not theoretically inconsistent that the whole industry creates value by trading on new information, correcting for mispricing, and accommodating liquidity needs.

Finally, we use our nonparametric approach to test several equilibrium predictions of the BG model. First, we find supportive evidence that the active industry is optimally sized. In the model, each fund chooses the optimal active size to maximize the value added. Consistent with this prediction, the steady state value added represents 76% of the optimal level, on average. Therefore, individual funds seem to internalize the impact of capacity constraints. In contrast, the lifecycle value added remains far from the optimal level. This result may not be surprising because investors do not observe \( a_i \) and \( b_i \)—they must learn about them using past data (Pastor and Stambaugh (2012)). During this learning phase, the money allocated by investors can be quite far from the fund optimal size. Learning effects may also explain why specific funds have more difficulty in maximizing their value added. This is the case for low expense funds for which the skill potential is high on average, but quite volatile. Therefore, investors may struggle to identify the best funds and reward them with additional flows.

Second, the gross alpha, defined as \( \alpha_i = E[\alpha_{i,t}] \), is a noisy measure of skill. Although it is widely used in previous work, we find that its cross-sectional distribution departs significantly from the distribution of each skill measure \( (a_i, b_i \text{ and } va_i) \). As pointed out by the BG model, the gross alpha is generally not informative about skill because it does not control for size. Therefore, some funds can be more skilled than their peers on every dimension (higher \( a_i \) and lower \( b_i \)) and deliver lower gross alphas as long as they manage a larger asset base (larger \( q_i \)).

Third, we examine whether the net alphas received by investors are equal to zero. The BG model predicts that skilled funds are in scarce supply and thus able to extract all the rent, leaving nothing to investors. Our nonparametric approach reveals that the net alphas cluster around zero. However, we observe several departures from the BG model. In particular, a majority of funds charge excessive fees to investors—a point noted by previous studies (e.g., Barras, Scaillet, and Wemers (2010)). Overall, our combined analysis of skill and performance reveals that mutual funds are not only skilled, but also in a stronger bargaining position than initially thought. Our results reveal that a minority of funds are "charlatans" that actually destroy value. Yet, these funds are still able to charge fees to investors.

Our work is related to several strands of the literature. Recent papers use parametric/Bayesian approaches to infer the distribution of fund alphas (e.g., Chen, Cliff, and
Zhao (2017), Jones and Shanken (2005), Harvey and Liu (2018a)) or their sensitivity to capacity constraints (Harvey and Liu (2018b)). Here, we apply a simple, nonparametric approach to multiple skill measures. Several studies apply the False Discovery Rate approach to measure the proportions of funds with non-zero performance (e.g., Avramov, Barras, and Kosowski (2013), Barras, Scaillet, and Wemers (2010), Ferson and Chen (2015)). This paper focuses on skill and estimates the entire distribution (not just the proportions), as well as its moments and quantiles. BG, BvB, and Pastor, Stambaugh, and Taylor (2015) propose the fd alpha and value added as measures of skill. We largely build on their analysis to define our skill measures. Finally, several studies provide evidence of capacity constraints at the aggregate level (e.g., Chen et al. (2004), Pastor, Stambaugh, and Taylor (2015)). Here, we examine the impact of capacity constraints at the individual fund level.

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

II The Measures of Mutual Fund Skill

A The Two Dimensions of Skill

We begin our analysis by presenting the measures of mutual fund skill. Our framework incorporates the widely held view that the fund industry is exposed to capacity constraints. As funds grow large, they may be unable to maintain the same returns because of increasing trading costs (e.g., BG, BvB, Pastor and Stambaugh (2012)). In a world with capacity constraints, individual funds can therefore be skilled along two dimensions. They can be skilled at (i) identifying profitable trades, and/or (ii) minimizing the cost of these trades as fund size increases.

To capture these two skill dimensions, we use the economic model of BG. We denote each fund by the subscript $i = 1, ..., n$, where $n$ denotes the total population size. We take each fund as the unit of analysis and denote it by the subscript $i = 1, ..., n$, where $n$ denotes the total population size. 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

---

2Following the tradition in the mutual fund literature, we take the fund as the unit of observation. Alternatively, we could also apply our approach to the fund managers (e.g., Pastor, Stambaugh, and Taylor (2019), Patel and Sarkissian (2017)).
lagged fund size. The total cost is modeled as a convex function of fund size to capture the impact of capacity constraint, i.e., \( TC_{i,t} = b_i q_{i,t-1}^2 \). Taking the difference between \( TR_{i,t} \) and \( TC_{i,t} \) and dividing by \( q_{i,t-1} \), we obtain the fund gross alpha:

\[
\alpha_{i,t} = a_i - b_i q_{i,t-1},
\]

where \( \alpha_{i,t} \) is time-varying as it dynamically responds to changes in fund size.

Building on Equation (1), we define the two skill dimensions as follows. We capture the first dimension using the first dollar (fd) alpha \( a_i \). This measure isolates the fund skill at identifying profitable ideas by determining the magnitude of the gross alpha when \( q_{i,t-1} = 0 \). In other words, we can interpret \( a_i \) as a "paper" return that is unencumbered by the drag of real world implementation (Perold and Salomon (1991)). We measure the second skill dimension using the size coefficient \( b_i \). This coefficient determines the sensitivity of the gross alpha to changes in fund size. Therefore, a low value of \( b_i \) signals the fund ability to mitigate capacity constraints.

A key feature of our framework is that skill potentially varies across funds, i.e., both measures \( a_i \) and \( b_i \) are fund specific. To capture this heterogeneity, we treat \( a_i \) and \( b_i \) not as fixed parameters, but as random realizations from the cross-sectional skill distributions \( \phi(a) \) and \( \phi(b) \). This contrasts with previous studies which typically impose restrictions on \( a_i \) and \( b_i \). For instance, it is common to assume that the size coefficient is constant across funds (e.g., Chen et al. (2004), Pastor, Stambaugh, and Taylor (2015)). Whereas this pooling assumption reduces estimation errors, it is a priori unclear why capacity constraints have the same impact on all funds. \(^3\)

As noted by BG, Equation (1) does not explicitly model the sources of variation in skill across funds. Skill can potentially vary because some funds have unique investment and trading abilities. For instance, they may receive more precise information signals or trade with high efficiency. Skill can also vary because funds follow specific strategies, such as investing in small cap stocks or trading at high frequencies. Both sources of variations are embedded in the skill measures \( a_i \) and \( b_i \). To see this point, we can write \( a_i = a_i^0 a(s_i) \) and \( b_i = b_i^0 b(s_i) \), where \( a_i^0, b_i^0 \) denote the unique fund skills, and \( a, b \) are functions that depend on a vector \( s_i \) that captures the characteristics of the fund strategy. For instance, Pastor, Stambaugh, and Taylor (2019) specify \( s_i \) as \([liq_i, turn_i]^\prime\), where \( liq_i \) and \( turn_i \) denote the levels of liquidity and turnover chosen by the fund.

Equation (1) is appealing because of its simplicity in capturing the two skill dimen-

\(^3\)To control for heterogeneity across funds, it is common to take the log of \( q_{i,t-1} \) (e.g., Chen et al. (2004), Harvey and Liu (2018b)) based on the assumption that a relative size change has the same impact for all funds. Instead of making this assumption, we allow \( b_i \) to vary across funds.
sions \(a_i\) and \(b_i\). As such, it leaves aside additional predictors that potentially affect the fund gross alpha, such as business cycle indicators, aggregate industry size, and other fund specific variables (age, family size).\(^4\) We can extend our baseline framework to accommodate richer alpha dynamics. To this end, we simply rewrite Equation (1) as

\[
\alpha_{i,t} = a_i - b_i q_{i,t-1} - \beta_i z_{i,t-1},
\]

where \(z_{i,t-1}\) is the \(Q\)-vector of additional predictors. In the empirical section of the paper, we examine several extensions of our baseline specification.

### B The Value Added

Our next measure is the value added which captures the value created by the fund from exploiting its skills. This measure aggregates the two skill dimensions into a single measure, i.e., it determines the economic rent associated with the combination of \(a_i\) and \(b_i\). In addition, the value added rests on a powerful economic interpretation. It is similar to the rent earned by a monopolist defined as the markup price of the good multiplied by the total quantities sold.

We consider two formulations of the value added. The first one proposed by Berk and van Binsbergen (2015) is the value added during the entire fund lifecycle (\(l\)):

\[
\text{va}_{i,l} = \plim_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} \alpha_{i,t} \cdot q_{i,t-1} = E[\alpha_{i,t} q_{i,t-1}] = a_i E[q_{i,t-1}] - b_i E[q_{i,t-1}^2],
\]

(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. In other words, \(\text{va}_{i,l}\) captures the average value added across the different size levels at which the fund operates. The second formulation is the value added once the fund reaches its average, steady state (\(ss\)) size \(E[q_{i,t-1}]\):

\[
\text{va}_{i,ss} = \plim_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} \alpha_{i,t} \cdot \plim_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} q_{i,t-1} = E[\alpha_{i,t}] E[q_{i,t-1}] = a_i E[q_{i,t-1}] - b_i E[q_{i,t-1}^2].
\]

(3)

The difference between the two formulations is equal to the covariance between the gross alpha and size: \(\text{va}_{i,l} - \text{va}_{i,ss} = \text{cov}(\alpha_{i,t}, q_{i,t-1}) = -b_i \text{var}(q_{i,t-1}) = -b_i (E[q_{i,t-1}^2] - E[q_{i,t-1}]^2)\). In the presence of capacity constraints, \(\text{cov}(\alpha_{i,t}, q_{i,t-1})\) is negative and \(\text{va}_{i,ss}\) is larger than \(\text{va}_{i,l}\).\(^5\) Similar to Equation (1), the two formulations of the value added

\(^4\)See, for instance, Chen et al. (2004), Kacperczyk, van Nieuwerburgh, and Veldkamp (2014), and Pastor, Stambaugh, and Taylor (2015, 2017)

\(^5\)This result is due to the Jensen inequality. We can write \(\text{va}_{i,l} = E[\text{va}_i(q_{i,t-1})]\) and \(\text{va}_{i,ss} = \text{va}_i(E[q_{i,t-1}])\), where \(\text{va}_i(q_{i,t-1}) = (a_i - b_i q_{i,t-1})q_{i,t-1}\). Because \(\text{va}_i(q_{i,t-1})\) is a concave function, we
are fund specific. To capture this heterogeneity, we treat \( va_{i,t} \) and \( va_{i,ss} \) as random realizations from the cross-sectional distributions \( \phi(va_l) \) and \( \phi(va_{ss}) \).

Previous studies commonly use the average gross alpha \( \alpha_i = E[\alpha_{i,t}] \) as a measure of aggregate skill.\(^6\) However, the gross alpha is likely to be a noisy measure of the value added because it does not control for fund size (BvB). Intuitively, 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. To illustrate, suppose that fund A is more skilled than fund B on every dimension (\( a_A > a_B, b_A < b_B \)), but chooses the same level of fees \( f_e \). If investors compete for performance such that the equality \( \alpha_i = f_e \) holds, both funds produce the same gross alphas (\( \alpha_A = \alpha_B \)). In this case, a comparison based on the gross alpha fails to capture any difference in skill.

III Overview of the nonparametric Approach

A General Motivation

We now describe the approach for estimating the cross-sectional skill distribution \( \phi(m) \), where \( m \in \{a, b, va_l, va_{ss}\} \) encompasses all four measures presented above. Our methodological contribution is to develop a nonparametric approach that imposes minimal structure on the skill distribution. As a result, it provides several key advantages.

