JOURNAL OF FINANCIAL AND QUANTITATIVE ANALYSIS Vol. 52, No. 4, Aug. 2017, pp. 1577–1604
COPYRIGHT 2017, MICHAEL G. FOSTER SCHOOL OF BUSINESS, UNIVERSITY OF WASHINGTON, SEATTLE, WA 98195
doi:10.1017/S002210901700045X

https://doi.org/10.1017/S002210901700045X Published online by Cambridge University Press

Mutual Fund Performance Evaluation and Best Clienteles

Stéphane Chrétien and Manel Kammoun*

Abstract

This paper investigates investor disagreement and clientele effects in performance evaluation by developing a measure that considers the best potential clienteles of mutual funds. In an incomplete market under law-of-one-price (LOP) and no-good-deal conditions, we obtain an upper bound on admissible performance measures that identifies the most favorable alpha. Empirically, we find that a reasonable investor disagreement leads to generally positive performance for the best clienteles. Performance disagreement by investors can be significant enough to change the average evaluation of mutual funds from negative to positive, depending on the clienteles.

I. Introduction

In today's mutual fund industry, there are thousands of funds that cater to different investors through their management style and other attributes. In incomplete markets, as investors can disagree about the attractiveness of funds, this catering might be worthwhile in leading the funds to find appropriate clienteles (i.e., the class of investors to whom they are the most valuable).

Recent research examines the effect of investor disagreement and heterogeneity on mutual fund performance evaluation. Studying the issue generally, Ferson and Lin (2014) find that taking into account heterogeneous preferences can lead to large valuation disagreement. In particular, they develop a bound on expected disagreement with a traditional alpha and show that such disagreement

*Chrétien (corresponding author), stephane.chretien@fsa.ulaval.ca, Faculty of Business Administration, Laval University; Kammoun, manel.kammoun@uqo.ca, Department of Administrative Sciences, Université du Québec en Outaouais. We thank Vikas Agarwal, Laurent Barras, Marie-Claude Beaulieu, Stephen Brown (the editor), Frank Coggins, Wayne Ferson, Juha Joenväärä, Laurenz Klipper, Jerchern Lin (the referee), Gabriel Power, and participants at the 2014 Mathematical Finance Days, the 2014 European Financial Management Association Annual Conference, the 2014 World Finance Conference, the 2014 Financial Management Association Annual Meeting, and seminars at Laval University, Ryerson University, University of Quebec at Montreal, and Université du Québec en Outaouais for helpful comments. We gratefully acknowledge financial support from the Faculty of Business Administration and Investors Group Chair in Financial Planning at Laval University and the Institut de Finance Mathématique de Montréal.

https://doi.org/10.1017/S002210901700045X Published online by Cambridge University Press

can be similar in importance to the widely documented effects of the benchmark choice problem and the statistical imprecision in estimates of alpha. They furthermore provide evidence that investor disagreement and heterogeneity are economically significant in the behavior of fund investors.

Many studies concentrate more precisely on identifying specific clienteles. Glode (2011) argues that mutual funds could be valuable to investors with high marginal utilities in difficult times by providing positive alphas in recessions. Bailey, Kumar, and Ng (2011) document that behavioral biases are factors of investor heterogeneity in the mutual fund industry. Del Guercio and Reuter (2014) find that the retail mutual fund market is formed from two broad clienteles that value funds differently: self-directed investors and investors acting with the help of brokers. In a literature review, Ferson (2010) emphasizes that measuring performance from the point of view of different clienteles is a challenge for future research.

Despite these contributions, the literature has not focused on the valuation that can be the most important for mutual funds, the one from their best potential clienteles. The goal of this paper is to provide additional evidence on investor disagreement and clientele effects in performance evaluation by developing and implementing a measure that considers the best potential clienteles of mutual funds. For our purpose, these clienteles are defined as the class of investors most favorable to a fund in the sense that they value the fund at an upper performance bound in a setup where the market is incomplete. Our "best clientele performance measure" thus not only considers investor disagreement but also focuses on the worthiest clienteles that a mutual fund could target.

