Skill and Value Creation in the Mutual Fund Industry

Laurent Barras, Patrick Gagliardini, and Olivier Scaillet*

This version, August 29, 2019

JEL Classification: G11, G12, C14, C33, C58.
Keywords: Mutual funds, skill, value added, nonparametric estimation, large unbalanced 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, Yong Chen, Jean-Daniel Fermanian, Francesco Franzoni, Laurent Fresard, Alexander Kempf, Gaelle Le Fol, Peter Limbach, Jordan Martel, 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, FMND 2019, QFFE 2019, the 2019 Cambridge Big Data Workshop, the 2019 CICF conference, the 2019 European Finance Association (EFA) Meeting as seminar participants at Geneva, McGill, McMaster, Lugano, Luxemburg, and ULB 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 distributions of mutual fund skill and its economic value. Our approach avoids the challenge of specifying these distributions and accommodates the need to study jointly multiple skill measures. Our results show that most funds (i) are skilled at detecting profitable trades, (ii) face tight capacity constraints, and (iii) extract positive value from their skills. We also show that investment and trading skills vary substantially across funds and are strongly correlated—two features that are partly driven by the fund strategies. Finally, we find that the fund industry (i) is not heavily concentrated, (ii) does a good job at maximizing profits, and (iii) is in a strong bargaining position vis-a-vis 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.[1] 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 unique investment and trading 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.

Measuring 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 how many funds create value from exploiting their skills. Third, it uncovers important features of the mutual fund industry regarding its levels of concentration, profits, and bargaining power vis-a-vis investors. Finally, it sheds light on the social value of active management. If funds are skilled, they may improve price efficiency and the allocation of resources in the economy.

In this paper, we propose an extensive analysis of individual fund skill. We develop a novel approach to estimate the entire distributions of both skill and its economic value. To measure skill, we build on the idea that funds have different abilities to invest and trade. To formalize this idea, we use the economic model of Berk and Green (2004; BG hereafter) in which the gross alpha α_{it} of each fund decreases with its size $q_{i,t-1}$: $\alpha_{i,t} = a_i - b_i q_{i,t-1}$. Using this model, we uncover two skill dimensions: (i) the first dollar (fd) alpha a_i captures the fund skill at generating investment ideas, and (ii) the size coefficient b_i captures the fund skill at mitigating capacity constraints.[2]

To measure the economic value of skill, we build on the concept of value added (see BvB). Similar to the rent of a monopolist, it captures the dollar profit earned by the fund from exploiting its skill. To distinguish between short- and long-term profits, we examine: (i) the lifecycle value added $va_i^l = E[\alpha_{i,t} q_{i,t-1}]$ which is measured over

[1]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).

[2]For instance, the Dimensional Fund Advisors (DFA) fund highlights its ability to efficiently trade small-cap stocks as a source of value creation. In addition to minimizing price impact when buying in the open market, DFA has developed an expertise in buying large blocks of shares from sellers in need of liquidity (Cohen (2002)).

the lifecycle of the fund, and (ii) the steady state value added $va_i^{ss} = E\left[\alpha_{i,t}\right] E\left[q_{i,t-1}\right]$ which is measured once the fund reaches its average size. In total, our framework jointly considers four different measures, which we denote by $m_i \in \{a_i, b_i, va_i^l, va_i^{ss}\}$.

Our estimation approach is nonparametric, simple, and supported by econometric theory. It is nonparametric as it imposes minimal structure of the shape of each distribution $\phi(m)$. It is simple to apply—intuitively, it is akin to computing an histogram using as inputs the estimated fund measures $\hat{m}_i$ $(i = 1, 2, ...)$ obtained from time-series regressions. Finally, it is supported by econometric theory because we formally derive the asymptotic properties of the estimated distribution. This allows us to conduct proper statistical inference in a large and possibly unbalanced panel of funds.

Our approach departs from standard Bayesian/parametric approaches recently used to infer the distribution of fund performance (e.g., Chen, Cliff, and Zhao (2017), Harvey and Liu (2018a), Jones and Shanken (2005)). These approaches require a full parametric specification of the distribution, as well as complex and computer-intensive methods (Gibbs sampling, Expectation Maximization). Both requirements make these approaches ill suited for the analysis of skill. First, they are prone to mispecification errors because theory offers little guidance to specify how skill varies across funds. Second, they have a limited ability to analyze multiple measures (a_i, b_i, va_i^l, and va_i^{ss}). Whereas a joint specification is subject to the curse of dimensionality, a separate specification is vulnerable to inconsistencies because all measures are intertwined.

The main estimation challenge is to adjust for the bias. Because the true fund measures m_i are unobservable, we can only rely on the estimated measures $\hat{m}_i$ to infer the distribution $\phi(m)$. 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 obtain the bias-adjusted distribution $\hat{\phi}^*(m)$. The bias adjustment is easy to interpret, available in closed form, and validated through an extensive Monte Carlo analysis. It also plays a major role in the estimation—we find that the unadjusted distribution $\hat{\phi}(m)$ overestimates the probability in the tails, does not capture the strong asymmetry in skill, and wrongly signals that the majority of funds fail to create value.

Our empirical analysis of US equity funds between January 1975 and December 2018 reveals several new insights. First, the ability of funds to detect profitable trades is both widespread and economically significant. Controling for the standard risk factors, we find that the fd alpha a_i is positive for 86% of the funds, and reaches 3.1% per year on average. Second, individual funds face increasing trading costs as they grow large—

the size coefficient b_i is positive for around 85% of the funds. This result provides strong support to models that emphasize the importance of capacity constraints in the mutual fund industry (e.g., BG, Pastor and Stambaugh (2012)). Third, skill generates substantial economic value. Around 60% of the funds create value over their lifecycle ($va_i^l > 0$), which is, on average, equal to \$1.7 mio. per year—a number similar to that reported by BvB (\$2.0 mio.). Fourth, the economic value increases dramatically once funds reach their average size—the steady state value added va_i^{ss} is positive for around 70% of the population, and is equal to \$7.3 mio. per year on average. The sharp difference of \$5.2 mio. highlights the economic importance of capacity constraints which lower the value added as size moves away from its average (e.g., $b_i > 0$ implies $va_i^l - va_i^{ss} < 0$).

Our fund-level analysis allows us to uncover two important properties of the skill dimensions. First, they vary significantly across individual funds. 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. Second, they are strongly correlated—the pairwise correlation between the estimated values $\hat{a}_i$ and $\hat{b}_i$ is equal to 0.82. This strong heterogeneity and correlation are partly driven by the strategy followed by funds (Pastor, Stambaugh, and Taylor (2019)). For instance, investing in small cap stocks involves illiquidity. As trading costs increase, it becomes more difficult to arbitrage any mispricing away. Consistent with this intuition, we find that small cap funds are more skilled at detecting profitable trades than large cap funds (higher a_i), but also more exposed to capacity constraints (higher b_i). Similarly, high turnover funds trade often which allows them to exploit more profitable opportunities (higher a_i) at the cost of higher trading costs (higher b_i).

The difficulty for funds to be skilled along the two dimensions implies that the industry is not concentrated. The top 5% of the funds only capture 20.8% of the total value added (versus 92.7% if a_i and b_i were uncorrelated). The skill correlation also highlights the importance of the value added in comparing funds. Because funds with more profitable ideas typically face higher trading costs, it is a priori unclear whether they dominate funds with more scalable investment strategies. Our empirical analysis reveals that this is indeed the case for small cap funds, but not for high turnover funds.

Finally, we use our novel approach to study three equilibrium predictions implied by the BG model. First, we find supportive evidence that funds maximize their value added. On average, the steady state value added va_i^{ss} represents 76% of the optimal value added level va_i^* defined as $\frac{a_i^2}{4b_i}$. Consistent with learning effects, we also find that the lifecycle value added va_i^l remains far from the optimal level. Because investors must learn about

a_i and b_i using past data, their short-term allocation can be quite different from the fund optimal size (Pastor and Stambaugh (2012)). Second, we confirm that the gross alpha $\alpha_i = E[\alpha_{i,t}]$ is a noisy measure of the value added. In the BG model, α_i can take any value because the fees chosen by funds are arbitrary (see BvB). We show that α_i is only informative about a_i, b_i, or va_i^* if all funds coordinate on specific fee setting policies. However, our comparison of low and high expense funds provides limited evidence of such coordination. Third, we clearly reject the prediction that the net alphas are equal to zero. This result resonates with the previous literature on performance and implies that additional elements beyond the BG model are necessary to explain why funds have both negative and positive alphas.

Overall, the extensive value created by the mutual fund industry has several key implications. It strongly suggests that active funds make financial prices more efficient and contribute to the allocation of real resources in the economy. Some funds may do so by participating to the primary market, or by improving the information efficiency of the secondary market (e.g., Bond, Edmans, and Goldstein (2012)). Importantly, it does not contradict the famous arithmetic of Sharpe (1991) under which the active industry cannot beat the market. As noted by Pedersen (2018), this rule breaks down because passive investors only hold a subset of the market and trade regularly (e.g., IPOs, index turnover). Therefore, it is not theoretically inconsistent that the industry as a whole beats the market by trading on information and accomodating liquidity needs. Finally, the combination of high value added and low performance reveals that mutual funds are not only skilled, but also in a strong bargaining position vis-a-vis their 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 nonparametric approach to multiple measures of skill and value added. 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 (2019)). This paper focuses on skill and estimates the entire distribution and its moments (not just the proportions). BG, BvB, and Pastor, Stambaugh, and Taylor (2015) introduce several measures of skill and economic value. We largely build on their framework to define our different 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 to capture the cross-sectional variation in trading skill.