First, our approach is largely immune to misspecification errors. This is not the case for standard Bayesian/parametric approaches as they require to fully specify the shape of the true distribution. In the context of skill, choosing the correct specification is challenging—whereas theory predicts that performance (net alphas) should cluster around zero, it offers no such guidance for skill. In principle, we can gain parametric flexibility by using normal mixture models (e.g., Chen, Cliff, and Zhao (2017), Harvey and Liu (2018a)). In practice, however, determining the correct number of mixtures is difficult because the parameters are estimated with significant noise (Cheng and Yang (2019)), and the statistical inference is technically involved (Chen (2017)).\(^7\)

Second, it allows for a joint analysis of all four skill measures. Such analysis is extremely challenging with Bayesian/parametric approaches because they require to correctly specify and estimating a multivariate distribution whose

\[
E[va_{i,t-1}] < va_i(E[q_{i,t-1}]).
\]

\(^6\)The gross alpha is examined, among others, by Baks, Metrick, and Wachter (2001), Barras, Scaillet, and Wermers (2010), Jensen (1968), Jones and Shanken (2005), Wermers (2000).

\(^7\)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)). In our setting, inference is even more complicated because we do not observe the true skill measures, but only the estimated ones.
marginals are potentially mixtures of distributions. To sidestep this challenge, it is tempting to specify and estimate each skill distribution separately. However, this procedure is likely to generate inconsistencies because the skill measures are theoretically related as per Equations (1)-(3).

Third, the implementation of the nonparametric approach is simple and fast. Intuitively, it is akin to computing an histogram using as inputs the estimated skill measure of each fund. In contrast, Bayesian/parametric approaches require sophisticated and computer-intensive Gibbs sampling and Expectation Maximization (EM) methods (e.g., Chen, Cliff, and Zhao (2017), Harvey and Liu (2018a), Jones and Shanken (2005)).

Fourth, it provides a unified framework for estimating the skill density function $\phi$, along with the other characterizations of the distribution, including the moments (e.g., mean, variance), the cumulative function $\Phi(x) = \text{prob}[m_i \leq x] = \int_{-\infty}^{x} \phi(m)dm$, and the distribution quantile $q(p) = \Phi^{-1}(p)$, where $p$ denotes the probability level.\footnote{To lighten notation, we do not subscript the density $\phi$ and the other quantities by the skill measure.}

Last but not least, it comes with a full-fledged econometric 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 determine its asymptotic properties and conduct proper statistical inference guided by theoretical results.

In the remaining sections, we present our nonparametric approach in more detail. For sake of brevity, we focus on the estimation of the skill density and relegate to the appendix the formal treatment of the three remaining estimators (moments, cdf, quantiles), as well as the proofs of the different econometric results.

B Estimation Procedure

B.1 Estimation of the Skill Measures

Our nonparametric estimation of the skill density $\phi(m)$ consists of three main steps. To begin, we estimate the skill measure of each fund $i$ in the population ($i = 1, \ldots, n$) using the following time-series regression:

$$r_{i,t} = \alpha_{i,t} + \beta_i^T f_t + \varepsilon_{i,t} = a_i - b_i g_{i,t-1} + \beta_i^T f_t + \varepsilon_{i,t},$$

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 $\varepsilon_{i,t}$ 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
\( \beta_i \) are random realizations from a continuum of funds. Under this sampling scheme, we can invoke cross-sectional limits to infer the density of each skill measure \( m.9^{10} \)

The vector of coefficients \( \hat{\gamma}_i = (\hat{a}_i, \hat{b}_i, \hat{\beta}_i)' \) for fund \( i (i = 1, ..., n) \) is computed as

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

where \( I_{i,t} \) is an indicator variable equal to one if \( r_{i,t} \) is observable (and zero otherwise), \( T \) is the total number of periods, \( T_i = \sum_{t=1}^{T} I_{i,t} \) 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,

\[
\tilde{a}_{1,i} = \frac{1}{T_i} \sum_{t=1}^{T} I_{i,t} q_{i,t-1}, \quad \tilde{a}_{2,i} = \frac{1}{T_i} \sum_{t=1}^{T} I_{i,t} q_{i,t-1}^2,
\]

we can then compute each of the four skill measures as

Fd alpha : \( \hat{m}_i = \hat{a}_i \),

Size coefficient : \( \hat{m}_i = \hat{b}_i \),

Value added (lifecycle) : \( \hat{m}_i = \hat{\alpha}_{i,l} = \hat{a}_i \tilde{a}_{1,i} - \hat{b}_i \tilde{a}_{2,i} \),

Value added (steady state) : \( \hat{m}_i = \hat{\alpha}_{i_ss} = \hat{a}_i \tilde{a}_{1,i} - \hat{b}_i \tilde{a}_{2,i} \).

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 \( 1_i^x \) equal to one if the following two conditions are met (and zero otherwise):

\[
1_i^x = 1 \{ CN_i \leq \chi_{1,T}, \tau_{i,T} \leq \chi_{2,T} \},
\]

where \( CN_i = \sqrt{\frac{eig_{max}(\hat{Q}_{x,i}^2)}{eig_{min}(\hat{Q}_{x,i})}} \) 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

\( ^9 \text{Gagliardini, Ossola, and Scaillet (2016) use a similar sampling scheme to develop testable applications of the arbitrage pricing theory in a large cross-section of assets.} \)

\( ^{10} \text{We can also apply our approach to estimate the cross-sectional distribution of the fund beta for each risk factor } k (k = 1, ..., K), \text{ denoted by } \phi(\beta_k). \text{ As explained below, the common practice of estimating } \phi(\beta_k) \text{ using the estimated betas is biased because of the error-in-variable (EIV) problem.} \)
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, the fund coefficients are estimated 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 \( n^\chi = \sum_{i=1}^{n} 1^X_i \).

**B.2 Kernel Density Estimation**

In the next step, we estimate the skill density function using a standard nonparametric approach based on kernel smoothing. The estimated density \( \hat{\phi} \) at a given point \( m \) is computed as

\[
\hat{\phi}(m) = \frac{1}{n^\chi h} \sum_{i=1}^{n} 1^X_i K\left( \frac{m_i - m}{h} \right),
\]

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 \( K(u) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{u^2}{2}\right) \) for our empirical analysis (see Silverman (1986)).

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 \( \sqrt{nh}(h^2T + (1/T)^3) \to 0 \), we have

\[
\sqrt{nh} \left( \hat{\phi}(m) - \phi(m) - bs(m) \right) \Rightarrow N(0, K_1 \phi(m)),
\]

and the bias term \( bs(m) \) is the sum of two components,

\[
bs_1(m) = \frac{1}{2} h^2 K_2 \phi^{(2)}(m),
\]

\[
bs_2(m) = \frac{1}{2T} \psi^{(2)}(m),
\]


\(12\) Similar to Equation (8), Okui and Yanagi (2018) consider a kernel estimator for the density of the mean and autocorrelation of random variables. However, their distributional results differ from those derived in our regression context aimed at measuring fund skill.
where $K_1 = \int K(u)^2du$, $K_2 = \int u^2K(u)du$, $\phi^{(2)}(m)$ 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 centered skill measure $\sqrt{T}(\hat{m}_i - m_i)$ equal to $\lim_{T \to \infty} \frac{T^2}{T} \sum_{t,s=1}^T I_{i,t} I_{i,s} u_{i,t} u_{i,s}$. For each skill measure, the term $u_{i,t}$ is given by

\[
\begin{align*}
\text{Fd alpha} & : u_{i,t} = e_1 Q_{x,i}^{-1} x_{i,t} \varepsilon_{i,t}, \\
\text{Size coefficient} & : u_{i,t} = e_2 Q_{x,i}^{-1} x_{i,t} \varepsilon_{i,t}, \\
\text{Value added (lifecycle)} & : u_{i,t} = E[q_{i,t-1}] e_1 Q_{x,i}^{-1} x_{i,t} \varepsilon_{i,t} + a_i(q_{i,t-1} - E[q_{i,t-1}]) \\
& \quad - E[q_{i,t-1}] e_2 Q_{x,i}^{-1} x_{i,t} \varepsilon_{i,t} - b_i(q_{i,t-1} - E[q_{i,t-1}]), \\
\text{Value added (steady state)} & : u_{i,t} = E[q_{i,t-1}] e_1 Q_{x,i}^{-1} x_{i,t} \varepsilon_{i,t} + a_i(q_{i,t-1} - E[q_{i,t-1}]) \\
& \quad - E[q_{i,t-1}] e_2 Q_{x,i}^{-1} x_{i,t} \varepsilon_{i,t} - b_i(2E[q_{i,t-1}] - E[q_{i,t-1}]), (12)
\end{align*}
\]

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. □

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 density estimation by adjusting for the bias term $bs(m)$. Equations (10)-(11) reveal that $bs(m)$ has two distinct components. The first component $bs_1$ is the smoothing bias, which is standard in nonparametric density estimation (e.g., Silverman (1986), Wand and Jones (1995)). The second component $bs_2$, which is referred to as the error-in-variable (EIV) bias, is non-standard—it arises because we estimate $\phi$ using the estimated skill 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 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 $(nT)^{-1/5}$; (ii) if $T$ is of comparable size or large relative to $n$ ($n^{2/5}/T \to 0$), $h^*$
is proportional to $n^{-\frac{1}{2}}$. Our Monte-Carlo analysis presented in the appendix reveals 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:

$$h^* = \left( \frac{K_2}{K_1} \int \phi^{(2)}(m)\psi^{(2)}(m)dm \right)^{-\frac{1}{4}} (n/T)^{-\frac{1}{4}}. \quad (13)$$

### B.3 Bias Adjustment

Our final step is to adjust the kernel density estimator $\hat{\phi}(m)$ for the bias. Building on the insights of Proposition III.1, we compute the two bias terms and the optimal bandwidth using a Gaussian reference model in which the fund skill measure $m_i$ and the log of the asymptotic variance $s_i = \log(S_i)$ are drawn from a bivariate normal distribution: $m_i \sim N(\mu_m, \sigma_m^2)$, $s_i \sim N(\mu_s, \sigma_s^2)$, and $corr(m_i, s_i) = \rho$.\(^{14}\)

This simple reference model has several appealing properties. First, the computation of the bias and the bandwidth is straightforward. 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, they both have closed-form expressions. This allows us to examine the determinants of the bias, and 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)).\(^{15}\) Similarly, the standard bootstrap usually seriously 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 closed-form expressions for the two bias components and the optimal bandwidth under the reference model as the number of funds $n$

---

\(^{13}\)The AMISE is defined as the integrated sum of the leading terms of the asymptotic variance and squared bias of the estimated density $\hat{\phi}(m)$.

\(^{14}\)A normal reference model underlies the celebrated Silverman rule of thumb for the choice of the bandwidth in standard non-parametric density estimation without the EIV problem. This rule gives $h^* = 1.06\sigma n^{-\frac{1}{2}}$, where $\sigma$ is the standard deviation of the observations (Silverman (1986)).

\(^{15}\)We can estimate the $r$th-derivative of a density $\phi$ 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 order $r$. 