We develop this new measure by combining the asset pricing bound literature with the stochastic discount factor (SDF) performance evaluation approach first proposed by Glosten and Jagannathan (1994) and Chen and Knez (1996). Cochrane and Saá-Requejo (2000) propose asset pricing bounds in an incomplete market under the law-of-one-price (LOP) condition of Hansen and Jagannathan (1991) and a no-good-deal condition that rules out investment opportunities with unreasonably high Sharpe ratios, termed "good deals." We obtain the best clientele alpha by adapting this approach to performance measurement and focusing on the upper bound.

Using Hansen's (1982) generalized method of moments (GMM), we estimate the best clientele SDF alphas with monthly returns of 2,786 actively managed U.S. open-ended equity mutual funds from Jan. 1984 to Dec. 2012. Our main results rely on a set of passive portfolios based on 10 industry portfolios, although they are similar for style portfolios or the market portfolio. Following the literature, we consider the disagreement generated by allowing a maximum Sharpe ratio (specifying the no-good-deal restriction) equal to the best Sharpe ratio from the passive portfolios plus either half the Sharpe ratio of the market index or its full value.

Empirically, we find that considering investor disagreement and focusing on the best potential clienteles lead to a generally positive performance for mutual funds. For example, with a disagreement that corresponds to an increase in admissible opportunities equivalent to half the market Sharpe ratio, the mean monthly best clientele alpha is equal to 0.236% (t-stat. $= 3.35$). Comparatively, the mean alpha when disagreement is ruled out is -0.179% (t-stat. $= -3.14$), a value similar

https://doi.org/10.1017/S002210901700045X Published online by Cambridge University Press

to the findings from standard measures based on representative investors. The spread of 0.415% between these values is comparable to the magnitude of the bound on the average disagreement documented by Ferson and Lin (2014). Accordingly, the proportions of positive and significantly positive alpha estimates increase from 20% to 78% and from 1% to 24%, respectively, when allowing for disagreement. To account for false discoveries, we implement the technique of Barras, Scaillet, and Wermers (BSW) (2010) and its extension proposed by Ferson and Chen (FC) (2015). For the best clienteles, we find that the proportions of skilled funds increase considerably, and the proportions of unskilled funds disappear.

Further augmenting the maximum Sharpe ratio improves even more the performance of the mutual funds for their best clienteles, and the results are robust to the use of simulated finite sample distributions for inference purposes. In addition, we show that increasing monthly Sharpe ratio opportunities by only 0.04 (approximately one third of the market Sharpe ratio) is sufficient for the best clienteles to give a nonnegative performance to mutual funds, on average. Hence, although the maximum Sharpe ratio is a subjective choice, our findings indicate that a reasonably small disagreement among investors is enough to generate a positive performance from the best potential clienteles of a majority of funds.

We also implement a conditional version of the best clientele measure and find that the inclusion of conditioning information does not alter our findings. Consistent with Moskowitz (2000), Kosowski (2011), and Glode (2011), we find evidence that the best clientele alpha estimates are more positive in recessions than in expansions. Finally, to further explore the issues of investor disagreement and clientele effects, we estimate a lower performance bound to obtain a "worst clientele alpha" and investigate the total performance disagreement resulting from the difference between best and worst clientele alphas. We find that the evidence from the worst clienteles is more negative than the one from standard measures and that total disagreement is economically and statistically significant.

Overall, the positive best clientele alpha estimates documented in this paper have two important implications for the mutual fund literature. First, as shown by Chen and Knez (1996) and Ferson and Lin (2014), a positive SDF alpha is required for the existence of investors who would want to buy a fund. However, a large number of studies document negative value added for actively managed mutual funds (see Fama and French (2010) and BSW (2010) for recent examples), with the growth of the industry termed a "puzzle" by Gruber (1996). In line with the continued popularity of these funds in practice, this paper empirically shows that a positive alpha exists for some clienteles for most funds.