The remainder of the paper is as follows. Section II presents the different measures of skill and value added. 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 Mutual Fund Skill and Value Creation

A The Two Skill Dimensions of Mutual Funds

We begin our analysis by describing how to measure mutual fund skill. Our framework incorporates the widely held view that mutual funds are exposed to capacity constraints (e.g., BG, Pastor and Stambaugh (2012), Perold and Salomon (1991)). In a world with capacity constraints, each fund faces increasing trading costs as it grows large. To generate high returns, it must therefore be skilled along two dimensions. The first dimension captures the fund skill at identifying profitable opportunities. The second dimension captures the fund skill at minimizing the price impact of its trades.

To measure the 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. For each fund, the total (benchmark-adjusted) revenue from active management is given by $TR_{i,t} = a_i q_{i,t-1}$, where $q_{i,t-1}$ denotes the lagged fund size. The total cost 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 $TR_{i,t} - TC_{i,t}$ and dividing by $q_{i,t-1}$, we obtain the fund gross alpha,

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

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

The BG model is particularly convenient because it provides a simple measurement of the two skill dimensions. The first dimension is measured by the first dollar (fd) alpha a_i. This measure isolates the profitability of the fund ideas by determining 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)). The second skill dimension is measured by the size coefficient b_i. This coefficient determines the sensitivity of the gross alpha to changes in fund size—a low value of b_i signals that the fund is able to mitigate capacity constraints.

A key feature of our framework is that we allow a_i and b_i to be fund specific. Therefore, we provide the first structural estimation of the BG model at the individual fund level. To do so, we treat a_i and b_i not as fixed parameters, but as random realizations from the cross-sectional skill distributions $\phi(a)$ and $\phi(b)$.[3] 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.[4]

Our framework for measuring skill calls for two main comments. First, Equation (1) embeds different sources of variation in skill that are not explicitly modeled. Skill can potentially vary because some funds have unique investment and trading abilities. For instance, some funds may be run by extremely talented managers or benefit from a high speed of information dissemination within their family (Cici, Jaspesen, and Kempf (2017)). Skill can also vary because funds follow specific strategies, such as investing in small cap stocks or trading at high frequencies. These different skill components may further reinforce each other through complementarity effects—for one, fund families may want to allocate their top managers to strategies which reward managerial skill the most (Fang, Kempf, and Trapp (2014)). To formalize these points, we can write $a_i = g_a(a_i^f, a_i^m, s_i)$ and $b_i = g_b(b_i^f, b_i^m, d_i)$ for given functions g_a and g_b, where a_i^f, b_i^f denote the unique skills of the fund family (f), a_i^m, b_i^m denote the unique skills of the fund manager (m), and the vector d_i captures the characteristics of the fund strategy. For instance, Pastor, Stambaugh, and Taylor (2019) specify d_i as $(liq_i, turn_i)'$, where liq_i and $turn_i$ denote the levels of liquidity and turnover chosen by the fund.

Second, we focus on Equation (1) because it provides a simple and natural starting point for capturing the two skill dimensions. However, we can easily extend our baseline framework by writing the gross alpha as $\alpha_{i,t} = a_i - b_i q_{i,t-1} - c_i' z_{i,t-1}$, where $z_{i,t-1}$ is the Q-vector of additional variables. The vector $z_{i,t-1}$ potentially includes (i) different functions of $q_{i,t-1}$ to model more complicated cost functions (e.g., fixed operating costs), and (ii) additional predictors to accommodate richer alpha dynamics (e.g, business cycle indicators, variables specific to the fund, the family, or the manager).[5] In the empirical section of the paper, we examine several alternatives to our baseline specification.

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

[4] To control for heterogeneity across funds, it is also 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 simply allow b_i to vary across funds.

[5] See, for instance, Chen et al. (2004), Kacperczyk, van Nieuwerburgh, and Veldkamp (2014), Fang, Kempf, and Trapp (2014), and Pastor, Stambaugh, and Taylor (2015, 2017).

B The Value Created by Mutual Funds

Equation (1) captures skill but does not measure the economic value associated with a_i and b_i. To address this issue, we use the value added (see BvB). This measure is defined as the economic rent (in dollars) earned by each fund from exploiting its skills. As such, it has an intuitive interpretation—it is similar to the rent of a monopolist measured as the markup price of the good multiplied by the total quantities sold.

To formalize this intuition, we consider two measures of the value added. First, we use the lifecycle (l) value added va_i^l initially proposed by BvB. This measure captures the value added by the fund during its entire lifecycle:

$$va_i^l = \underset{T\to\infty}{\text{plim}} \frac{1}{T}\sum_{t=1}^{T} \alpha_{i,t} \cdot q_{i,t-1} = E\left[\alpha_{i,t} q_{i,t-1}\right] = a_i E[q_{i,t-1}] - b_i E[q_{i,t-1}^2], \tag{2}$$

where $E[q_{i,t-1}]$ and $E[q_{i,t-1}^2]$ denote the time-series averages of the fund size and its squared value. In other words, va_i^l captures the average value added across the different size levels at which the fund operates. Second, we propose a new measure referred to as the steady state (ss) value added va_i^{ss}. Contrary to va_i^l, this measure captures the value added once the fund reaches its average, or steady state size $E[q_{i,t-1}]$:

$$\begin{aligned} va_i^{ss} &= \underset{T\to\infty}{\text{plim}} \frac{1}{T}\sum_{t=1}^{T} \alpha_{i,t} \cdot \underset{T\to\infty}{\text{plim}} \frac{1}{T}\sum_{t=1}^{T} q_{i,t-1} = E\left[\alpha_{i,t}\right] E\left[q_{i,t-1}\right] \\ &= a_i E[q_{i,t-1}] - b_i E[q_{i,t-1}]^2. \end{aligned} \tag{3}$$

Similar to Equation (1), we allow the value added in Equations (2) and (3) to be fund specific. To capture the heterogeneity across funds, we treat va_i^l and va_i^{ss} as random realizations from the cross-sectional distributions $\phi(va^l)$ and $\phi(va^{ss})$.

In a world with capacity constraints, the steady state value added is greater than the lifecycle value added. To see this point, we can write their difference as $va_i^l - va_i^{ss} = cov(\alpha_{i,t}, q_{i,t-1}) = -b_i var(q_{i,t-1})$, and note that the covariance between the gross alpha and size is negative when b_i is positive. Therefore, the comparison of the two measures quantifies the value loss as the fund progressively converges towards the steady state.

It is tempting to estimate how much value the fund creates relative to its total size. This measure, which is obtained by dividing va_i^{ss} by $E[q_{i,t-1}]$, is nothing else that the average gross alpha $\alpha_i = E[\alpha_{i,t}]$.[6] However, the gross alpha is a noisy measure of value

[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). It is commonly defined as the intercept α_i^u from the regression of $r_{i,t}$ on the risk factors f_t: $r_{i,t} = \alpha_i^u + \beta_i' f_t + \varepsilon_{i,t}$. Whereas

creation precisely because it does not control for differences in fund size (BvB)—a point that we formally examine below. Such differences arise naturally because funds have different skill levels and may choose different fee setting policies. Therefore, using the gross alpha is akin to measuring the monopolist rent with the markup price of the goods, regardless of how much quantity is sold.

We illustrate this point with a simple example in which (i) fund A has more investment and trading skills than fund B ($a_A > a_B$, $b_A < b_B$), but (ii) chooses the same level of fees f_e (see the appendix for a detailed description). If investors compete for performance (as in the BG model), they allocate money to both funds until they break even. As a result, both funds produce the same gross alphas: $\alpha_A = \alpha_B = f_e$. This implies that the gross alpha fails to capture that fund A creates more value than fund B. It also fails to capture that fund A is more skilled than fund B on every dimension. the differences in skill—fund A has more ideas (higher a_i) and trades more efficiently (lower b_i).

III Overview of the Nonparametric Approach

A General Motivation

We now describe the approach for estimating the cross-sectional 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 should cluster around zero, it offers no such guidance for skill/value added. 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]

α_i^u is not equal to α_i because of the omitted size variable $q_{i,t-1}$, the difference is negligible if $q_{i,t-1}$ is weakly correlated with f_t (which is typically the case in the data).

[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)). Here, inference is even more complicated because we do not observe the true skill measures, but only the estimated ones.

Second, it allows for a joint analysis of all four measures. Such analysis is extremely challenging with Bayesian/parametric approaches because they involve the daunting task of correctly specifying and estimating a multivariate distribution whose marginals are potentially mixtures of distributions. To sidestep this challenge, it is tempting to specify and estimate each distribution separately. However, this procedure is likely to generate inconsistencies because the measures of skill and value added 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 ϕ, along with the other characterizations of the distribution, including the cumulative distribution function $\Phi(x) = P[m_i \leq x] = \int_{-\infty}^{x} \phi(u)du$, the moments (e.g., mean, variance), and the distribution quantile $Q(p) = \Phi^{-1}(p)$, where p denotes the probability level.