---
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 $\sqrt{n h (h^2 T + (1/T)^{1/4})} \to 0$, the two bias components under the reference model are equal to

$$bs_1^\nu(m) = \left[ \frac{1}{2} K_2 h^2 \frac{1}{\sigma^2_m} (\bar{m}_1^2 - 1) \right] \frac{1}{\sigma^2_m} \varphi(\bar{m}_1),$$

$$bs_2^\nu(m) = \left[ \frac{1}{2T} \exp(\bar{\mu}_s) \frac{1}{\sigma^2_m} (\bar{m}_2^2 - 1) \right] \frac{1}{\sigma^2_m} \varphi(\bar{m}_2),$$

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

$$h^* = \left[ \frac{K_2}{K_1 2\sqrt{\pi}} \frac{3}{4 \sigma^2_m} \left( \frac{\rho^4 \sigma^4_s}{12} - \rho^2 \sigma^2_s + 1 \right) \exp\left( \bar{\mu}_s (1 - \frac{\rho^2}{2}) \right) \right]^{-\frac{1}{2}} (n/T)^{-\frac{1}{4}}. \tag{16}$$

**Proof.** See the appendix.

Proposition III.2 yields several insights. First, 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 $bs_1^\nu(m)$. With a population of several thousand funds, the smoothing term becomes negligible for all values of $m$. In contrast, $bs_2^\nu(m)$ 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. Second, the magnitude of the EIV bias depends on the variances of the true versus estimated skill, measured as $\sigma^2_m$ and $\sigma^2_m = \frac{1}{T} \exp(\bar{\mu}_s)$. As $\sigma_m$ increases relative to $\sigma_m$, the EIV bias becomes less severe because it makes the cross-sectional variation of the estimated skill more aligned with that of the true skill (and vice-versa). The appendix contains a detailed comparative static analysis of the EIV bias for our skill measures.

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

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

where $\hat{u}_{i,t}$ is obtained by plugging the estimated quantities for the chosen skill 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 \( \hat{b}_{s_1}(m) \), \( \hat{b}_{s_2}(m) \), and the optimal bandwidth \( h^* \). Finally, we remove the bias terms from the unadjusted density in Equation (8) to obtain the bias-adjusted density estimator

\[
\hat{\phi}^*(m) = \hat{\phi}(m) - \hat{b}_{s_1}(m) - \hat{b}_{s_2}(m).
\]

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 \( bs_2(m) \approx bs_2(m) \)). 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 skill density \( \phi \). As long as \( \phi \) peaks around its mean, this derivative takes negative values in the center and positive values in the tails—exactly like the function \( bs_2(m) \).\(^{16}\) Second, our extensive Monte-Carlo analysis calibrated on the data reveal that the bias-adjusted density captures the true density remarkably well (see the appendix).\(^{17}\)

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 gross return and size time-series for the entire population and different groups of funds sorted on investment style (small/large cap and growth/value), and characteristics (low/high expense ratios and low/high turnover).

To estimate the regression for each fund in Equation (4), we use the four-factor model of Cremers, Petajisto, and Zitzewitz (2012; CPZ hereafter) 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. The CPZ model depart from the model of Carhart (1997) in two respects: (i) \( r_{m,t} \) is proxied by the excess return of the S&P500 (instead of the CRSP market index), and (ii) the size and

\(^{16}\) The two terms \( bs_2(m) \) and \( bs_2(m) \) only differs if \( \phi \) 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.

\(^{17}\) 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.
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.\textsuperscript{18}

The motivation for using the CPZ 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, CPZ show that the Russell 2000 produces a Carhart alpha of -2.4% per year over the period 1980-2005. Therefore, if a fund uses the Russell 2000 as a benchmark, it is likely 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 2,291 funds. The appendix provides more detail on the construction of the mutual fund dataset.

## B Summary Statistics

Table I reports summary statistics for our mutual fund sample. For the entire population and each fund group, we construct a value weighted portfolio of all funds and report the first four moments of its gross excess returns in Panel A. In the entire population, the portfolio achieves a risk-return tradeoff similar to that of the aggregate stock market with a mean and volatility equal to 8.0% and 14.8% per year. It also exhibits a negative skewness (-0.7) and a positive kurtosis (5.3). The results are similar across groups, except for the small cap portfolio which produces higher levels for the mean and volatility.

In Panel B, we report the estimated portfolio betas on the four factors in the CPZ model. Consistent with intuition, small cap funds are heavily exposed to the size factor (0.78). Whereas growth funds are negatively exposed to the value factor (-0.35), the opposite holds for value funds (0.22). We also find that high expense and high turnover funds tilt toward small cap, growth stocks.

Please insert Table I here

\textsuperscript{18} Because the factors in the CPZ model are not available between January 1975 and December 1978, we replace them with the values obtained from the Carhart model. Focusing instead on the period January 1979-December 2018 does not change our main results.
V Empirical Results

A The Two Dimensions of Skill

A.1 Analysis in the Entire Population

We begin our analysis by examining the two skill dimensions—the fd alpha and size coefficient. We estimate \( a_t \) and \( b_t \) for each fund via ordinary least squares, and then use our nonparametric approach to infer the cross-sectional skill distributions \( \phi(a) \) and \( \phi(b) \). To describe their properties, we compute the bias-adjusted estimates of (i) the moments (mean, variance, skewness, kurtosis), (ii) the proportions of funds with negative and positive skill measures denoted by \( \hat{\pi}^- \) and \( \hat{\pi}^+ \), and (iii) the distribution quantiles at 5% and 95% denoted by \( \hat{q}^5 \) and \( \hat{q}^{95} \) (see the appendix for the computations). To ease interpretation, we standardize \( \hat{b}_t \) for each fund so that it corresponds to the change in gross alpha for a one standard deviation change in size. The summary statistics for \( \phi(a) \) and \( \phi(b) \) are shown in Panels A and B of Table II.

The results for the first skill dimension reveals two key insights. First, there is overwhelming evidence that individual funds produce profitable trading ideas. The fd alpha is positive for 86% of the funds in the population, and economically large with an average level of 3.1% per year. Second, we see a large heterogeneity in skill across funds. In particular, a minority of funds exhibit stellar investment skills—5% of them exhibit a fd alpha above 8.2% per year, which is 2.6 times larger than the average.

The cross-sectional distribution of the second skill dimension exhibits similar features. More than 85% of the funds have a positive size coefficient whose magnitude is typically large. On average, a one standard deviation increase in size reduces the gross alpha by 1.4% per year. These results provide strong support to models that emphasize the importance of capacity constraints for mutual funds (e.g., BG, Pastor and Strambaugh (2012)). We also observe a strong heterogeneity as funds largely differ in their ability to mitigate capacity constraints—a finding that is inconsistent with the common assumption that the size coefficient remains constant across funds.

Please insert Table II here

A.2 Comparison between Fund Groups

The strong cross-sectional variation in skill can potentially be driven by the specific strategy followed by individual funds. The results shown in Table II for different fund groups confirm that it is indeed the case. The average value of the fd alpha varies
between 1.8% and 4.6% per year across the different groups. Similarly, the average value of the size coefficient ranges between 0.8% and 1.7%. However, there is still a substantial variation in skill left unexplained by the fund strategy, i.e., the volatility of $a_i$ and $b_i$ is similar to its population level. This implies that each group contains funds with unique abilities to identify highly profitable ideas and trade more efficiently.

Figure 1 further shows that the two size groups exhibit the biggest skill gap. Small cap funds largely dominate large cap funds along the first skill dimension (FD alpha). There are several interpretations for this result. Small cap stocks potentially exhibit greater mispricing because of limits to arbitrage—as noted by Hong, Lim, and Stein (2000), these stocks are largely untouched by mutual funds. In addition, small cap stocks have a higher idiosyncratic volatility and may therefore provide more opportunities for stock picking (e.g., Duan, Yu, and McLean (2009)). On the contrary, small cap funds are largely dominated along the second skill dimension (size coefficient). This finding resonates with previous studies that show that, on average, small cap funds face tighter capacity constraints (e.g., Chen et al. (2004), Pastor, Stambaugh, and Taylor (2015)).

We document a similar pattern for funds sorted on characteristics. Specifically, funds with high expense ratios/turnover tend to generate more profitable ideas, but follow strategies that are difficult to scale up. This result is partly driven by the size effect documented above because high expense/turnover funds tilt their portfolios toward small cap stocks (see Table I).

Please insert Figure 1 here

A.3 The Correlation between the Two Skill Dimensions

A key insight of our cross-sectional analysis is that the two skill dimensions are strongly correlated. In the entire population, a simple (biased) correlation calculation based on the estimated coefficients $\hat{a}_i$ and $\hat{b}_i$ yields a coefficient of 0.82. Across fund groups, the correlation between the average values of $\hat{a}_i$ and $\hat{b}_i$ is even larger at 0.90.

This positive correlation is partly explained by the fund investment strategy. As discussed in Section II, a specific strategy determines the fund characteristics $s_i$ (e.g., liquidity, turnover) which, in turn, jointly affect the two skill measures (Pastor, Stambaugh, and Taylor (2019)). For instance, investing in small cap stocks involves a higher degree of illiquidity. This characteristic increases trading costs and makes mispricing more difficult to arbitrage away. As a result, small cap funds exhibit higher levels of $a_i$ and $b_i$ than large cap funds. Similarly, $a_i$ and $b_i$ are typically larger among high turnover funds as they choose to exploit more ideas at the cost of trading more often.
The implications of this positive correlation are twofold. First, it becomes essential to aggregate the two skill dimensions to assess the overall skill level. Funds with the most profitable investment ideas are also likely to be the ones with the most binding capacity constraints. Therefore, it is a priori unclear whether they dominate funds that are able to scale up less profitable trading ideas.

Second, it clarifies the relation between skill and characteristics such as expense ratios and turnover. Whereas previous studies argue that these characteristics signal superior skill (e.g., Kacperczyk, van Nieuwerburgh, and Veldkamp (2014), Pastor, Stambaugh, and Taylor (2017)), other studies favor the opposite interpretation (e.g., Elton et al. (1993), Gil-Bazo and Ruiz-Verdu (2009)). Our results show that both conclusions hold provided that the right dimension of skill is examined—high expense/turnover funds are skilled at generating profitable ideas, but unskilled at mitigating capacity constraints.

B The Value Added

B.1 Analysis of the Entire Population

We now turn to the analysis of the lifecycle and steady state value added. We estimate $v_{ai,l}$ and $v_{ai,ss}$ for each fund, and then use our nonparametric approach to infer the cross-sectional distributions $\phi(v_{ai})$ and $\phi(v_{a_{ss}})$. As shown in Equations (2)-(3), both $v_{ai,l}$ and $v_{ai,ss}$ depend on the two skill dimensions and thus provide an economically-motivated approach for aggregating them. The summary statistics for $\phi(v_{ai})$ and $\phi(v_{a_{ss}})$ are reported in Panels A and B of Table IV.

Individual funds create significant value from their investment and trading decisions. On average, the annual value added is equal to 1.7 mio. over the fund lifecycle, and 7.3 mio. once the fund reaches its average size. In addition, the large majority of funds create value—the proportions $\hat{\pi}_{va_{i}}^+$ and $\hat{\pi}_{va_{ss}}^+$ are close to 60% and 70%, respectively. Consistent with our previous analysis, $v_{ai,l}$ is typically lower than $v_{ai,ss}$ because funds are exposed to capacity constraints (i.e., $\text{cov}(\alpha_{it}, q_{i,t-1}) < 0$). Put differently, the value added is significantly lower than its steady state level when $q_{i,t-1}$ is below average, but only marginally higher when $q_{i,t-1}$ is above average.

A minority of funds produce a negative value added or, equivalently, a negative gross alpha. These funds are either "charlatans", i.e., funds without any investment ideas ($a_i < 0$), or funds that grow too large to maintain a positive alpha ($a_i < b_i q_{it-1}$).