Second, although the effects of the benchmark choice problem and the statistical imprecision in estimates of alpha can change the performance assessment of a fund, these issues do not generally change the conclusion regarding the negative value added for the mutual fund industry. This paper shows that investor disagreement can be significant enough to change the average evaluation of mutual funds from negative to positive, depending on the clienteles. It thus reinforces the analysis of Ferson and Lin (2014) on the economic importance of investor disagreement and clientele effects.

The remainder of this paper is organized as follows: Section II develops the best clientele performance measure. Section III offers our methodology for estimating performance values and summarizing the results. Section IV describes the mutual fund data and the passive portfolio returns. Section V provides and interprets our empirical results. Finally, Section VI presents conclusions.

II. Performance Measure for the Best Potential Clienteles

A. Basic Performance Setup

We measure the performance, or alpha, with the SDF approach:

$$\alpha_{\mathrm{MF},t} = \mathrm{E}_t[m_{t+1} R_{\mathrm{MF},t+1}] - 1, \tag{1}$$

where m_{t+1} is the SDF of an investor interested in valuing the mutual fund with gross return $R_{\mathrm{MF},t+1}$, and the expectation operator $\mathrm{E}_t[\,]$ is understood to be conditional on the investor's or the public information set. Taking the unconditional expectation on both sides, and dropping time subscripts except when needed to avoid ambiguity, the expected alpha is given by

$$\alpha_{\mathrm{MF}} = \mathrm{E}[m\, R_{\mathrm{MF}}] - 1. \tag{2}$$

Glosten and Jagannathan (1994) and Chen and Knez (1996) were the first to propose SDF alphas for performance evaluation. The SDF approach does not require any assumptions about complete markets, utility functions, or aggregation. In the context of performance evaluation, Ferson (2010) argues that it is general enough to properly account for heterogeneous investors and differentially informed managers. In contrast, for a traditional regression-based alpha, a positive (negative) value does not necessarily imply buying (selling) the fund, and a manager with superior information does not necessarily generate a positive value. Ferson ((2010), p. 227) thus concludes that "the SDF alpha seems to be on the most solid theoretical footing, and should probably get more attention than it has in the literature."

Unfortunately, the SDFs of different investors are not observable. Hence, the literature typically uses the SDF from a representative investor obtained through economic assumptions and equilibrium conditions. This approach provides a unique performance evaluation that can be relevant for all investors. However, it rules out investor disagreement that occurs when one client views the performance of a fund differently from another client (e.g., Ferson and Lin (2014)). Furthermore, it exposes the results to the benchmark choice problem, because the selected performance model does not necessarily price correctly passive portfolios (e.g., Chen and Knez (1996), Ahn, Cao, and Chrétien (2009), and Cremers, Petajisto, and Zitzewitz (2013)).

Instead, in this paper, we impose an economically relevant structure on the set of SDFs of all investors to obtain a restricted set useful to identifying the most favorable performance. Let M represent this restricted set. Under the assumption that it is constrained enough to be closed and convex, Chen and Knez (1996) and Ahn et al. (2009) demonstrate that it is possible to find an upper bound on the

https://doi.org/10.1017/S002210901700045X Published online by Cambridge University Press

performance of a fund:

$$\bar{\alpha}_{\mathrm{MF}} = \sup_{m \in M} \mathrm{E}[m\, R_{\mathrm{MF}}] - 1, \tag{3}$$

where $\bar{\alpha}_{\mathrm{MF}}$ represents the upper bound on the expected alpha, the highest average performance value that can be found from the heterogeneous investors considered in M.