Last but not least, it comes with a full-fledged inferential theory. We derive the asymptotic distribution of each estimator as the numbers of funds n and return observations T grow large (simultaneous double asymptotics with $n, T \to \infty$). We can therefore determine its asymptotic properties and conduct proper statistical inference guided by theoretical results.

B Estimation Procedure

B.1 Estimation of the Different Measures

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

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

where $r_{i,t}$ is the fund gross excess return (before fees) over the 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 β_i are random realizations from a continuum of funds. Under this sampling scheme, we can invoke cross-sectional limits to infer the density of each skill measure m.[8,9]

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

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

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, $\overline{q}_{i,1} = \frac{1}{T_i} \sum_{t=1}^{T} I_{i,t} q_{i,t-1}$, $\overline{q}_{i,2} = \frac{1}{T_i} \sum_{t=1}^{T} I_{i,t} q_{i,t-1}^2$, we can then infer each of the four measures as

$$\begin{aligned}
\text{Fd alpha} &: \quad \hat{m}_i = \hat{a}_i, \\
\text{Size coefficient} &: \quad \hat{m}_i = \hat{b}_i, \\
\underset{\text{(lifecycle)}}{\text{Value added}} &: \quad \hat{m}_i = \widehat{va}_i^{l} = \hat{a}_i \overline{q}_{i,1} - \hat{b}_i \overline{q}_{i,2}, \\
\underset{\text{(steady state)}}{\text{Value added}} &: \quad \hat{m}_i = \widehat{va}_i^{ss} = \hat{a}_i \overline{q}_{i,1} - \hat{b}_i \overline{q}_{i,1}^2.
\end{aligned} \tag{6}$$

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

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

where $CN_i = \sqrt{eig_{max}\left(\hat{Q}_{x,i}\right) / eig_{\min}\left(\hat{Q}_{x,i}\right)}$ is the condition number of the matrix $\hat{Q}_{x,i}$ defined as the ratio of the largest to smallest eigenvalues eig_{max} and eig_{min}, $\tau_{i,T} = T/T_i$ is the inverse of the relative sample size T_i/T, and $\chi_{1,T}$, $\chi_{2,T}$ denote the two threshold values. The first condition $\{CN_i \leq \chi_{1,T}\}$ excludes funds for which the time series

[8] 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.

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

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 $n_\chi = \sum_{i=1}^n \mathbf{1}_i^\chi$.

B.2 Kernel Density Estimation

In the next step, we estimate the skill density function using a standard nonparametric approach based on kernel smoothing.[10] The estimated density $\hat{\phi}$ at a given point m is computed as

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

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

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

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

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

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

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

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

where $K_1 = \int K(u)^2 du$, $K_2 = \int u^2 K(u) du$, $\phi^{(2)}(m)$ is the second derivative of the density $\phi(m)$ and $\psi^{(2)}(m)$ is the second derivative of the function $\psi(m) = \omega(m)\phi(m)$

[10]See, for instance, Ait-Sahalia (1996), Ait-Sahalia and Lo (1998), and Stanton (1997) for applications of kernel density estimation in finance.

with $\omega(m) = E[S_i | m_i = m]$. *The term* S_i *is the asymptotic variance of the estimated centered measure* $\sqrt{T}(\hat{m}_i - m_i)$ *equal to* $\underset{T\to\infty}{plim}(\frac{\tau_{i,T}^2}{T}\sum_{t,s=1}^{T} I_{i,t}I_{i,s}u_{i,t}u_{i,s})$. *For each measure, the term* $u_{i,t}$ *is given by*

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

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

Proof. See the appendix. ■

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 in nonparametric statistics—it arises because we estimate ϕ using the estimated measures instead of the true ones (i.e., $\hat{m}_i$ instead of m_i).

Finally, Proposition III.1 provides guidelines for the choice of the bandwidth. We show in the appendix 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:[11] (i) if T is small relative to n ($n^{2/5}/T \to \infty$), h^* is proportional to $(nT)^{-\frac{1}{3}}$; (ii) if T is large relative to n ($n^{2/5}/T \to 0$), h^* is proportional to

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

$n^{-\frac{1}{5}}$. 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:

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

B.3 Bias Adjustment

Our final step is to adjust the kernel density estimator $\hat{\phi}(m)$ for the bias. To do so, we apply a Gaussian reference model to compute the two bias terms and the optimal bandwidth given in Proposition III.1.[12] Under this model, the fund measure m_i and the log of the asymptotic variance $s_i = \log(S_i)$ are drawn from a bivariate normal distribution where $m_i \sim N(\mu_m, \sigma_m^2)$, $s_i \sim N(\mu_s, \sigma_s^2)$, and $corr(m_i, s_i) = \rho$.

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

These benefits are not shared by a fully nonparametric approach in which the bias terms are inferred from Equations (10)-(11) via a nonparametric estimation of the second-order derivatives $\phi^{(2)}$ and $\psi^{(2)}$. Estimating these derivative terms is notoriously difficult and generally leads to large estimation errors (e.g., Wand and Jones (1995; ch. 2)).[13] 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 Gaussian reference model as the number of

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

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

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

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

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

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

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

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

Proof. See the appendix ■

Equations (14)-(15) imply that the smoothing bias is negligible, whereas the EIV bias is not. As the total number of funds n increases, h^* shrinks towards zero, which reduces the magnitude of $bs_1^r(m)$. With a population of several thousand funds, the smoothing term becomes negligible for all values of m. In contrast, $bs_2^r(m)$ depends on the number of observations T because it arises from the gap between $\hat{m}_i$ and m_i. Therefore, the EIV bias remains significant even if the fund population is large. Another insight from Equation (15) is that the magnitude of the EIV bias depends on the variances of $\hat{m}_i$ and m_i which are defined as σ_m^2 and $\sigma_{\hat{m}}^2 = \frac{1}{T}\exp(\bar{\mu}_s)$. As σ_m increases relative to $\sigma_{\hat{m}}$, the EIV bias becomes less severe because it makes the cross-sectional variation of the estimated measure $\hat{m}_i$ more aligned with that of the true measure m_i (and vice-versa). The appendix contains a detailed comparative static analysis of the magnitude of the EIV bias.

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

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

where $\hat{u}_{i,t}$ is obtained by plugging the estimated quantities for the chosen measure in

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

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

An important question is whether the EIV bias obtained with the normal reference model provides a good approximation of the true bias (i.e., whether $bs_2^r(m) \approx bs_2(m)$). Two compelling arguments show that this is the case. First, Proposition III.1 shows that the true bias $bs_2(m)$ is a function of the second-order derivative of the true density ϕ. As long as ϕ peaks around its mean, this derivative takes negative values in the center and positive values in the tails—exactly like the function $bs_2^r(m)$.[14] 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).[15]

With the bias-adjusted density at hand, we can easily estimate the cumulative distribution function (cdf), moments, and quantiles via numerical integration. For instance, the proportion of funds with a negative m_i is given by $\hat{\pi}^- = \int_{-\infty}^{0} \hat{\phi}^*(u)du$. An alternative approach is to directly use the bias-adjusted estimators of all these quantities (cdf, moments, quantiles) which are derived in the appendix. Whereas both approaches are asymptotically equivalent, the Monte-Carlo analysis reveals that the numerical integration produces a lower Mean Squared Error (MSE). Motivated by these results, we use this approach in the empirical section of the paper.

IV Data Description

A Mutual Fund Data and Benchmark Model

We conduct our analysis on the entire population of open-end actively managed US equity funds. We collect monthly data on net returns and size, as well as annual data on fees, turnover, and investment objectives from the CRSP database between January 1975 and December 2018. This allows us to construct the time-series of the gross return

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

[15]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.

and size for the population and different groups (small/large cap, 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 departs 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 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.[16]

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, the Russell 2000 has a negative alpha of -2.4% per year over the period 1980-2005. Therefore, small cap fund that use this index as a benchmark are likely be classified as unskilled under the Carhart model.[17]

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. To this end, we construct an equally-weighted portfolio of all existing funds at the start of each month. In Panel A, we report the first four moments and first-order autocorrelation of the portfolio gross excess returns. In the entire population, the portfolio achieves a risk-return tradeoff similar to that of the aggregate stock market (8.8% and 15.3% per year). It also exhibits a negative skewness (-0.75) and a positive kurtosis (5.34). The results are similar across groups, except for the small cap portfolio which produces higher mean and volatility.

In Panel B, we report the estimated portfolio betas and adjusted R^2 obtained with the CPZ model. We find that small cap funds are heavily exposed to the size factor (0.77), which is also the case for high turnover funds. (0.50). Finally, Panel C reports

[16]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.

[17]In the appendix, we re-estimate the distributions of skill and value added using alternative models. With the exception of small-cap funds, we find that the results remain largely unchanged.

additional characteristics that include the average number of funds in the portfolio and its average size, fees, turnover, and age (obtained by averaging across funds and then over time). Consistent with intuition, small cap funds manage a small asset base—the average size is equal to \$371 mio. versus \$1,286 mio. for large cap funds. We also find that high turnover funds trade very aggressively. The annual average turnover reaches 175% versus 43% for low turnover funds.

Please insert Table I here

V Empirical Results

A Skill and Value Creation in the Mutual Fund Industry

A.1 The Two Skill Dimensions

We begin our analysis with a bird's eye view of the two skill dimensions among mutual funds—the fd alpha and size coefficient. We estimate a_i and b_i for each fund (as per Equation (12)), 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). The summary statistics for $\phi(a)$ and $\phi(b)$ are shown in Panel A of Table II. To ease interpretation, we standardize $\hat{b}_i$ for each fund in Table II so that it corresponds to the change in gross alpha for a one standard deviation change in size.