19The average for the lifecycle value added is similar to the number reported by BvB, which is equal to 2.0 mio. per year (Table 7). Note that their baseline number of 3.2 mio. per year cannot be compared to ours because it includes the fund diversification services (i.e., it is based on a comparison between the fund gross return and the net index returns).
To distinguish between these two cases, we compare the proportions $\hat{\pi}_a$ and $\hat{\pi}_{va} = (\hat{\pi}_{a1} + \hat{\pi}_{a2})/2$. We find that 14% of the funds in the population are charlatans ($\hat{\pi}_a = 14.0\%$), which represents around 40% of the funds with negative value added ($\hat{\pi}_{va} = 14.0\% / (41.9+31.2)/2 = 40.0\%$). It is a priori surprising that the value added is negative given that funds always have the option to invest passively and earn a zero alpha. It could be the case that charlatans take active positions to hide their lack of skill and charge higher fees (Berk and van Binsbergen (2018)). Another possibility is that some funds may grow too large as investors and managers learn about the skill measures $a_i$ and $b_i$—a point we revisit below.

Another insight from Table IV is that the mutual fund industry is not heavily concentrated—at the steady state, the top 5% of the funds only capture 20.8% of it. This departs from the prediction that concentration should increase as technology improves the ability of investors to detect skilled funds (Garleanu and Pedersen (2018)). This lack of concentration is a natural consequence of the correlation between $a_i$ and $b_i$. Because it is difficult for individual funds to be skilled along the two dimensions, they generally have a limited potential to create value. A simple calculation confirms this point—if we assume that $a_i$ and $b_i$ were uncorrelated, we would observe a high degree of concentration as the top 5% would capture 92.7% of the value added.21

Please insert Table III here

B.2 Comparison between Fund Groups

Focusing on the value added allows us to determine which types of funds achieve the most profitable combinations of investment and trading skills. As shown in Table III and Figure 2, the small cap group unambiguously creates more value (lifecycle and steady state) than the large cap group. For one, it produces a higher average (5.7 vs 2.0 mio.) and a higher skill proportion ($\hat{\pi}_{va} = 72.8\%$ vs 53.0%). Taken together, these results imply that the skill of small cap funds at identifying profitable trades more than compensates for their greater exposure to capacity constraints. We also find that the low turnover group largely dominates the high turnover group. Contrary to small cap

---

20We denote the total value added for the population and the top 5% as $\hat{V} = n\hat{\mu}_{va,ss}$ and $\hat{V}(5) = n \cdot 0.05\hat{E}(va_{i,ss} | va_{i,ss} > q_{va,ss})$, where $\hat{\mu}_{va,ss}$ is the average value added. We compute the expectation term via a numerical integration of $\hat{\phi}(va_{ss})$ to obtain $\hat{V}(5)/\hat{V} = 22.8\%$.

21For each fund $i$, we simulate 10,000 values of $a_i$ and $b_i$ by drawing them independently from the vectors of estimated positive $a$ alphas and size coefficients. We then compute the value added by assuming that funds choose their level optimally such that $va_i = a^2_i/4b_i$ (as per Equation (19) below). Finally, we compute the ratio $\hat{V}(5)/\hat{V}$ to obtain 92.7%.
funds, low turnover funds create value through another skill dimension, i.e., their ability to scale up their investment strategy.

The comparison is more subtle for funds with different expense ratios. On average, the low expense group creates significantly more value than the high expense group (8.9 vs 3.6 mio.). However, its value added distribution is more spread out, which implies a higher proportion of funds that destroy value ($\tilde{\pi}_{va} = 39.0\%$ vs 32.0%). This proportion difference cannot be explained by a higher concentration of charlatans—$\tilde{\pi}_{va}$ is similar in both groups (Table II). Instead, it suggests a capital misallocation as some low expense funds grow too large to maintain a positive gross alpha.

C Impact of the EIV Bias

The EIV bias adjustment largely changes the shape of the unadjusted distribution $\hat{\phi}(m)$. First, it removes probability mass from the tails. Intuitively, using the estimated skill measures introduces noise and thus inflates the probability of observing extreme skill levels. Second, it induces a positive skewness because the correlation $\rho$ between each skill measure and estimation variance is positive ($\hat{\rho} \approx 0.25$). In other words, funds with better investment ideas (large $a_i$), tighter capacity constraints (large $b_i$), and larger value added (large $va_i$) tend to hold concentrated portfolios with higher volatility.

To quantify these adjustments, Table IV compares the bias adjusted and unadjusted distributions for all skill measures. Apart from the mean which is not subject to the EIV bias, the differences are striking. The unadjusted quantiles are implausibly large because they are heavily influenced by large observations. For one, the spread between the two quantiles for the lifecycle value added is 2.3 times larger than the adjusted spread. In addition, the unadjusted distribution fails to capture the strong asymmetry in skill across individual funds.

Put together, these results change the economic interpretation of the results. For instance, the unadjusted statistics for the lifecycle value added lead to the wrong conclusion that the majority of funds destroy value ($\tilde{\pi}_{va,l} = 54.2\%$). In addition, they reveal that 22.5% of the funds have a negative size coefficient—a suspiciously large number given that equity funds typically do not trade in OTC markets where transaction costs decrease with size (Pedersen (2015; ch. 5)). In contrast, the adjusted statistics capture the asymmetric nature of capacity constraints—whereas $b_i$ is close to zero for unconstrained funds ($q_{b_i} = -0.6\%$), it can rise significantly for funds facing tight capacity

---

22 Formally, Equation (15) shows that when $\rho$ is positive, the bias adjustment induces positive skewness because the probability mass is not transferred around the mean, but to its right ($\mu_m + \rho \sigma_m \sigma_s > \mu_m$).
D Equilibrium Considerations

D.1 Do Skilled Funds Maximize the Value Added?

We now study the equilibrium implications of the BG model for mutual fund skill. This model considers the interaction between a set of skilled funds in scarce supply and a large number of rational investors that compete for performance. Solving for the equilibrium yields several intuitive predictions about skill that we can examine empirically using our nonparametric approach.

The first one is that individual funds choose a size $q_i^*$ at which the value added is maximized. In the BG model, each fund has investment ideas ($a_i > 0$), but a limited ability to scale up its strategy ($b_i > 0$). Its objective is to maximize profits $v$ under the constraint that the fees $f_{e,i}$ are equal to the gross alpha $a_i$ (investors break even). Therefore, maximizing profits is equivalent to maximizing the value added, i.e., $v = f_{e,i} q_i = a_i = v a_i$. Replacing $a_i$ with $a_i - b_i q_i$ and using the first order condition $\frac{\partial v a_i}{\partial q_i} = 0$, we have $q_i^* = \frac{a_i}{2b_i}$, and the optimal value added $v a_i^*$ is given by $a_i q_i^* - b_i q_i^2 = \frac{a_i^2}{4b_i}$.

We begin our analysis by estimating the cross-sectional distribution of the optimal value added $\mu(v a^*)$. To apply our nonparametric approach, we build on our proposition III.1 and simply write the estimated skill measure $\hat{\mu}_i$ and the error term $u_{i,t}$ as

$$\hat{\mu}_i = \hat{v} a_i^* = \frac{\hat{a}_i^2}{4\hat{b}_i}, \quad u_{i,t} = \frac{2a_i}{4\hat{b}_i} e'_i Q_{x,i}^{-1} x_{i,t} \varepsilon_{i,t} - \frac{a_i^2}{4\hat{b}_i} e'_i Q_{x,i}^{-1} x_{i,t} \varepsilon_{i,t},$$

where we impose the restriction that $\hat{a}_i$ and $\hat{b}_i$ are positive (which holds for 72% of the funds). This condition guarantees that the optimal value added is well defined (i.e., $\hat{v} a_i^*$ is positive). The summary statistics for $\phi(v a^*)$ are shown in Panel A of Table V.

The maximum value that funds could potentially create is large—for the average fund, it reaches 14.1 mio. per year. It also varies significantly, both in the population and across fund groups. In the population, 5% of the funds create more than 50 mio. per year. Across fund groups, we find that low expense/turnover funds have the highest skill potential with an average of 21.0 and 19.3 mio. per year. Overall, these results further emphasize the importance of accounting for the heterogeneity across funds.

Next, we test the prediction of the BG model by measuring the difference between
the optimal value added $va_i^*$ and its actual level measured either with $va_{i,t}$ or $va_{i,ss}$. We then estimate the distributions $\phi(va^* - va_i)$ and $\phi(va^* - va_{ss})$ nonparametrically with $\hat{m}_i$ equal to $\hat{va}^*_i - \hat{va}_{i,t}$ ($\hat{va}^*_i - \hat{va}_{i,ss}$), and $u_{i,t}$ equal to the difference between the error terms in Equations (20) and (12).

Consistent with the model prediction, Panel B of Table V shows that the steady state value added is close to the optimal level. The average difference between $va^*_i$ and $va_{i,ss}$ is equal to 3.4 mio. per year, which means that funds extract more than 75% of the optimal profits once they reach their average size. The strong pairwise correlation of 0.94 between $\hat{va}_{i,ss}$ and $\hat{va}^*_i$ confirms that funds with higher skill potential do create more value. In contrast, the lifecycle value added is far from the optimal level—the average difference between $va^*_i$ and $va_{i,t}$ reaches 13.5 mio. per year. An intuitive explanation for this large gap is the presence of learning effects. Because investors do not observe the skill dimensions $a_i$ and $b_i$, they must learn about them using past data (Pastor and Stambaugh (2012)). Therefore, the amount of money they are willing to invest can be quite different from the optimal size $q^*_i$ at which the value added is maximized. Whereas the impact of learning decreases as the fund reaches its steady state level, it is likely to have a large impact during the fund lifecycle.

Learning effects may also explain why some groups of funds leave more money on the table. For instance, there is significant uncertainty about skill among low expense funds—this group exhibits the highest cross-sectional volatility of $va^*_i$ (35.1 mio. per year). It may therefore be more difficult for investors to detect the most skilled funds and reward them with additional flows. Consistent with this analysis, low expense funds exhibit the largest gap between the optimal and actual value added.

Please insert Table V here

D.2 Is the Gross Alpha a Noisy Measure of Skill?

The second prediction of the BG model is that the gross alpha $\alpha_i$ is a noisy measure of skill. In the model, funds can choose any level of fees $f_{e,i}$ without changing the optimal value $va_i^*$. If, for instance, a fund chooses low fees, it receives additional money from investors, $q_i(f_{e,i}) - q_i^*$, which can be passively invested to keep $va_i^*$ unchanged. Therefore, arbitrary fees imply arbitrary gross alpha (given the equality $f_{e,i} = \alpha_i$).

The gross alpha is only informative about skill in particular cases where all funds coordinate on the same fee setting policy. In the appendix, we show that the gross alpha is proportional to (i) the $fd$ alpha only if all funds set $q_i(f_{e,i}) = q_i^*$ (optimal size), (ii) the size coefficient only if all funds set $q_i(f_{e,i}) = q_i^{*2}$ (squared optimal size), and (iii) the
value added only if all funds set \( q_i(f_{e,i}) = \bar{q} \) (same size).\(^{23}\) Each of these policies implies specific predictions regarding fees and size that are summarized in Panel A of Table VI. For instance, the policy same size predicts a constant size across funds.

Overall, we find limited evidence that the gross alpha is informative about skill. First, Panel B shows that the statistics on fund fees and size are not consistent with any of the policies mentioned above. Instead, we find that fund size tends to be large and negatively related to fees (the pairwise correlation equals -0.23 across funds). This is consistent with the analysis of Habib and Johnson (2016) who note that some funds charge low fees and manage a large asset base to mitigate several institutional constraints.\(^{24}\)

Please insert Table VI here

Second, the cross-sectional distribution of the gross alpha \( \phi(\alpha) \) departs significantly from the skill distributions. To estimate \( \phi(\alpha) \), we compute the gross alpha of each fund from the time-series regression \( r_{i,t} = \alpha_i + \beta_{i,t} f_t + \varepsilon_{i,t} \), and apply our nonparametric approach with

\[
\hat{m}_i = \hat{\alpha}_i, \quad \hat{u}_{i,t} = \varepsilon_{i,t} Q_x^{-1} x_t, \tag{21}
\]

where \( x_t = (1, f_t', \bar{f}_t') \) and \( Q_x = E[x_t x_t'] \). As shown in Table VII, the variation in the gross alpha, both within and across groups, bears little resemblance with the one reported for the skill measures \( (\alpha_i, b_i, va_{i,t}, va_{i,ss}) \). Consistent with this result, the average correlation between the estimated gross alpha and each skill measure only equals 0.36.