By considering a set of SDFs, as opposed to selecting a unique SDF, our setup results in a finite range of performance values (e.g., Chen and Knez (1996), Ahn et al. (2009)). Hence, it allows for investor disagreement that can occur in incomplete markets. As argued by Chen and Knez ((1996), p. 529), "given that mutual funds are set up to satisfy different clienteles, such an evaluation outcome may not be unrealistic." Empirically, Ferson and Lin (2014) find that disagreement is in fact economically important. Although we also examine the lower bound, we focus on the upper bound because it can be interpreted as the performance from the class of investors most favorable to the fund (in a valuation sense). We can thus evaluate whether funds add value from the perspective of their best potential clienteles. In particular, Chen and Knez (1996) and Ferson and Lin (2014) show that if this value is positive, there are some investors who would want to buy the funds, with an optimal investment proportional to the alpha.

B. Restricting the SDFs

One key to our approach is restricting the set of SDFs in an economically meaningful way. We impose two conditions that the SDFs of mutual fund investors should meet. The first condition is the LOP condition, as discussed by Hansen and Jagannathan (1991). The SDFs used for performance measurement should price correctly passive portfolios or basis assets:

$$\mathrm{E}[m\, \mathbf{R_K}] - \mathbf{1} = 0, \tag{4}$$

where $\mathbf{R_K}$ is a vector of gross returns on K passive portfolios, and $\mathbf{1}$ is a $K \times 1$ unit vector. It is plausible to assume that investors would agree that passively managed portfolios should have zero average alphas. The main benefit of imposing the LOP condition is to alleviate the previously mentioned benchmark choice problem. To see how this condition restricts the set of SDFs, Hansen and Jagannathan (1991) show that it implies a minimum SDF standard deviation that is related to the highest Sharpe ratio attainable in the passive portfolios.[1]

Although the LOP condition provides important restrictions on the set of SDFs, it is not sufficient to make it closed and convex and thus would allow an infinite range of performance values, as discussed in Chen and Knez (1996). The second condition we impose is the no-good-deal condition of Cochrane and Saá-Requejo (2000). Specifically, the SDFs should not allow investment opportunities with Sharpe ratios that are too high:

$$\frac{\mathrm{E}[R_j - R_F]}{\sigma[R_j - R_F]} < \bar{h}, \tag{5}$$

[1]Other restrictions on SDFs that can be developed from this condition include, for example, the bounds of Snow (1991) on selected higher SDF moments and the bound of Chrétien (2012) on the SDF autocorrelation.

https://doi.org/10.1017/S002210901700045X Published online by Cambridge University Press

where R_j is the return on any potential asset j, $\bar{h}$ is the maximum Sharpe ratio allowable, and R_F is the risk-free rate. We thus stipulate that investors would find it implausible that investment opportunities could provide Sharpe ratios that are too high, making them good deals. If available, these deals would be unlikely to survive because investors would quickly grab them up.

The reasons why high Sharpe ratios should be ruled out are discussed by Ross (1976), MacKinlay (1995), Cochrane and Saá-Requejo (2000), and Ross (2005), among others. Ross (1976) argues that high Sharpe ratios are unreasonable from the perspective of the capital asset pricing model (CAPM) and thus rules them out in studying deviations from the arbitrage pricing theory. In developing a specification test for multifactor models, MacKinlay (1995) uses a bound on the maximum Sharpe ratio, arguing that high ratios are unlikely from the perspective of risk-based models. Cochrane and Saá-Requejo (2000) argue that implausibly high Sharpe ratio opportunities should be rapidly exploited by investors. Unless there are limits to exploiting them, their presence would imply implausibly high investor risk aversion. A similar argument is formalized in Ross (2005).