We find overwhelming evidence that mutual funds are able to detect profitable investment ideas. The fd alpha is positive for 85.8% of the funds in the population, and economically large with an average level of 3.1% per year. At the same time, individual funds face increasing trading costs as they grow large. Around 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. In terms of level, a \$100 mio. increase in size lowers the gross alpha by 0.2%.[18] These results provide strong support to models that emphasize the importance of capacity constraints for

[18]We have $\frac{\bar{b}}{\bar{\sigma}_q}100$ mio= $\frac{1.4}{6.5}$ = 0.2% per year, where $\bar{b}$ = 1.4% is the average size coefficient, and $\bar{\sigma}_q$ = \$650 mio. is the average volatility of fund size (i.e., time-series volatility averaged across funds). This number provides a lower bound for the average impact of a \$100 mio. increase because of the Jensen inequality, i.e., $E[\frac{b_i}{\sigma_{qi}}] > \frac{E[b_i]}{E[\sigma_{qi}]} = \frac{\bar{b}}{\bar{\sigma}_q}$.

mutual funds (e.g., BG, Pastor and Strambaugh (2012)).

Whereas the presence of funds with a negative size coefficient ($b_i < 0$) is not consistent with the BG model, the economic significance of this phenomenon is weak. These funds only represent 14.3% of the population and their size coefficient are close to zero (i.e., $\hat{Q}(5\%) = -0.6\%$ per year). For these specific funds, it could be the case that the cost structure $TC_{i,t}$ is more complex than in the BG model (e.g., fixed costs). In the appendix, we find that adding fixed costs (using $1/q_{i,t-1}$) and non-linear size effects (using $q^2_{i,t-1}$ and $q^3_{i,t-1}$) reduces the number of funds with negative $\hat{b}_i$ from 521 to 364 and 260, respectively. Therefore, alternative cost specifications help to reduce but do not eliminate all funds with negative $\hat{b}_i$.

Please insert Table II here

A.2 The Value Added

Next, we examine the value created by mutual funds—the lifecycle and steady state value added. For each fund, we estimate va^l_i and va^{ss}_i as a function of the two skill dimensions a_i and b_i (as per Equation (12)), and then use our nonparametric approach to infer the cross-sectional distributions $\phi(va^l)$ and $\phi(va^{ss})$. The summary statistics for both distributions are reported in Panel B Table II.

Overall, individual funds create significant value from their investment and trading decisions. Over their lifecycle, close to 60% of the funds produce a positive value added which, on average, reaches \$1.7 mio. per year—a number that is comparable to the value of \$2.0 mio. reported by BvB (their Table 7).[19] The total value created by mutual funds is even more striking at the steady state. Once funds reach their average size, the value added is positive for 70% of the population and is, on average, equal to \$7.3 mio. per year. Consistent with our previous analysis, va^{ss}_i is typically larger than va^l_i because b_i is typically positive (see Panel A).[20] The strong magnitude of this gap—\$5.2 mio. per year on average—further emphasizes the economic importance of capacity constraints.

The minority of funds with negative value added is comprised of two types: (i) "charlatans" that have no profitable investment ideas ($a_i < 0$), and (ii) funds that grow too large to maintain a positive alpha ($a_i < b_i q_{i,t-1}$). With a negative proportion for the fd alpha equal to 14%, we conclude that charlatans represent around 40% of the

[19] Their baseline average of 3.2 mio. per year is not comparable because it includes the fund diversification services (i.e., the benchmark returns are computed on a net basis (not gross)).

[20] We can interpret this result using the Jensen inequality. We can write $va^l_i = E[va_i(q_{i,t-1})]$ and $va^{ss}_i = va_i(E[q_{i,t-1}])$, where $va_i(q_{i,t-1}) = (a_i - b_i q_{i,t-1})q_{i,t-1}$. Because $va_i(q_{i,t-1})$ is a concave function, we have $E[va_i(q_{i,t-1})] < va_i(E[q_{i,t-1}])$.

funds with negative value added.[21] It is a priori surprising that the value added is negative given that funds have the option to invest passively. It could be the case that charlatans take active positions to hide their lack of skill (Berk and van Binsbergen (2019)). 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.

A.3 Impact of the EIV Bias

Accounting for the EIV bias adjustment is essential because it largely changes the shape of the cross-sectional distributions. First, it removes probability mass from the tails of the unadjusted density $\hat{\phi}(m)$. Intuitively, using the estimated skill measures $\hat{m}_i$ introduces noise and thus inflates the probability of observing extreme skill levels. Second, it induces a positive skewness because the correlation ρ between each skill measure and estimation variance is positive ($\hat{\rho} \approx 0.25$).[22] In other words, funds with better investment ideas (high a_i), tighter capacity constraints (high b_i), and larger value added (high va_i) tend to hold concentrated portfolios with higher volatility.

To quantify these adjustments, Table III and Figure 1 compare the bias adjusted and unadjusted distributions $\hat{\phi}^*(m)$ and $\hat{\phi}(m)$ for (i) the two skill dimensions (Panel A), and (ii) the value added (Panel B). 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 ($\hat{\pi}^- = 54.2\%$). 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, it can rise significantly for funds facing tight capacity constraints

[21] The average proportion of funds with negative value added $\hat{\pi}_{va}^-$ (lifecycle and steady state) is equal to 36.6% = (41.9+31.1)/2. Therefore, the proportion of charlatans is equal to $\hat{\pi}_a^- / \hat{\pi}_{va}^- = 40.0\%$ (14.0/36.6).

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

$(\hat{Q}(95\%) = 3.7\%)$.[23]

Please insert Table III and Figure 1 here

B A Deeper Look at Individual Fund Skill

B.1 The Variation in Skill across Funds

An important insight from Table II (Panel A) is that the two skill dimensions vary significantly across funds. For instance, we find that some funds exhibit stellar investment skills—5% of them exhibit a fd alpha above 8.1% per year, which is 2.6 times larger than the average. We also find that funds largely differ in their ability to mitigate capacity constraints as the volatility of b_i is as large as the average (1.4% per year). This finding is inconsistent with the commonly used panel regression approach which imposes a constant size coefficient b across all funds. It also explains why this approach provides weak statistical evidence of capacity constraints (as discussed by Pastor, Stambaugh, and Taylor (2015)). When b_i varies across funds, this variation inflates the standard deviation of the estimated $\hat{b}$ (Pesaran and Yagamata (2008)). As a result, $\hat{b}$ may not be statistically significant even if most funds are exposed to capacity constraints.

As discussed in Section II, the cross-sectional variation in skill is potentially driven by the specific strategies followed by funds. To examine this issue, we measure skill among funds with different portolio liquidity (small/large cap) and turnover (low/high turnover)—two key determinants of the fund strategy (Pastor, Stambaugh, and Taylor (2019)). The results of this analysis are summarized in Table IV and Figure 2.

Overall, the results confirm that the two skill dimensions vary across strategies. The average values of the fd alpha and size coefficient vary between 1.8% and 4.7% per year, and between 1.0% and 1.8% per year. At the same time, we still observe substantial cross-fund variation in skill within each group. This implies that some funds exhibit unique investment and trading skills that go beyond the specific strategies they follow.

We find that small cap funds largely dominate large cap funds along the first skill dimension, whereas the opposite holds for the second skill dimension. These results are consistent with the difference in liquidity between the two groups. Because small cap stocks are illiquid, they are likely to exhibit greater mispricing (higher a_i)—as noted by Hong, Lim, and Stein (2000), these stocks are largely untouched by mutual funds.[24]

[23]More generally, our analysis implies that any boxplot obtained with standard statistics softwares must be interpreted with caution if the variable of interest is an estimated quantity.

[24]Small cap stocks may also yield a higher fd alpha because have a higher idiosyncratic volatility and may therefore provide more opportunities for stock picking (e.g., Duan, Yu, and McLean (2009)).

At the same time, the cost of trading small cap stocks is higher which makes capacity constraints more severe (higher b_i).

We document a similar pattern for high versus low turnover funds. By rebalancing their portfolio more often, high turnover funds are able to exploit a larger number of profitable opportunities (higher a_i). However, they alo incur higher transaction costs (higher b_i). These findings clarify the relation between skill and turnover. Whereas previous studies argue that turnover signals superior skill (e.g., Chen, Jegadeesh, and Wermers (2000), Grinblatt and Titman (1989)), other studies favor the opposite interpretation (e.g., Carhart (1997), Elton et al. (1993)). Our results show that both conclusions hold provided that the right dimension of skill is examined—high turnover funds are skilled at generating profitable ideas, but unskilled at mitigating capacity constraints.

Please insert Table IV and Figure 2 here

B.2 The Correlation between the Two Skill Dimensions

Our cross-sectional analysis of skill uncovers another important result: the two skill dimensions are strongly correlated. In the entire population, the pairwise correlation between the estimated coefficients $\hat{a}_i$ and $\hat{b}_i$ reaches 0.82. This correlation is partly explained by the fund strategy—as shown in Table IV, liquidity and turnover change the two skill dimensions simultaneously. In other words, a_i and b_i are correlated because they both depend on the vector $d_i = (liq_i, turn_i)'$ that characterizes the fund strategy.