Please insert Table VII here

D.3 Do Funds Extract All the Rent from their Skill?

The third prediction of the BG model is that the net alpha \( \alpha_i^* \) earned by investors is equal to zero. In the model, skilled funds are in scarce supply, which allows them extract all the rent by setting fees equal to the gross alpha \( (\alpha_i = f_{e,i}) \). This behaviour implies that investors are left with a zero surplus (once the learning process is completed).

\(^{23}\)Our analysis largely builds on that of BvB which already discusses the relation between the gross alpha and the value added. Our contribution is to include all four skill measures (fd alpha, size coefficient, value added, gross alpha) and examine their relations under different fee setting policies.

\(^{24}\)The Investment Company Act imposes diversification rules on 75% of the portfolio 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.
To examine this issue, we apply our nonparametric approach to infer the cross-sectional distribution of the net alpha $\phi(\alpha^n)$. We compute the net alpha of each fund using the time-series regression $r_{i,t}^n = \alpha_{i,t}^n + \beta_{i,t}^f r_f + \varepsilon_{i,t}$, where the net excess return $r_{i,t}^n$ is equal to the gross excess return minus the fees. Then, we write $\hat{m}_i$ and $u_{i,t}$ as

$$\hat{m}_i = \hat{\alpha}_i^n,$$  \hspace{1cm} \hspace{1cm} (23)

$$u_{i,t} = e'_i Q^{-1} x_t \varepsilon_{i,t}.$$  \hspace{1cm} \hspace{1cm} (24)

Contrary to the analysis of skill, we are not the first to estimate the net alpha distribution $\phi(\alpha^n)$. However, the nonparametric approach proposed here brings several important advantages. As discussed in Section III, it is easier to apply and less prone to misspecification errors than standard Bayesian/parametric approaches (e.g., Chen, Cliff, and Zhao (2017), Harvey and Liu (2018a)). It also departs from the False Discovery Rate approach which estimates the proportions of non-zero alpha funds by counting the number of large alpha $t$-statistics (Barras, Scaillet, and Wermers (2010)). Our nonparametric approach estimates the entire distribution (not just the proportions), and has more power to detect funds with alphas close to zero (see the appendix).

The summary statistics for $\phi(\alpha^n)$ reported in Table VII provide some support to the BG model. The net alpha typically clusters around zero—it ranges between $\pm 1.5\%$ per year for around 70% of the funds.25 In addition, the proportion of positive alpha funds drops significantly from 68.4% to 34.7% when we move from a gross to a net basis.

However, the relation between the gross alpha and fees is quite loose. In particular, we find ample evidence of negative performance ($\pi^{-}_{\alpha^n} = 65.3\%$) and 5% of the funds produce a net alpha below -2.7% per year. To explain why a subset of the fund population generates large non-zero alphas, we therefore need additional elements beyond the BG model. The presence of search costs could be one of them (Garleanu and Pederson (2018)). If investors have to spend resources to detect skilled funds, they need to be compensated for these costs and earn a positive net alpha in equilibrium. Another element is the existence of unsophisticated investors to which funds charge excessively high fees (e.g., Christoffersen and Musto (2002), Gruber (1996)). The behaviour of these investors drives a wedge between gross alphas and fees that is left unexplained by any rational model.

Please insert Table VIII here

---

25 Consistent with the equality between the gross alpha and fees, Table VI further shows that fund groups with higher gross alphas charge higher fees (the pairwise correlation equals 0.57).
E Overall Implications

Our empirical analysis reveals that a large majority of funds create value from exploiting their skills and contribute to make equity prices more informative. This role is socially valuable because it improves the real allocation of resources. Individual funds help channel funds to the most promising new companies. They may do so directly by participating to initial public offerings, or indirectly via the secondary market. As discussed by Cochrane (2013) and Pedersen (2018), active funds make secondary markets more liquid and informationally efficient, which is likely to reduce the cost of capital in the primary market. In addition, active funds can have an impact on the economy if the cash flows of firms depend on the efficiency of the secondary market—a point summarized by Bond, Edmans, and Goldstein (2012). For instance, managers may learn from equity prices and improve their real investment decisions. They may also be better incentivized to exert effort if it is accurately reflected in prices.

The structure of the active industry determines how this social function is performed. First, the industry is not heavily concentrated because a minority of funds are skilled along the two dimensions (i.e., few have a high $a_i$ and low $b_i$). Therefore, each fund contributes to making prices more efficient. Second, mutual funds optimize their active size because they internalize the impact of capacity constraints. This implies that financial prices are less efficient than in a fully competitive equilibrium in which the active industry would be larger.

Our results that the active industry as a whole creates value is not inconsistent with the famous arithmetic of Sharpe (1991). This rule states that if passive investors do not trade and hold the market, the aggregate return of active investors must be equal to the market return. In reality, passive investors only hold a subset of the market as they concentrate on stocks included in indices. In addition, they must trade regularly when companies issue new shares or when the composition of indices changes. Therefore, even if we assume that mutual funds are representative of the entire population of active investors, they can still add value on aggregate by trading based on information and providing liquidity to passive investors (Pedersen (2018)). To illustrate this point, consider a set of passive investors that track the Russell 2000. Because this index exhibits an annual turnover close to 50% per year, passive investors are forced to rebalance their portfolios. When active investors accomodate these trades, they collectively gain between between 0.4% and 0.8% per year (Petajisto (2011)).

Finally, our analysis reveals strong evidence of skill and weak evidence of performance. We therefore conclude that individual funds are not only skilled—they are also in a strong bargaining position vis-a-vis the investors. This position might actually be
stronger than initially thought. As shown in Table III, a minority of funds actually
destroy value. Yet, these funds can still charge fees to investors which fail to detect
charlatans and incorporate the negative impact of capacity constraints.

F Additional Results

F.1 Alternative Asset Pricing Models

Our estimation of mutual fund skill obviously depends on the choice of the asset pric-
ing model. To examine the issue, we repeat our analysis using the four-factor model
of Carhart (1997) and the five-factor model of Fama and French (2015). Overall, the
distributions of the two skill dimensions remain largely unchanged. We observe two no-
ticeable differences. First, the average fd alpha among small cap funds drops from 4.6%
to 3.1% per year under the Carhart model. This is consistent with the analysis of CPZ
who show that the Carhart model assigns a negative alpha to the Russell 2000. Second,
the proportion of funds with a positive fd alpha decreases from 86.0% to 74.8% 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

As noted by Pastor, Stambaugh, and Taylor (2015), the estimated skill measures
\( \hat{\alpha}_i \) and \( \hat{\beta}_i \) can potentially be biased because the return residual \( \varepsilon_{i,t} \) is positively correlated
with the change in size \( \varepsilon_{qi,t} \). Whereas this bias vanishes asymptotically, it may have a
significant impact for funds with a small sample size (Stambaugh (1999)). To control
for this bias, we use the approach of Amihud and Hurvich (2004) and add \( \hat{\varepsilon}_{qi,t} \) to the
set of regressors in Equation (4) (see the appendix for details). The empirical results
reveal the estimated skill dimensions remain largely unchanged.

F.3 Alternative Predictors of the Gross Alpha

We examine alternative specifications to model the dynamics of the gross alpha. To
begin, we follow Harvey and Liu (2018b) and use the industry-adjusted size \( q_{it-1}^{ref} \), defined
as the ratio between the size of the fund and that of the active fund industry. The
rationale for this specification is that the impact of fund-level capacity constraints may
vary with the overall size of the industry. We find that our results remain largely unchanged.

We also examine whether the gross alpha is driven by industry-wide capacity con-
straints: \( \alpha_{i,t} = \alpha_i - b_i q_{t-1} \), where \( q_{t-1} \) is defined as the ratio of the industry size on the
total market capitalization. We confirm the result of Pastor, Stambaugh, and Taylor (2015) that industry-wide capacity constraints have a negative impact. However, this model is difficult to estimate at the individual fund level because the coefficients are poorly estimated—the condition number $CN_i$ increases significantly, which implies that only 373 funds satisfy the selection criterion in Equation (6).

Next, we test whether the relation between $\alpha_{i,t}$ and $q_{i,t-1}$ is nonlinear. Such nonlinearities occur if the fund has fixed operating costs $F_i$ (e.g., fixed costs of acquiring information). To examine this issue, we estimate the following model: $\alpha_{i,t} = (a_i q_{i,t-1} - b_i q_{i,t-1}^2 - F_i)/q_{i,t-1} = a_i - b_i q_{i,t-1} - F_i \frac{1}{q_{i,t-1}}$. Whereas the distributions for $a_i$ and $b_i$ remain largely unchanged, we find little support for this specification, i.e., $F_i$ has the wrong negative sign for the majority of the funds.

Finally, we include the age of the fund as an additional predictor of the gross alpha: $\alpha_{i,t} = a_i - b_i q_{i,t-1} - c_i age_{i,t-1}$. Consistent with Pastor, Stambaugh, and Taylor (2015), we find that age has a negative impact as 71% of them have a positive coefficient $c_i$. Controlling for age also increases the proportion of funds exposed to capacity constraints as $\hat{\pi}_b$ increases from 85.9% to 91.7%.

F.4 Investor Learning and Skill Priors

Our nonparametric approach yields estimates of the entire cross-sectional skill distribution. Therefore, it provides relevant information for modeling the prior distributions of $a_i$ and $b_i$ in an empirical Bayes setting. For instance, BG examine the prior distribution that investors have on the fd alpha. Calibrating their model using data on fund returns, survival rates, and flows, they find that around 80% of the funds achieve a positive fd alpha—a proportion that is very close to the one documented in Table II ($\hat{\pi}_a^+ = 86\%$). More recently, Pastor and Stambaugh (2012) elicit the joint prior distribution of $a_i$ and $b_i$ by setting their correlation equal to zero to ease Bayesian estimation. In contrast, the empirical evidence suggests that $a_i$ and $b_i$ are strongly correlated. Therefore, investors in their model take more time to learn because they believe that the gross alpha distribution is more spread out than the one inferred from the data.

VI Conclusion

In this paper, we apply a new approach for estimating the entire skill distribution across mutual funds. Our approach is nonparametric and thus particularly suited to the analysis of skill. It avoids the challenge of correctly specifying the skill distribution, and allows us to jointly examine multiple skill measures, including the two skill dimensions
(the fd alpha and size coefficient), and the two formulations of the value added (lifecycle and steady state). In addition to its flexibility, our approach is simple to implement, applicable to the different characterizations of the skill distribution (e.g., moments, quantiles), and supported by econometric theory.

Our analysis brings several insights into the active fund industry. First, it is skilled—around 85% of the funds are skilled at detecting profitable trades, and around 70% exhibit a positive value added once they reach the steady state size. Second, it is not heavily concentrated because the two skill dimensions—the fd alpha and size coefficient—are positively correlated. In other words, it is difficult for funds to have both investment and trading skills. Third, the active industry is close to its optimal size because funds internalize the impact of capacity constraints. As a result, its size is smaller than in a fully competitive equilibrium. Finally, it is in good bargaining position vis-a-vis the investors because we find strong evidence of skill, but weak evidence of performance.