The implications of this positive correlation are twofold. First, the mutual fund industry is not heavily concentrated. We find that the top 5% of the funds only capture 20.8% of the total value added at the steady state.[25] This lack of concentration arises because few funds are skilled along the two dimensions, which limit their potential to create value. A simple calculation confirms this point—if a_i and b_i were uncorrelated, we would observe a high degree of concentration, i.e., the top 5% would capture 92.7% of the total value added.[26] This result contrasts with the high level of asset concentration observed among fund families (Berk, van Binsbergen, and Liu (2017)). One plausible explanation is that there are cost benefits in forming large families such as

[25] We denote the total value added for the population and the top 5% as $\hat{V} = n\hat{\mu}_{va}^{ss}$ and $\hat{V}_{top} = n \cdot 0.05\hat{E}(va_i^{ss} | va_i^{ss} > \hat{Q}(95\%))$, 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 $\hat{V}_{top}/\hat{V} = 22.8\%$.

[26] For each fund i, we simulate 10,000 values of a_i and b_i by drawing them independently from the vectors of estimated positive fd alphas and size coefficients. We then compute the value added by assuming that funds choose their level optimally such that $va_i = a_i^2/(4b_i)$ (as per Equation (19) below). Finally, we compute the ratio $\hat{V}_{top}/\hat{V}$ to obtain 92.7%.

shared resources and improved trading commissions (Chen et al. (2004)).

Second, it becomes essential to aggregate the two skill dimensions to determine which types of funds create more value. Because funds with more profitable ideas typically face tighter capacity constraints, it is a priori unclear whether they dominate funds able to scale up less profitable trading ideas. To examine this issue, Table V compares the value added (lifecycle and steady state) among the different fund groups. We find that the skill of small cap funds at identifying profitable trades more than compensates for their greater exposure to capacity constraints. At the steady state, these funds produce a positive value more often ($\hat{\pi}^+ = 80.9\%$ vs 62.4%) and its level is higher on average (\$7.6 mio. vs \$4.5 mio.). We also find that low turnover funds unambiguously dominate high turnover funds (both in terms of average and proportion). In short, small cap funds and low turnover funds both create more value, but rely on different skill dimensions—investment skills for the former versus trading skills for the latter.

Please insert Table V here

C Equilibrium Considerations

C.1 Do Skilled Funds Maximize their Value Added?

We now study the equilibrium implications of the BG model which result from the interaction between (i) a set of skilled funds in scarce supply and (ii) a large number of rational investors that compete for performance. Solving for the equilibrium yields several intuitive predictions about skill and value added that we can examine empirically using our nonparametric approach.

The first one is that mutual funds maximize the value added from exploiting their skills. 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 π_i under the constraint that the fees $f_{e,i}$ are equal to the gross alpha $\alpha_i = a_i - b_i q_i$ (so that investors break even). Maximizing profits is equivalent to maximizing the value added because $\pi_i = f_{e,i} q_i = \alpha_i q_i = va_i$. Replacing α_i with $a_i - b_i q_i$ and using the first order condition $\frac{\partial va_i}{\partial q_i} = 0$, we can write the optimal size as $q_i^* = \frac{a_i}{2b_i}$, and the optimal value added as $va_i^* = a_i q_i^* - b_i q_i^{*2} = \frac{a_i^2}{4b_i}$.

As a prelude to our main analysis, we estimate the cross-sectional distribution of the optimal value added, denoted by $\phi(va^*)$. The nonparametric estimation procedure described in Section III remains unchanged—the only difference pertains to the definition of the estimated measure $\hat{m}_i$ and its associated error term $u_{i,t}$:

$$\hat{m}_i = \widehat{va}_i^* = \frac{\hat{a}_i^2}{4\hat{b}_i}, \tag{19}$$

$$u_{i,t} = \frac{2a_i}{4b_i} e_1' Q_{x,i}^{-1} x_{i,t} \varepsilon_{i,t} - \frac{a_i^2}{4b_i^2} e_2' Q_{x,i}^{-1} x_{i,t} \varepsilon_{i,t}, \tag{20}$$

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 we focus on skilled funds for which the optimal value added is well defined (i.e., $\widehat{va}_i^*$ is positive). The summary statistics for $\phi(va^*)$ in Table VI (Panel A) reveal that the maximum rent that funds could earn reaches an average of \$14.1 mio. per year. The optimal value added also varies significantly, both in the population and across fund groups. For instance, we find that low turnover funds produce the combination of skill with the highest profit potential.

Next, we turn to the first prediction of the model by measuring the difference between the optimal value added va_i^* and its observed level, measured either with $va_{i,l}$ or $va_{i,ss}$. To do so, we estimate the distributions $\phi(va^* - va_l)$ and $\phi(va^* - va_{ss})$ using our nonparametric approach where $\hat{m}_i$ is given by

$$\hat{m}_i = \widehat{va}_i^* - \widehat{va}_{i,l} = \frac{\hat{a}_i^2}{4\hat{b}_i} - \hat{a}_i \overline{q}_{i,1} - \hat{b}_i \overline{q}_{i,2}, \tag{21}$$

$$\hat{m}_i = \widehat{va}_i^* - \widehat{va}_{i,ss} = \frac{\hat{a}_i^2}{4\hat{b}_i} - \hat{a}_i \overline{q}_{i,1} - \hat{b}_i \overline{q}_{i,1}^2, \tag{22}$$

and $u_{i,t}$ is set equal to the difference between the error term for $\widehat{va}_i^*$ in Equation (20) and the error term for $\widehat{va}_{i,l}$ or $\widehat{va}_{i,ss}$ in Equation (12).

Table VI (Panel B) reveals that the lifecycle value added is very far from its optimal level—the average difference between va_i^* and $va_{i,l}$ reaches \$13.5 mio. per year. An intuitive explanation for this large gap is the presence of learning effects. If 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 level at which the value added is maximized.

In this context, the analysis of the steady state value added is informative. Whereas the impact of learning can be large, it should weaken once funds reach their average size. This analysis provides strong support to the BG model. We find that the average difference between va_i^* and $va_{i,ss}$ drops to \$3.4 mio. per year. Therefore, funds extract more than 75% of the optimal profits once they reach their average size. The strong pairwise correlation of 0.94 between $\widehat{va}_{i,ss}$ and $\widehat{va}_i^*$ confirms that funds with higher skill potential do create more value.[27]

Our results further show that 79% of the funds have an average size above its optimal level (the average difference equals \$199 mio.). In the BG model, $\overline{q}_{i,1}$ being above q_i^* does not imply that funds fail to maximize the value added—if they simply invest the additional amount $\overline{q}_{i,1} - q_i^*$ passively, they can keep va_i^* unchanged. However, this is not the case—for these funds, $va_{i,ss}$ is \$3.4 mio. lower than va_i^* on average. This suggests that fund managers are also unsure of their skill levels and must learn about them alongside with investors.

Please insert Table VI here

C.2 Is the Gross Alpha a Noisy Measure of the Value Added?

The second prediction of the BG model is that the gross alpha α_i is a noisy measure of the value added. The model implies that the gross alpha must be equal to fees because investors break even. However, it does not pin down the level of fees—lower fees simply lead to a higher fund size without changing total profits $va_i^* = a_1 q_1^* - b_1 q_1^{*2} = \frac{a_i^2}{4b_i}$. Given that fees are arbitrary, the gross alpha can a priori take any value.

The information contained in the gross alpha depends on whether all funds coordinate on three specific fee setting policies.[28] Under Policy 1, the gross alpha is informative about the value added. In this case, all funds set fees such that the size is constant across all funds at $\bar{q}$, which yields $\alpha_i = \frac{a_1 q_1^* - b_1 q_1^{*2}}{\bar{q}} = \frac{1}{\bar{q}} \frac{a_i^2}{4b_i} \div va_i^*$. Under Policy 2, the gross alpha is informative about the first skill dimension (fd alpha). All funds set fees such that the size equals the optimal level q_i^*, which yields $\alpha_i = \frac{a_i q_i^* - b_i q_i^{*2}}{q_i^*} = \frac{a_i}{2} \div a_i$. Under Policy 3, the gross alpha is informative about the second skill dimension (size coefficient). Here, all funds set fees such that the size equals its squared optimal level q_i^{*2}, which yields $\alpha_i = \frac{a_i q_i^* - b_i q_i^{*2}}{q_i^{*2}} = b_i$. This analysis is summarized in Table VII.

Please insert Table VII here

To assess the information contained in the gross alpha, we compare funds with different levels of expenses (fees). For each group, we apply our nonparametric approach to estimate the cross-sectional distributions of each measure (gross alpha, value added,

[27] As noted by Berk, van Binsbergen, and Liu (2017), fund families can facilitate the convergence towards the optimal value added by reallocating capital across managers (via promotions and demotions).

[28] Here, we build on the analysis of BvB and extend it to include all four measures: the gross alpha, the value added, the fd alpha, and the size coefficient.

and the two skill dimensions). For the gross alpha, we define $\hat{m}_i$ and $u_{i,t}$ as:

$$\begin{aligned}
\hat{m}_i &= \hat{\alpha}_i = \hat{a}_i - \hat{b}_i \overline{q}_{i,1}, \\
u_{i,t} &= e_1' Q_x^{-1} x_t \varepsilon_{i,t} - E[q_{i,t-1}] e_2' Q_{x,i}^{-1} x_{i,t} \varepsilon_{i,t} - b_i (q_{i,t-1} - E[q_{i,t-1}]).
\end{aligned}$$

Consistent with the BG model, Table VIII (Panel A) shows that the gross alpha and the value added are essentially unrelated. On average, high expense funds exhibit higher gross alpha than low expense funds (1.2% vs 0.7% per year). However, they produce significantly lower value added on average (\$5.4 vs \$11.7 mio. for va_i^{ss}). In other words, funds do not set fees such that they all manage the same size (Policy 1)—instead, we observe a substantial variation in size across funds (the volatility equals \$2,049 mio.).

In contrast, Panel B reveals that the gross alpha is related to the fd alpha. On average, high expense funds have a significantly higher fd alpha (4.0% vs 2.2%). In addition, $\hat{\alpha}_i$ and $\hat{a}_i$ are positively correlated (0.49). These results suggest that some funds set fees so as to reach their optimal size q_i^* (Policy 2). At the same time, the moderate correlation between $\hat{\alpha}_i$ and $\hat{a}_i$ leaves room for alternative fee policies. For instance, Habib and Johnson (2016) note that some funds prefer to charge low fees and manage a large asset base to mitigate several institutional constraints.[29] Consistent with this analysis, we find that fund size is negatively correlated with fees (-0.23).

Please insert Table VIII here

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

The third prediction of the BG model focuses on the net alpha earned by investors defined as $\alpha_i^n = \alpha_i - f_{e,i}$. In the model, skilled funds have bargaining power because they are in scarce supply. Therefore, they extract all the rent from their skills, leaving investors with a zero net alpha.

To examine this prediction, we estimate the cross-sectional distribution of the net alpha $\phi(\alpha^n)$. Applying our nonparametric approach, we define $\hat{m}_i$ and $u_{i,t}$ as

$$\begin{aligned}
\hat{m}_i &= \hat{\alpha}_i^n = \hat{a}_i - \hat{b}_i \overline{q}_{i,1} - \bar{f}_{e,i}, && (23) \\
u_{i,t} &= e_1' Q_x^{-1} x_t \varepsilon_{i,t} - E[q_{i,t-1}] e_2' Q_{x,i}^{-1} x_{i,t} \varepsilon_{i,t} \\
&\quad - b_i (q_{i,t-1} - E[q_{i,t-1}]) - (f_{e,i,t} - \bar{f}_{e,i}), && (24)
\end{aligned}$$

[29]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.

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

We are not the first to estimate the entire net alpha distribution—recent studies have used standard parametric approaches to estimate $\phi(\alpha^n)$ (e.g., Chen, Cliff, and Zhao (2017), Harvey and Liu (2018a)). As discussed in Section III, using a nonparametric approach brings several advantages. For one, it is less prone to misspecification errors because it does not require to specify the shape of $\phi(\alpha^n)$. It also simpler and faster to apply because it does not require complex and computer-intensive algorithms such as Expectation Maximization (EM) methods.

Table IX shows that the fund alphas cluster around zero—the average is equal to -0.5% per year and 70% of the funds have net alphas between ±1.5% per year. However, several aspects of the alpha distribution are not consistent with the BG model. Consistent with the previous literature, we find ample evidence of negative performance ($\hat{\pi}^- = 66.3\%$). For some funds, the negative alphas are economically highly significant ($\hat{Q}(5\%) = -2.7\%$ per year). In addition, a third of the funds deliver positive alphas ($\hat{\pi}^+ = 34.7\%$)—a number that is significantly higher than the one obtained with the False Discovery Rate (Barras, Scaillet, and Wermers (2010)). This difference reflects the improved power of the nonparametric approach in detecting funds with alphas close to zero.

Overall, we need additional elements beyond the BG model to explain the shape of $\phi(\alpha^n)$. Positive alphas can be rationalized by the presence of search costs. If investors have to spend resources to detect skilled funds, Garleanu and Pedersen (2018) show that they need to be compensated for these costs in equilibrium. For negative alphas, a common view is that funds are able to exploit unsophisticated investors and charge them excessive fees (e.g., Christoffersen and Musto (2002), Gruber (1996)). The behaviour of these investors drives a positive wedge between fees and gross alphas that is left unexplained by any rational model.

Please insert Table IX here

D Overall Implications for the Mutual Fund Industry

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 allocation of resources to the most promising new companies. Individual funds may do so directly by participating to initial public offerings. More indirectly, they can improve the liquidity and efficiency of secondary markets, which is likely to reduce the cost of capital in the primary market (Cochrane (2013), Pedersen (2018)). 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, we find that mutual funds do a good job at maximizing their value added. In other words, they internalize the impact of capacity constraints on their portfolio returns. This implies that 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 (i) do trade regularly when companies issue new shares (IPOs, SEOs) and the composition of popular indices changes, (ii) do not hold the market because they concentrate on a subset of assets included in the indices they track. Therefore, even if we assume that mutual funds represent the only active investors, they can still beat the market by trading based on information and providing liquidity to passive investors (Pedersen (2018)).[30]

Finally, we provide strong evidence that the value added is positive and the net alpha is negative. We therefore conclude that individual funds are not only skilled—they are also in a strong bargaining position vis-a-vis investors. Our results show that this position may be stronger than initially thought. Whereas a common explanation for negative alphas is that funds overcharge relative to their skill, some of them actually destroy value (i.e., their value added is negative). Yet, these funds still charge fees to investors which fail to detect charlatans and/or incorporate the negative impact of capacity constraints.

[30]To illustrate this point, suppose that a set of passive investors track the Russell 2000. Because this index exhibits an annual turnover close to 50% per year, theses 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)).

E Additional Results

E.1 Alternative Asset Pricing Models

Our estimation of mutual fund skill possibly depends on the choice of the asset pricing 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 noticeable differences. First, the average fd alpha among small cap funds drops from 4.6% to 3.1% per year under the Carhart model, which is consistent with the analysis of CPZ. 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.

E.2 Fund Size and the Small Sample Bias

As noted by Pastor, Stambaugh, and Taylor (2015), the estimated size coffefficient $\hat{b}_i$ can potentially be biased because the return residual $\varepsilon_{i,t}$ is positively correlated with the change in size $\varepsilon_{qi,t}$, i.e., $\varepsilon_{i,t} = \phi_i \varepsilon_{qi,t} + \upsilon_{i,t}$. Whereas this bias vanishes asymptotically, it may have a significant impact for funds with a small sample size. To control for this bias, we use the approach of Amihud and Hurvich (2004) and add a proxy for $\varepsilon_{qi,t}$ to the set of regressors in Equation (4) (see the appendix for details). Theory predicts that the bias should be negative because $E[\hat{b}_i - b_i] = -E[\hat{\rho}_{qi} - \rho_{qi}]\phi > 0$ (Stambaugh (1999)). Consistent with this prediction, we find that the average size coefficient increases from 1.4% to 1.6% per year. Therefore the results strengthen the importance of capacity constraints for the mutual fund industry.

E.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: $\alpha_{i,t} = a_i - b_i q^{rel}_{i,t-1}$, where $q^{rel}_{i,t-1}$ is 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 relation between the gross alpha and size 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 constraints: $\alpha_{i,t} = a_i - b_i q_{t-1}$, where q_{t-1} is defined as the ratio of the industry size on the total market capitalization. Whereas we confirm the result of Pastor, Stambaugh, and

Taylor (2015), 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).

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 68% of them have a positive coefficient c_i. Controling for age also leaves the proportions of funds with positive fd alpha and size coefficient largely unchanged.

E.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 ($\hat{\pi}^+ = 85.8\%$). 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. The empirical evidence suggests that a_i and b_i are strongly correlated. Therefore, investors in their model possibly 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 cross-sectional distributions of skill and its economic value. Our approach is nonparametric and thus particularly suited to the analysis of skill. It avoids the challenge of correctly specifying each distribution, and allows us to jointly examine multiple measures, including two skill dimensions (fd alpha and size coefficient), and two measures of the value added (lifecyle and steady state value added). In addition to its flexibility, our approach is simple to implement, applicable to the different characterizations of each distribution (e.g., moments, quantiles), and supported by econometric theory.

Our 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, the industry is not heavily concentrated because the two skill dimensions are positively correlated. In other words, it is difficult for funds to have both investment and trading skills. Third, it does a good job at maximizing the rent from exploiting skill. This implies the size of the industry is smaller than in a fully competitive equilibrium as funds internalize the impact of capacity constraints. Finally, the industry has a strong bargaining power vis-a-vis investors because we find the value created by funds is high, whereas the performance received by investors is low.

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 for capturing the heterogeneity across stocks (risk exposure, commonality in liquidity), or in corporate finance for capturing the heterogeneity across firms (investment and financing decisions), and, more recently, in household finance for capturing the heterogeneity across households (time preference, risk aversion; see Calvet et al. (2019)).

References

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

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

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

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

[5] Baks, K. P., A. Metrick, and J. Wachter, 2001, Should Investors Avoid all Actively Managed Mutual Funds? A Study in Bayesian Performance Evaluation, Journal of Finance 56, 45-85.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

[20] Christoffersen, S.E. K., and D. K. Musto, 2002, Demand Curves and the Pricing of Money Management, Review of Financial Studies 15, 1495-1524.