Whereas our paper focuses on skill, our nonparametric approach has potentially wide applications in finance and economics. 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 when we want to capture the heterogeneity across stocks (e.g., risk exposure, commonality in liquidity), or in corporate finance when we want to capture the heterogeneity across firms (e.g., investment and financing decisions).
References


Panel A reports the average number of funds and the first four moments of the portfolio gross excess return for all funds in the population, four styles groups (small cap, large cap, growth, value), and four characteristic-sorted groups (low expense, high expense, low turnover, high turnover). Panel B reports the estimated portfolio betas on the market, size, value, and momentum factors, as well as the adjusted $R^2$ using the Cremers, Petajisto, and Zitzewitz benchmark model. All statistics are computed using monthly data between January 1975 and December 2018.

### Panel A: Gross Excess Return

<table>
<thead>
<tr>
<th></th>
<th>Average Nb. Funds</th>
<th>Mean (Ann.)</th>
<th>Volatility (Ann.)</th>
<th>Skewness</th>
<th>Kurtosis</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>All Funds</strong></td>
<td>937</td>
<td>8.0</td>
<td>14.8</td>
<td>-0.7</td>
<td>5.3</td>
</tr>
<tr>
<td><strong>Investment Styles</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Small-cap</td>
<td>188</td>
<td>9.8</td>
<td>18.7</td>
<td>-0.6</td>
<td>5.0</td>
</tr>
<tr>
<td>Large-cap</td>
<td>394</td>
<td>8.0</td>
<td>14.6</td>
<td>-0.7</td>
<td>5.2</td>
</tr>
<tr>
<td>Growth</td>
<td>401</td>
<td>8.3</td>
<td>16.4</td>
<td>-0.7</td>
<td>5.1</td>
</tr>
<tr>
<td>Value</td>
<td>242</td>
<td>7.9</td>
<td>13.6</td>
<td>-0.7</td>
<td>5.4</td>
</tr>
<tr>
<td><strong>Fund Characteristics</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low Expense</td>
<td>306</td>
<td>8.0</td>
<td>14.4</td>
<td>-0.7</td>
<td>5.2</td>
</tr>
<tr>
<td>High Expense</td>
<td>232</td>
<td>8.6</td>
<td>16.3</td>
<td>-0.8</td>
<td>5.0</td>
</tr>
<tr>
<td>Low Turnover</td>
<td>182</td>
<td>7.9</td>
<td>14.8</td>
<td>-0.8</td>
<td>5.4</td>
</tr>
<tr>
<td>High Turnover</td>
<td>181</td>
<td>9.1</td>
<td>16.6</td>
<td>-0.6</td>
<td>5.0</td>
</tr>
</tbody>
</table>

### Panel B: Estimated Betas

<table>
<thead>
<tr>
<th></th>
<th>Market</th>
<th>Size</th>
<th>Value</th>
<th>Momentum</th>
<th>Adj. $R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>All Funds</strong></td>
<td>0.93</td>
<td>0.25</td>
<td>-0.11</td>
<td>0.01</td>
<td>0.98</td>
</tr>
<tr>
<td><strong>Investment Styles</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Small-cap</td>
<td>0.98</td>
<td>0.78</td>
<td>-0.22</td>
<td>0.06</td>
<td>0.97</td>
</tr>
<tr>
<td>Large-cap</td>
<td>0.95</td>
<td>0.14</td>
<td>-0.06</td>
<td>0.01</td>
<td>0.99</td>
</tr>
<tr>
<td>Growth</td>
<td>0.95</td>
<td>0.34</td>
<td>-0.35</td>
<td>0.03</td>
<td>0.97</td>
</tr>
<tr>
<td>Value</td>
<td>0.91</td>
<td>0.12</td>
<td>0.22</td>
<td>-0.01</td>
<td>0.98</td>
</tr>
<tr>
<td><strong>Fund Characteristics</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low Expense</td>
<td>0.93</td>
<td>0.19</td>
<td>-0.05</td>
<td>0.01</td>
<td>0.98</td>
</tr>
<tr>
<td>High Expense</td>
<td>0.94</td>
<td>0.42</td>
<td>-0.27</td>
<td>0.02</td>
<td>0.97</td>
</tr>
<tr>
<td>Low Turnover</td>
<td>0.93</td>
<td>0.23</td>
<td>-0.07</td>
<td>-0.01</td>
<td>0.97</td>
</tr>
<tr>
<td>High Turnover</td>
<td>0.95</td>
<td>0.39</td>
<td>-0.31</td>
<td>0.11</td>
<td>0.95</td>
</tr>
</tbody>
</table>
Panel A contains the summary statistics on the cross-sectional distribution of the first skill dimension (the first dollar (fd) alpha) for all funds, four styles groups (small cap, large cap, growth, value), and four characteristic-sorted groups (low expense, high expense, low turnover, high turnover). It reports the first four moments, the proportions of funds with a negative and positive fd alpha, and the distribution quantiles at 5% and 95%. Panel B repeats the analysis for the second skill dimension (the size coefficient). All cross-sectional estimates are adjusted for bias (smoothing and EIV) using our non-parametric approach.

### Panel A: First Dollar Alpha

<table>
<thead>
<tr>
<th></th>
<th>Mean (Ann.)</th>
<th>Volatility (Ann.)</th>
<th>Skewness</th>
<th>Kurtosis</th>
<th>Negative</th>
<th>Positive</th>
<th>5%</th>
<th>95%</th>
</tr>
</thead>
<tbody>
<tr>
<td>All Funds</td>
<td>3.1</td>
<td>3.4</td>
<td>1.6</td>
<td>4.4</td>
<td>14.0</td>
<td>86.0</td>
<td>-1.3</td>
<td>8.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Investment Styles</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Small Cap</td>
<td>4.6</td>
<td>3.5</td>
<td>2.7</td>
<td>3.8</td>
<td>6.9</td>
<td>93.1</td>
<td>-0.8</td>
<td>10.1</td>
</tr>
<tr>
<td>Large Cap</td>
<td>1.8</td>
<td>2.5</td>
<td>2.4</td>
<td>17.8</td>
<td>19.9</td>
<td>80.1</td>
<td>-1.5</td>
<td>5.5</td>
</tr>
<tr>
<td>Growth</td>
<td>3.0</td>
<td>3.8</td>
<td>2.0</td>
<td>6.7</td>
<td>17.8</td>
<td>82.2</td>
<td>-2.0</td>
<td>8.6</td>
</tr>
<tr>
<td>Value</td>
<td>3.7</td>
<td>3.7</td>
<td>1.9</td>
<td>4.5</td>
<td>12.1</td>
<td>87.9</td>
<td>-1.2</td>
<td>9.5</td>
</tr>
<tr>
<td>Fund Characteristics</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low Expense</td>
<td>2.1</td>
<td>3.3</td>
<td>1.5</td>
<td>5.9</td>
<td>22.5</td>
<td>77.5</td>
<td>-2.4</td>
<td>6.8</td>
</tr>
<tr>
<td>High Expense</td>
<td>4.0</td>
<td>4.3</td>
<td>1.5</td>
<td>2.9</td>
<td>15.3</td>
<td>84.7</td>
<td>-1.7</td>
<td>10.3</td>
</tr>
<tr>
<td>Low Turnover</td>
<td>2.8</td>
<td>4.4</td>
<td>2.1</td>
<td>8.7</td>
<td>21.9</td>
<td>78.1</td>
<td>-2.8</td>
<td>8.6</td>
</tr>
<tr>
<td>High Turnover</td>
<td>3.6</td>
<td>5.1</td>
<td>2.0</td>
<td>5.2</td>
<td>21.0</td>
<td>79.0</td>
<td>-3.2</td>
<td>11.0</td>
</tr>
</tbody>
</table>

### Panel B: Size Coefficient

<table>
<thead>
<tr>
<th></th>
<th>Mean (Ann.)</th>
<th>Volatility (Ann.)</th>
<th>Skewness</th>
<th>Kurtosis</th>
<th>Negative</th>
<th>Positive</th>
<th>5%</th>
<th>95%</th>
</tr>
</thead>
<tbody>
<tr>
<td>All Funds</td>
<td>1.4</td>
<td>1.4</td>
<td>2.2</td>
<td>4.8</td>
<td>14.1</td>
<td>85.9</td>
<td>-0.6</td>
<td>3.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Investment Styles</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Small Cap</td>
<td>1.7</td>
<td>1.7</td>
<td>1.8</td>
<td>5.8</td>
<td>13.5</td>
<td>86.5</td>
<td>-0.9</td>
<td>4.4</td>
</tr>
<tr>
<td>Large Cap</td>
<td>1.0</td>
<td>1.0</td>
<td>2.8</td>
<td>11.9</td>
<td>18.0</td>
<td>82.0</td>
<td>-0.5</td>
<td>2.6</td>
</tr>
<tr>
<td>Growth</td>
<td>1.5</td>
<td>1.7</td>
<td>2.4</td>
<td>4.7</td>
<td>17.1</td>
<td>82.9</td>
<td>-1.0</td>
<td>4.1</td>
</tr>
<tr>
<td>Value</td>
<td>1.4</td>
<td>1.4</td>
<td>1.5</td>
<td>3.4</td>
<td>13.9</td>
<td>86.1</td>
<td>-0.7</td>
<td>3.7</td>
</tr>
<tr>
<td>Fund Characteristics</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low Expense</td>
<td>0.8</td>
<td>1.1</td>
<td>0.8</td>
<td>1.6</td>
<td>23.0</td>
<td>77.0</td>
<td>-0.8</td>
<td>2.6</td>
</tr>
<tr>
<td>High Expense</td>
<td>1.6</td>
<td>1.9</td>
<td>1.5</td>
<td>1.4</td>
<td>17.9</td>
<td>82.1</td>
<td>-1.1</td>
<td>4.6</td>
</tr>
<tr>
<td>Low Turnover</td>
<td>1.0</td>
<td>1.7</td>
<td>1.6</td>
<td>3.5</td>
<td>27.7</td>
<td>72.3</td>
<td>-1.5</td>
<td>3.5</td>
</tr>
<tr>
<td>High Turnover</td>
<td>1.5</td>
<td>2.3</td>
<td>2.0</td>
<td>8.6</td>
<td>22.0</td>
<td>78.0</td>
<td>-1.5</td>
<td>4.7</td>
</tr>
</tbody>
</table>
Table III
Cross-Sectional Distribution of the Value Added

Panel A contains the summary statistics on the cross-sectional distribution of the lifecycle value added for all funds, four styles groups (small cap, large cap, growth, value), and four characteristic-sorted groups (low expense, high expense, low turnover, high turnover). It reports the first four moments, the proportions of funds with a negative and positive value added, and the distribution quantiles at 5% and 95%. Panel B repeats the analysis for the steady state value added. All cross-sectional estimates are adjusted for bias (smoothing and EIV) using our non-parametric approach.