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

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

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

[24] Duan, Y., G. Hu, and R.D. McLean, 2009, When is Stock Picking Likely to be Successful? Evidence from Mutual Funds, Financial Analysts Journal 65, 1-11.

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

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

[27] Fang, J., A. Kempf, and M. Trapp, 2014, Fund Manager Allocation, Journal of Financial Economics 111, 661-674.

[28] Ferson, W., and Y. Chen. 2019. How Many Good and Bad Fund Managers are There, Really? in C.F. Lee, editor, the Handbook of Financial Economics, Mathematics, Statistics and Technology, ch. 108.

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

[30] Garleanu, N., and L. Pedersen, 2018, Efficiently Inefficient Markets for Assets and Asset Management, Journal of Finance 73, 1663-1712.

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

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

[33] Grinblatt, M., and S. Titman, 1989, Mutual Fund Performance: An Analysis of Quarterly Portfolio Holdings, Journal of Business 62, 394–415.

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

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

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

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

[38] Harvey, C., and Y. Liu, 2018a, Detecting Repeatable Performance, Review of Financial Studies, Forthcoming.

[39] Harvey, C., and Y. Liu, 2018b, Decreasing Returns to Scale, Fund Flows, and Performance, Working Paper.

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

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

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

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

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

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

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

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

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

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

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

[51] Pastor, L., R. F. Stambaugh, and L. Taylor, 2019, Fund Tradeoffs, Working Paper.

[52] Pedersen, L.H, 2015. Efficiently Inefficient. Princeton University.

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

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

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

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

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

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

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

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

[61] Stanton, R., 1997, A Nonparametric Model of Term Structure Dynamics and the Market Price of Interest Rate Risk, Journal of Finance 52, 1973-2002.

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

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

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

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

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

Panel B: Estimated Betas

	Market	Size	Value	Momentum	Adj. R2
All Funds	0.93	0.25	-0.11	0.01	0.98
Investment Styles					
Small-cap	0.98	0.78	-0.22	0.06	0.97
Large-cap	0.95	0.14	-0.06	0.01	0.99
Growth	0.95	0.34	-0.35	0.03	0.97
Value	0.91	0.12	0.22	-0.01	0.98
Fund Characteristics					
Low Expense	0.93	0.19	-0.05	0.01	0.98
High Expense	0.94	0.42	-0.27	0.02	0.97
Low Turnover	0.93	0.23	-0.07	-0.01	0.97
High Turnover	0.95	0.39	-0.31	0.11	0.95

Table II
Cross-Sectional Distribution of the Two Skill Dimensions

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

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

Panel B: Size Coefficient

	Moments				Proportions (%)		Quantiles (Ann.)	
	Mean (Ann.)	Volatility (Ann.)	Skewness	Kurtosis	Negative	Positive	5%	95%
All Funds	1.4	1.4	2.2	4.8	14.1	85.9	-0.6	3.7
Investment Styles								
Small Cap	1.7	1.7	1.8	5.8	13.5	86.5	-0.9	4.4
Large Cap	1.0	1.0	2.8	11.9	18.0	82.0	-0.5	2.6
Growth	1.5	1.7	2.4	4.7	17.1	82.9	-1.0	4.1
Value	1.4	1.4	1.5	3.4	13.9	86.1	-0.7	3.7
Fund Characteristics								
Low Expense	0.8	1.1	0.8	1.6	23.0	77.0	-0.8	2.6
High Expense	1.6	1.9	1.5	1.4	17.9	82.1	-1.1	4.6
Low Turnover	1.0	1.7	1.6	3.5	27.7	72.3	-1.5	3.5
High Turnover	1.5	2.3	2.0	8.6	22.0	78.0	-1.5	4.7

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

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

Panel B: Steady State Value Added

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

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

	Moments				Proportions (%)		Quantiles (Ann.)	
	Mean (Ann.)	Volatility (Ann.)	Skewness	Kurtosis	Negative	Positive	5%	95%
With Bias Adjustement	3.1	3.4	1.6	4.4	14.0	86.0	-1.3	8.2
Without Bias Adjustement	3.1	5.0	0.5	7.1	21.4	78.6	-4.0	11.3
Difference	0.0	-1.6	1.2	-2.8	-7.4	7.4	2.7	-3.1

Panel B Size Coefficient

	Moments				Proportions (%)		Quantiles (Ann.)	
	Mean (Ann.)	Volatility (Ann.)	Skewness	Kurtosis	Negative	Positive	5%	95%
With Bias Adjustement	1.4	1.4	2.2	4.8	14.1	85.9	-0.6	3.7
Without Bias Adjustement	1.4	2.2	0.7	7.2	22.5	77.5	-1.6	5.1
Difference	0.0	-0.8	1.5	-2.3	-8.3	8.3	1.0	-1.5

Panel C: Lifecycle Value Added

	Moments				Proportions (%)		Quantiles (Ann.)	
	Mean (Ann.)	Volatility (Ann.)	Skewness	Kurtosis	Negative	Positive	5%	95%
With Bias Adjustement	1.7	10.6	5.8	47.8	41.9	58.1	-4.3	13.9
Without Bias Adjustement	1.7	17.9	3.5	35.4	54.2	45.8	-14.7	26.3
Difference	0.0	-7.3	2.4	12.4	-12.3	12.3	10.4	-12.4

Panel D: Steady State Value Added

	Moments				Proportions (%)		Quantiles (Ann.)	
	Mean (Ann.)	Volatility (Ann.)	Skewness	Kurtosis	Negative	Positive	5%	95%
With Bias Adjustement	7.3	18.6	5.7	42.1	31.2	68.8	-3.8	34.5
Without Bias Adjustement	7.3	22.2	5.0	37.9	36.3	63.7	-6.4	41.1
Difference	0.0	-3.7	0.7	4.1	-5.1	5.1	2.6	-6.6

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

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

Panel B: Difference with Lifecycle and Steady State Value Added

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

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

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

Panel B: Summary Statistics for Fund Fees and Size

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

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.

	Moments				Proportions (%)		Quantiles (Ann.)	
	Mean (Ann.)	Volatility (Ann.)	Skewness	Kurtosis	Negative	Positive	5%	95%
All Funds	0.7	1.5	0.3	8.5	31.6	68.4	-1.4	3.0
Investment Styles								
Small-cap	1.7	1.8	1.9	2.1	15.8	84.2	-0.9	4.5
Large-cap	0.2	0.9	0.5	4.4	43.2	56.8	-1.2	1.6
Growth	0.6	1.9	-0.4	8.8	37.7	62.3	-2.0	3.2
Value	1.2	1.6	1.6	4.8	22.0	78.0	-1.1	3.8
Fund Characteristics								
Low Expense	0.5	1.4	1.0	2.0	38.5	61.5	-1.6	2.7
High Expense	1.2	2.4	0.3	7.1	28.4	71.6	-1.8	4.6
Low Turnover	0.9	1.7	1.0	3.0	27.9	72.1	-1.5	3.6
High Turnover	1.0	2.3	1.1	5.0	34.5	65.5	-2.0	4.4

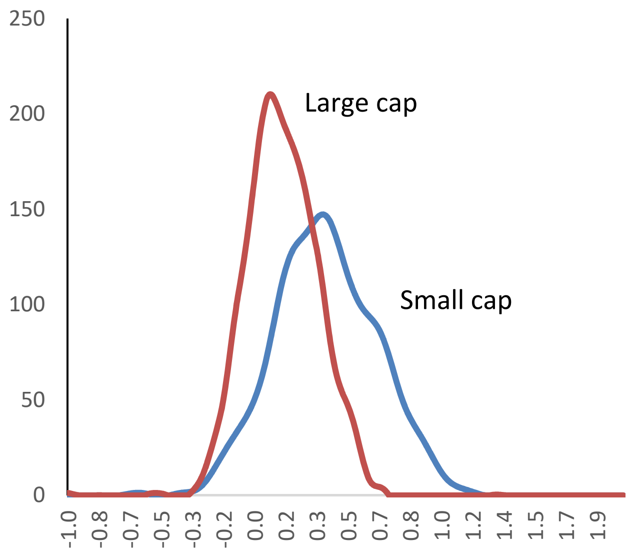
Table VIII
Cross-Sectional Distribution of Net Alpha

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.