### Panel A: Lifecycle Value Added

<table>
<thead>
<tr>
<th></th>
<th>Mean (Ann.)</th>
<th>Volatility (Ann.)</th>
<th>Skewness</th>
<th>Kurtosis</th>
<th>Negative</th>
<th>Positive</th>
<th>5%</th>
<th>95%</th>
</tr>
</thead>
<tbody>
<tr>
<td>All Funds</td>
<td>1.7</td>
<td>10.6</td>
<td>5.8</td>
<td>47.8</td>
<td>41.9</td>
<td>58.1</td>
<td>-4.3</td>
<td>13.9</td>
</tr>
<tr>
<td><strong>Investment Styles</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Small Cap</td>
<td>3.8</td>
<td>11.5</td>
<td>4.4</td>
<td>22.8</td>
<td>35.3</td>
<td>64.7</td>
<td>-4.5</td>
<td>19.1</td>
</tr>
<tr>
<td>Large Cap</td>
<td>-0.5</td>
<td>9.5</td>
<td>5.1</td>
<td>50.6</td>
<td>56.3</td>
<td>43.7</td>
<td>-5.5</td>
<td>6.3</td>
</tr>
<tr>
<td>Growth</td>
<td>1.9</td>
<td>15.4</td>
<td>4.0</td>
<td>26.8</td>
<td>46.1</td>
<td>53.9</td>
<td>-6.6</td>
<td>20.0</td>
</tr>
<tr>
<td>Value</td>
<td>2.9</td>
<td>15.0</td>
<td>4.7</td>
<td>31.4</td>
<td>35.4</td>
<td>64.6</td>
<td>-4.4</td>
<td>18.7</td>
</tr>
<tr>
<td><strong>Fund Characteristics</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low Expense</td>
<td>6.1</td>
<td>35.0</td>
<td>4.6</td>
<td>23.7</td>
<td>42.5</td>
<td>57.5</td>
<td>-7.6</td>
<td>45.0</td>
</tr>
<tr>
<td>High Expense</td>
<td>1.7</td>
<td>8.6</td>
<td>1.8</td>
<td>3.5</td>
<td>38.4</td>
<td>61.6</td>
<td>-3.4</td>
<td>11.6</td>
</tr>
<tr>
<td>Low Turnover</td>
<td>4.6</td>
<td>30.3</td>
<td>0.6</td>
<td>13.8</td>
<td>32.6</td>
<td>67.4</td>
<td>-5.7</td>
<td>29.9</td>
</tr>
<tr>
<td>High Turnover</td>
<td>1.3</td>
<td>15.8</td>
<td>1.5</td>
<td>1.3</td>
<td>48.0</td>
<td>52.0</td>
<td>-8.6</td>
<td>15.2</td>
</tr>
</tbody>
</table>

### Panel B: Steady State Value Added

<table>
<thead>
<tr>
<th></th>
<th>Mean (Ann.)</th>
<th>Volatility (Ann.)</th>
<th>Skewness</th>
<th>Kurtosis</th>
<th>Negative</th>
<th>Positive</th>
<th>5%</th>
<th>95%</th>
</tr>
</thead>
<tbody>
<tr>
<td>All Funds</td>
<td>7.3</td>
<td>18.6</td>
<td>5.7</td>
<td>42.1</td>
<td>31.2</td>
<td>68.8</td>
<td>-3.8</td>
<td>34.5</td>
</tr>
<tr>
<td><strong>Investment Styles</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Small Cap</td>
<td>7.6</td>
<td>12.8</td>
<td>3.4</td>
<td>14.4</td>
<td>19.1</td>
<td>80.9</td>
<td>-2.2</td>
<td>29.0</td>
</tr>
<tr>
<td>Large Cap</td>
<td>4.5</td>
<td>14.1</td>
<td>6.2</td>
<td>48.4</td>
<td>37.6</td>
<td>62.4</td>
<td>-3.7</td>
<td>25.6</td>
</tr>
<tr>
<td>Growth</td>
<td>9.1</td>
<td>24.9</td>
<td>4.4</td>
<td>22.4</td>
<td>34.9</td>
<td>65.1</td>
<td>-4.7</td>
<td>45.9</td>
</tr>
<tr>
<td>Value</td>
<td>9.8</td>
<td>24.0</td>
<td>5.5</td>
<td>37.1</td>
<td>25.3</td>
<td>74.7</td>
<td>-3.4</td>
<td>41.3</td>
</tr>
<tr>
<td><strong>Fund Characteristics</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low Expense</td>
<td>11.7</td>
<td>36.9</td>
<td>4.5</td>
<td>21.7</td>
<td>35.5</td>
<td>64.5</td>
<td>-6.6</td>
<td>65.5</td>
</tr>
<tr>
<td>High Expense</td>
<td>5.4</td>
<td>10.6</td>
<td>3.5</td>
<td>14.0</td>
<td>25.6</td>
<td>74.4</td>
<td>-3.2</td>
<td>23.0</td>
</tr>
<tr>
<td>Low Turnover</td>
<td>9.9</td>
<td>30.8</td>
<td>3.0</td>
<td>17.4</td>
<td>21.8</td>
<td>78.2</td>
<td>-3.2</td>
<td>48.4</td>
</tr>
<tr>
<td>High Turnover</td>
<td>7.1</td>
<td>19.1</td>
<td>4.3</td>
<td>21.4</td>
<td>31.5</td>
<td>68.5</td>
<td>-4.8</td>
<td>33.4</td>
</tr>
</tbody>
</table>
### Table IV

**Impact of the Error-in-Variable (EIV) Bias**

This table compares the cross-sectional skill distributions with and without the adjustment for the Error-in-Variable (EIV bias). Panel A shows the summary statistics on the cross-sectional distribution of the first dollar (fd) alpha for all funds, four styles groups (small cap, large cap, growth, value), and four characteristic-sorted groups (low expense, high expense, low turnover, high turnover). It reports the first four moments, the proportions of funds with a negative and positive fd alpha, and the distribution quantiles at 5% and 95%. Panels C to D repeat the analysis for the size coefficient, the lifecycle value added, and the steady state value added.

#### Panel A: First Dollar Alpha

<table>
<thead>
<tr>
<th></th>
<th>Moments</th>
<th>Proportions (%)</th>
<th>Quantiles (Ann.)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean (Ann.)</td>
<td>Volatility (Ann.)</td>
<td>Skewness</td>
</tr>
<tr>
<td>With Bias Adjustement</td>
<td>3.1</td>
<td>3.4</td>
<td>1.6</td>
</tr>
<tr>
<td>Without Bias Adjustement</td>
<td>3.1</td>
<td>5.0</td>
<td>0.5</td>
</tr>
<tr>
<td>Difference</td>
<td>0.0</td>
<td>-1.6</td>
<td>1.2</td>
</tr>
</tbody>
</table>

#### Panel B: Size Coefficient

<table>
<thead>
<tr>
<th></th>
<th>Moments</th>
<th>Proportions (%)</th>
<th>Quantiles (Ann.)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean (Ann.)</td>
<td>Volatility (Ann.)</td>
<td>Skewness</td>
</tr>
<tr>
<td>With Bias Adjustement</td>
<td>1.4</td>
<td>1.4</td>
<td>2.2</td>
</tr>
<tr>
<td>Without Bias Adjustement</td>
<td>1.4</td>
<td>2.2</td>
<td>0.7</td>
</tr>
<tr>
<td>Difference</td>
<td>0.0</td>
<td>-0.8</td>
<td>1.5</td>
</tr>
</tbody>
</table>

#### Panel C: Lifecycle Value Added

<table>
<thead>
<tr>
<th></th>
<th>Moments</th>
<th>Proportions (%)</th>
<th>Quantiles (Ann.)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean (Ann.)</td>
<td>Volatility (Ann.)</td>
<td>Skewness</td>
</tr>
<tr>
<td>With Bias Adjustement</td>
<td>1.7</td>
<td>10.6</td>
<td>5.8</td>
</tr>
<tr>
<td>Without Bias Adjustement</td>
<td>1.7</td>
<td>17.9</td>
<td>3.5</td>
</tr>
<tr>
<td>Difference</td>
<td>0.0</td>
<td>-7.3</td>
<td>2.4</td>
</tr>
</tbody>
</table>

#### Panel D: Steady State Value Added

<table>
<thead>
<tr>
<th></th>
<th>Moments</th>
<th>Proportions (%)</th>
<th>Quantiles (Ann.)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean (Ann.)</td>
<td>Volatility (Ann.)</td>
<td>Skewness</td>
</tr>
<tr>
<td>With Bias Adjustement</td>
<td>7.3</td>
<td>18.6</td>
<td>5.7</td>
</tr>
<tr>
<td>Without Bias Adjustement</td>
<td>7.3</td>
<td>22.2</td>
<td>5.0</td>
</tr>
<tr>
<td>Difference</td>
<td>0.0</td>
<td>-3.7</td>
<td>0.7</td>
</tr>
</tbody>
</table>
Table V
Optimal Versus Actual Value Added

The table compares the optimal value added with the two formulations of the actual value added (lifecycle, steady state). Panel A contains the summary statistics on the cross-sectional distribution of the lifecycle value added for all funds, four styles groups (small cap, large cap, growth, value), and four characteristic-sorted groups (low expense, high expense, low turnover, high turnover). It reports the first four moments, the proportions of funds with a negative and positive value added, and the distribution quantiles at 5% and 95%. Panel B reports the mean and volatility, and the distribution quantiles at 5% and 95% of the difference between the optimal and lifecycle value added. Panel C repeats the analysis for the steady state value added. All cross-sectional estimates are adjusted for bias (smoothing and EIV) using our non-parametric approach.

### Panel A: Optimal Value Added

<table>
<thead>
<tr>
<th></th>
<th>Mean (Ann.)</th>
<th>Volatility (Ann.)</th>
<th>Skewness</th>
<th>Kurtosis</th>
<th>Negative (%)</th>
<th>Positive (%)</th>
<th>5%</th>
<th>95%</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>All Funds</strong></td>
<td>14.1</td>
<td>24.7</td>
<td>5.9</td>
<td>42.5</td>
<td>0.0</td>
<td>100.0</td>
<td>0.5</td>
<td>50.5</td>
</tr>
<tr>
<td><strong>Investment Styles</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Small Cap</td>
<td>11.4</td>
<td>15.8</td>
<td>6.9</td>
<td>60.1</td>
<td>0.0</td>
<td>100.0</td>
<td>0.8</td>
<td>37.3</td>
</tr>
<tr>
<td>Large Cap</td>
<td>12.4</td>
<td>27.0</td>
<td>6.4</td>
<td>50.9</td>
<td>0.0</td>
<td>100.0</td>
<td>0.4</td>
<td>48.8</td>
</tr>
<tr>
<td>Growth</td>
<td>16.4</td>
<td>30.3</td>
<td>6.8</td>
<td>64.5</td>
<td>0.0</td>
<td>100.0</td>
<td>0.6</td>
<td>61.6</td>
</tr>
<tr>
<td>Value</td>
<td>16.6</td>
<td>26.7</td>
<td>4.3</td>
<td>20.2</td>
<td>0.0</td>
<td>100.0</td>
<td>0.9</td>
<td>55.7</td>
</tr>
<tr>
<td><strong>Fund Characteristics</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low Expense</td>
<td>21.0</td>
<td>35.1</td>
<td>3.2</td>
<td>10.1</td>
<td>0.0</td>
<td>100.0</td>
<td>0.9</td>
<td>72.0</td>
</tr>
<tr>
<td>High Expense</td>
<td>10.3</td>
<td>12.4</td>
<td>5.1</td>
<td>28.9</td>
<td>0.0</td>
<td>100.0</td>
<td>0.6</td>
<td>32.4</td>
</tr>
<tr>
<td>Low Turnover</td>
<td>19.3</td>
<td>23.5</td>
<td>4.2</td>
<td>27.5</td>
<td>0.0</td>
<td>100.0</td>
<td>1.5</td>
<td>60.5</td>
</tr>
<tr>
<td>High Turnover</td>
<td>15.1</td>
<td>24.0</td>
<td>4.0</td>
<td>11.7</td>
<td>0.0</td>
<td>100.0</td>
<td>0.6</td>
<td>52.0</td>
</tr>
</tbody>
</table>

### Panel B: Difference with Lifecycle and Steady State Value Added

<table>
<thead>
<tr>
<th></th>
<th>Moments</th>
<th>Quantiles (Ann.)</th>
<th>vs Lifecycle Value Added</th>
<th>Moments</th>
<th>Quantiles (Ann.)</th>
<th>vs Steady State Value Added</th>
<th>Moments</th>
<th>Quantiles (Ann.)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean (Ann.)</td>
<td>Volatility (Ann.)</td>
<td>5%</td>
<td>95%</td>
<td>Mean (Ann.)</td>
<td>Volatility (Ann.)</td>
<td>5%</td>
<td>95%</td>
</tr>
<tr>
<td><strong>All Funds</strong></td>
<td>13.5</td>
<td>17.7</td>
<td>0.9</td>
<td>42.3</td>
<td>3.4</td>
<td>4.7</td>
<td>1.5</td>
<td>10.0</td>
</tr>
<tr>
<td><strong>Investment Styles</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Small Cap</td>
<td>9.9</td>
<td>11.4</td>
<td>1.1</td>
<td>28.0</td>
<td>3.0</td>
<td>4.3</td>
<td>3.5</td>
<td>9.4</td>
</tr>
<tr>
<td>Large Cap</td>
<td>13.8</td>
<td>19.5</td>
<td>0.9</td>
<td>44.0</td>
<td>3.2</td>
<td>4.1</td>
<td>1.2</td>
<td>9.0</td>
</tr>
<tr>
<td>Growth</td>
<td>17.1</td>
<td>26.2</td>
<td>1.1</td>
<td>57.0</td>
<td>4.5</td>
<td>7.0</td>
<td>0.8</td>
<td>14.9</td>
</tr>
<tr>
<td>Value</td>
<td>14.9</td>
<td>16.8</td>
<td>1.2</td>
<td>44.2</td>
<td>4.0</td>
<td>4.4</td>
<td>8.7</td>
<td>12.2</td>
</tr>
<tr>
<td><strong>Fund Characteristics</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low Expense</td>
<td>21.3</td>
<td>29.6</td>
<td>1.4</td>
<td>71.9</td>
<td>8.4</td>
<td>15.2</td>
<td>0.8</td>
<td>31.9</td>
</tr>
<tr>
<td>High Expense</td>
<td>10.1</td>
<td>8.8</td>
<td>1.2</td>
<td>28.5</td>
<td>3.1</td>
<td>2.8</td>
<td>6.1</td>
<td>8.6</td>
</tr>
<tr>
<td>Low Turnover</td>
<td>20.0</td>
<td>27.0</td>
<td>1.5</td>
<td>61.6</td>
<td>6.9</td>
<td>9.3</td>
<td>8.7</td>
<td>23.9</td>
</tr>
<tr>
<td>High Turnover</td>
<td>18.6</td>
<td>24.6</td>
<td>1.2</td>
<td>58.7</td>
<td>5.3</td>
<td>5.7</td>
<td>2.8</td>
<td>15.4</td>
</tr>
</tbody>
</table>
Table VI
Fund Fees and Size

Panel A describes the specific fee setting policies under which the gross alpha is informative about the skill dimensions or the value added. Each policy yields specific predictions regarding either fees or size. Panel B contains the summary statistics on the cross-sectional distribution of fund fees and size for all funds, four styles groups (small cap, large cap, growth, value), and four characteristic-sorted groups (low expense, high expense, low turnover, high turnover). It reports the mean, volatility, and the distribution quantiles at 5% and 95%.

Panel A: Fund Fees and Size under Specific Fee Setting Policies

<table>
<thead>
<tr>
<th>Fee Setting Policy</th>
<th>Scheme I (optimal size)</th>
<th>Scheme II (size coefficient)</th>
<th>Scheme III (value added)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fees are set such that</td>
<td>The fund size equals the optimal size</td>
<td>The fund size equals the squared optimal size</td>
<td>The fund size equals the median size</td>
</tr>
<tr>
<td>Main Prediction</td>
<td>Fees vary across funds to allow them to reach their optimal size</td>
<td>Fees are tiny to allow funds to reach their squared optimal size</td>
<td>Size is constant across all funds</td>
</tr>
<tr>
<td>Does the Gross Alpha Measure Skill?</td>
<td>First-Dollar Alpha (1st skill dimension)</td>
<td>Size Coefficient (2nd skill dimension)</td>
<td>Value Added</td>
</tr>
</tbody>
</table>

Panel B: Summary Statistics for Fund Fees and Size

<table>
<thead>
<tr>
<th></th>
<th>Mean Fees</th>
<th>Mean Size</th>
<th>Volatility Fees</th>
<th>Volatility Size</th>
<th>Quantile 5% Fees</th>
<th>Quantile 5% Size</th>
<th>Quantile 95% Fees</th>
<th>Quantile 95% Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>All Funds</td>
<td>1.25</td>
<td>785</td>
<td>0.39</td>
<td>2049</td>
<td>0.66</td>
<td>40</td>
<td>1.95</td>
<td>2901</td>
</tr>
<tr>
<td>Investment Styles</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Small-cap</td>
<td>1.36</td>
<td>388</td>
<td>0.35</td>
<td>614</td>
<td>0.87</td>
<td>43</td>
<td>2.00</td>
<td>1299</td>
</tr>
<tr>
<td>Large-cap</td>
<td>1.17</td>
<td>1062</td>
<td>0.37</td>
<td>2841</td>
<td>0.63</td>
<td>41</td>
<td>1.88</td>
<td>3937</td>
</tr>
<tr>
<td>Growth</td>
<td>1.30</td>
<td>791</td>
<td>0.39</td>
<td>2020</td>
<td>0.77</td>
<td>42</td>
<td>2.02</td>
<td>3323</td>
</tr>
<tr>
<td>Value</td>
<td>1.19</td>
<td>971</td>
<td>0.38</td>
<td>2481</td>
<td>0.61</td>
<td>42</td>
<td>1.87</td>
<td>3738</td>
</tr>
<tr>
<td>Fund Characteristics</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low Expense</td>
<td>0.83</td>
<td>1487</td>
<td>0.16</td>
<td>3594</td>
<td>0.48</td>
<td>51</td>
<td>1.03</td>
<td>6475</td>
</tr>
<tr>
<td>High Expense</td>
<td>1.72</td>
<td>359</td>
<td>0.28</td>
<td>622</td>
<td>1.39</td>
<td>36</td>
<td>2.21</td>
<td>1128</td>
</tr>
<tr>
<td>Low Turnover</td>
<td>1.18</td>
<td>1280</td>
<td>0.35</td>
<td>3215</td>
<td>0.67</td>
<td>51</td>
<td>1.83</td>
<td>4961</td>
</tr>
<tr>
<td>High Turnover</td>
<td>1.33</td>
<td>613</td>
<td>0.38</td>
<td>1163</td>
<td>0.81</td>
<td>48</td>
<td>1.99</td>
<td>2308</td>
</tr>
</tbody>
</table>
### Table VII
**Cross-Sectional Distribution of Gross Alpha**

The table contains the summary statistics on the cross-sectional distribution of the gross alpha for all funds, four styles groups (small-cap, large-cap, growth, value), and four characteristic-sorted groups (low-expense, high-expense, low-turnover, high turnover). It reports the first four moments, the proportions of funds with a negative and positive gross alpha, and the distribution quantiles at 5% and 95%. All cross-sectional estimates are adjusted for bias (smoothing and EIV) using our non-parametric approach.

<table>
<thead>
<tr>
<th></th>
<th>Moments</th>
<th>Proportions (%)</th>
<th>Quantiles (Ann.)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean (Ann.)</td>
<td>Volatility (Ann.)</td>
<td>Skewness</td>
</tr>
<tr>
<td>All Funds</td>
<td>0.7</td>
<td>1.5</td>
<td>0.3</td>
</tr>
<tr>
<td><strong>Investment Styles</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Small-cap</td>
<td>1.7</td>
<td>1.8</td>
<td>1.9</td>
</tr>
<tr>
<td>Large-cap</td>
<td>0.2</td>
<td>0.9</td>
<td>0.5</td>
</tr>
<tr>
<td>Growth</td>
<td>0.6</td>
<td>1.9</td>
<td>-0.4</td>
</tr>
<tr>
<td>Value</td>
<td>1.2</td>
<td>1.6</td>
<td>1.6</td>
</tr>
<tr>
<td><strong>Fund Characteristics</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low Expense</td>
<td>0.5</td>
<td>1.4</td>
<td>1.0</td>
</tr>
<tr>
<td>High Expense</td>
<td>1.2</td>
<td>2.4</td>
<td>0.3</td>
</tr>
<tr>
<td>Low Turnover</td>
<td>0.9</td>
<td>1.7</td>
<td>1.0</td>
</tr>
<tr>
<td>High Turnover</td>
<td>1.0</td>
<td>2.3</td>
<td>1.1</td>
</tr>
</tbody>
</table>
### Table VIII
**Cross-Sectional Distribution of Net Alpha**

The table contains the summary statistics on the cross-sectional distribution of the net alpha for all funds, four styles groups (small-cap, large-cap, growth, value), and four characteristic-sorted groups (low-expense, high-expense, low-turnover, high turnover). It reports the first four moments, the proportions of funds with a negative and positive net alpha, and the distribution quantiles at 5% and 95%. All cross-sectional estimates are adjusted for bias (smoothing and EIV) using our non-parametric approach.

<table>
<thead>
<tr>
<th></th>
<th>Moments</th>
<th>Proportions (%)</th>
<th>Quantiles (Ann.)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean (Ann.)</td>
<td>Volatility (Ann.)</td>
<td>Skewness</td>
</tr>
<tr>
<td>All Funds</td>
<td>-0.5</td>
<td>1.6</td>
<td>-1.0</td>
</tr>
<tr>
<td>Investment Styles</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Small-cap</td>
<td>0.3</td>
<td>1.7</td>
<td>0.9</td>
</tr>
<tr>
<td>Large-cap</td>
<td>-1.0</td>
<td>0.9</td>
<td>-1.5</td>
</tr>
<tr>
<td>Growth</td>
<td>-0.7</td>
<td>1.9</td>
<td>-1.6</td>
</tr>
<tr>
<td>Value</td>
<td>0.0</td>
<td>1.5</td>
<td>0.4</td>
</tr>
<tr>
<td>Fund Characteristics</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low Expense</td>
<td>-0.4</td>
<td>1.4</td>
<td>0.2</td>
</tr>
<tr>
<td>High Expense</td>
<td>-0.5</td>
<td>2.4</td>
<td>-0.9</td>
</tr>
<tr>
<td>Low Turnover</td>
<td>-0.2</td>
<td>1.6</td>
<td>0.0</td>
</tr>
<tr>
<td>High Turnover</td>
<td>-0.4</td>
<td>2.3</td>
<td>-0.1</td>
</tr>
</tbody>
</table>
Figure 1
Cross-sectional Distributions of the Two Skill Dimensions: Analysis across Fund Groups

Panel A plots the cross-sectional densities of the first dollar alpha for small cap and large cap funds. Panel B compares growth and value funds. Panel C compares low expense and high expense funds. Finally, Panel D compares low turnover and high turnover funds. All the estimated densities are adjusted for bias (smoothing and EIV) using our non-parametric approach.
Figure 1
Cross-sectional Distributions of the Two Skill Dimensions: Analysis across Fund Groups (Continued)

Panel A plots the cross-sectional densities of the size coefficient for small cap and large cap funds. Panel B compares growth and value funds. Panel C compares low expense and high expense funds. Finally, Panel D compares low turnover and high turnover funds. All the estimated densities are adjusted for bias (smoothing and EIV) using our non-parametric approach.