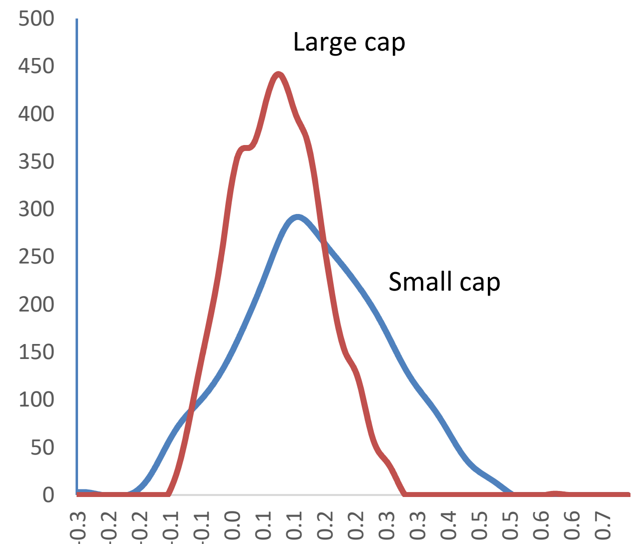
	Moments				Proportions (%)		Quantiles (Ann.)	
	Mean (Ann.)	Volatility (Ann.)	Skewness	Kurtosis	Negative	Positive	5%	95%
All Funds	-0.5	1.6	-1.0	9.7	65.3	34.7	-2.7	1.7
Investment Styles								
Small-cap	0.3	1.7	0.9	4.2	45.2	54.8	-2.2	3.1
Large-cap	-1.0	0.9	-1.5	4.3	85.6	14.4	-2.5	0.5
Growth	-0.7	1.9	-1.6	9.9	67.2	32.8	-3.4	1.9
Value	0.0	1.5	0.4	4.8	49.5	50.5	-2.2	2.4
Fund Characteristics								
Low Expense	-0.4	1.4	0.2	1.1	62.6	37.4	-2.4	1.8
High Expense	-0.5	2.4	-0.9	10.0	60.5	39.5	-3.6	3.0
Low Turnover	-0.2	1.6	0.0	3.6	58.7	41.3	-2.6	2.3
High Turnover	-0.4	2.3	-0.1	5.5	59.7	40.3	-3.5	3.1

Figure 1
Cross-sectional Distributions of the Two Skill Dimensions: Analysis across Fund Groups

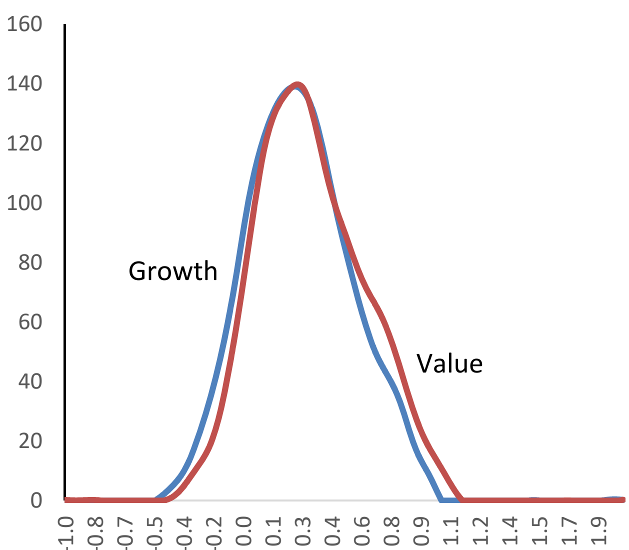
Panel A plots the cross-sectional densities of the first dollar alpha for small cap and large cap funds. Panel B compares growth and value funds. Panel C compares low expense and high expense funds. Finally, Panel D compares low turnover and high turnover funds. All the estimated densities are adjusted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



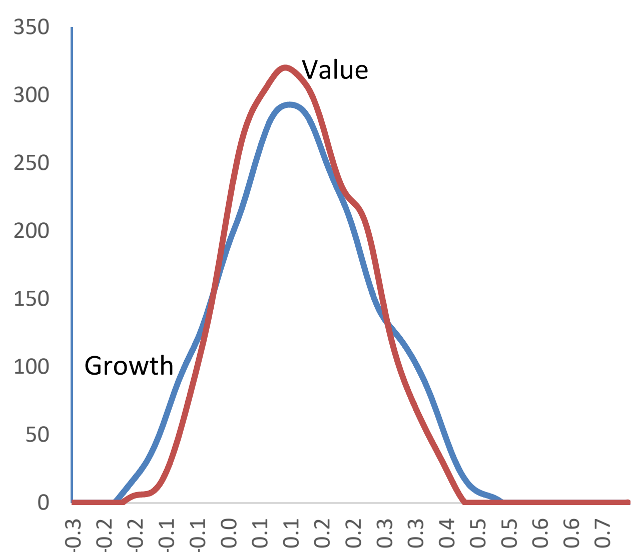
Panel A: Size - Fd Alpha



Panel B: Size - Size Coefficient



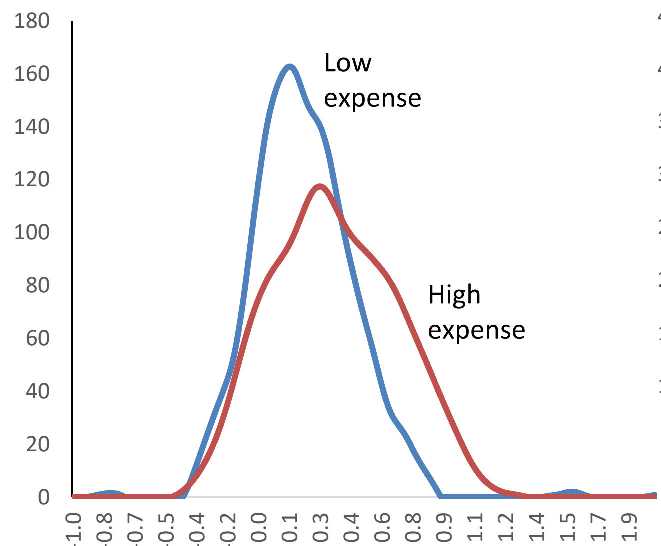
Panel C: Growth/Value – Fd Alpha



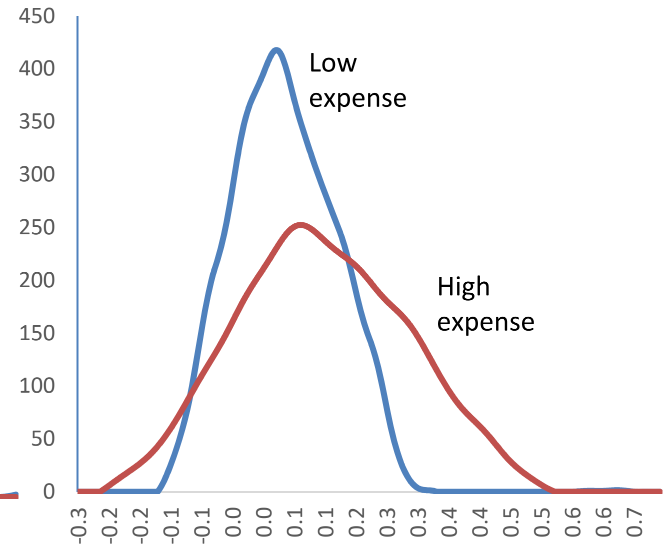
Panel D: Growth/Value – Size Coefficient

Figure 1
Cross-sectional Distributions of the Two Skill Dimensions:
Analysis across Fund Groups (Continued)

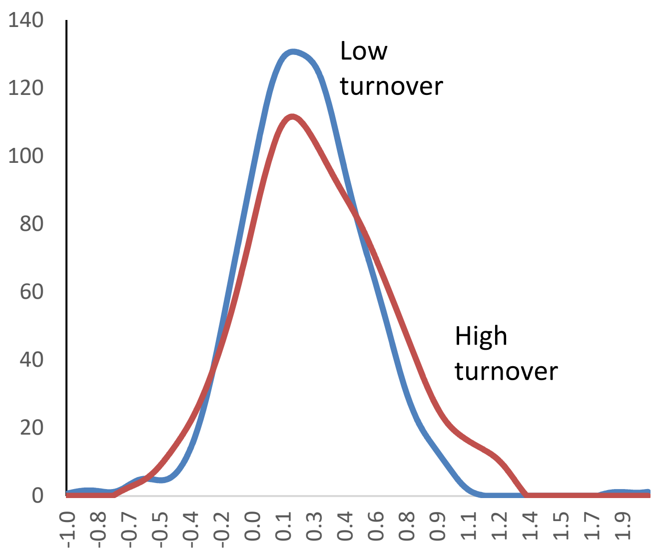
Panel A plots the cross-sectional densities of the size coefficient for small cap and large cap funds. Panel B compares growth and value funds. Panel C compares low expense and high expense funds. Finally, Panel D compares low turnover and high turnover funds. All the estimated densities are adjusted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



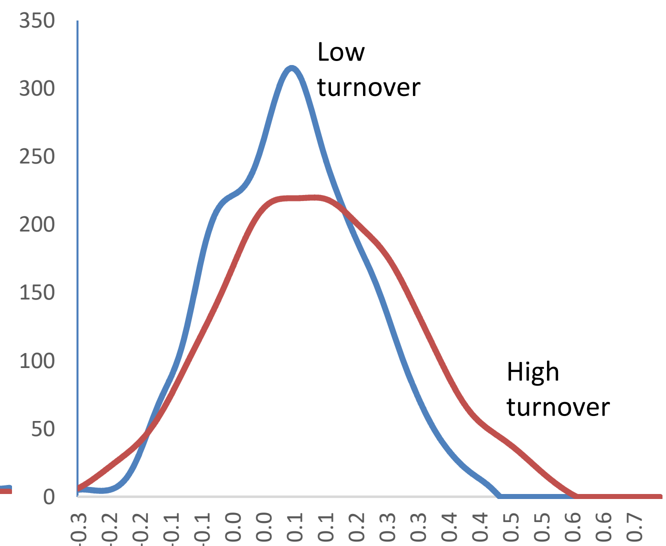
Panel E: Expenses – Fd Alpha



Panel F: Expenses – Size Coefficient



Panel G: Turnover – Fd Alpha



Panel H: Turnover – Size Coefficient