By extension of the work of Hansen and Jagannathan (1991), Cochrane and Saá-Requejo (2000) show that the no-good-deal condition restricts the set of SDFs by limiting its second moment:

$$\mathrm{E}[m^2] \quad \leq \quad \frac{(1+\bar{h}^2)}{R_F^2}. \tag{6}$$

Put differently, the assumption that investors would necessarily want to take part at the margin in implausibly good deals implies that their SDF satisfies the restriction $\sigma[m] \leq \bar{h}/R_F$. As shown by Cochrane and Saá-Requejo (2000), this restriction and the LOP condition make the set of SDFs closed and convex and thus allow the existence of price bounds.

There are other restrictions that could be imposed on SDFs to result in price bounds.[2] In this paper, we select the no-good-deal condition for the following reasons. First, since being introduced by Sharpe (1966), the Sharpe ratio has a long history of relevancy in performance evaluation. Due to its simplicity and intuitive appeal, it is a commonly used measure both in practice and in academic studies. Second, the literature offers some guidance on the choice of the maximum Sharpe ratio. In contrast, there is little guidance for the maximum gain–loss ratio, and there are often not enough restrictions imposed by the no-arbitrage (NA) condition (so NA bounds are typically wide). Third, the Sharpe ratio captures approximate arbitrage opportunities as well as the gain–loss ratio when returns are normally distributed, which is reasonable for our sample of monthly equity mutual fund returns.[3] Fourth, the no-good-deal framework offers a closed-form solution for the

[2]For example, Hansen and Jagannathan (1991) discuss the NA condition that excludes nonpositive SDFs by ruling out arbitrage opportunities. Chen and Knez (1996) show that it is sufficient to obtain a finite range of performance values, and Ahn et al. (2009) study NA performance bounds for mutual funds. As an alternative, Bernardo and Ledoit (2000) introduce a maximum gain–loss ratio condition that rules out approximate arbitrage opportunities. They show that it leads to a restriction on the minimum and maximum SDF values.

[3]Koski and Pontiff (1999) find that most equity mutual funds do not use derivatives and that the distributional characteristics of users and nonusers are similar, with monthly skewness and excess kurtosis relatively close to 0.

https://doi.org/10.1017/S002210901700045X Published online by Cambridge University Press

performance bounds. This solution facilitates and accelerates its implementation in comparison to numerical-only solutions obtained when imposing the NA condition or the maximum gain–loss ratio condition. This advantage should not be neglected given that our large-scale empirical investigation considers thousands of mutual funds.[4]

C. Best Clientele Performance Measure

Considering our basic performance setup and our restrictions on the set of SDFs, the upper bound on performance evaluation can be found by solving the following problem:

$$\bar{\alpha}_{\mathrm{MF}} = \sup_{m \in M} \mathrm{E}[m\, R_{\mathrm{MF}}] - 1, \tag{7}$$

$$\text{subject to} \quad \mathrm{E}[m\, \mathbf{R_K}] = \mathbf{1},\ \mathrm{E}[m^2] \leq \frac{(1+\bar{h}^2)}{R_F^2}. \tag{8}$$

Cochrane and Saá-Requejo (2000) show that this problem has the following solution:

$$\bar{\alpha}_{\mathrm{MF}} = \mathrm{E}[\bar{m}\, R_{\mathrm{MF}}] - 1, \tag{9}$$

with

$$\bar{m} = m^* + v w, \tag{10}$$

$$m^* = \mathbf{a}' \mathbf{R_K}, \tag{11}$$

$$w = R_{\mathrm{MF}} - \mathbf{c}' \mathbf{R_K}, \tag{12}$$

where

$$\mathbf{a}' = \mathbf{1}' \mathrm{E}[\mathbf{R_K} \mathbf{R_K}']^{-1}, \tag{13}$$

$$\mathbf{c}' = \mathrm{E}[R_{\mathrm{MF}} \mathbf{R_K}'] \mathrm{E}[\mathbf{R_K} \mathbf{R_K}']^{-1}, \tag{14}$$

$$v = \sqrt{\frac{\left(\frac{(1+\bar{h}^2)}{R_F^2} - \mathrm{E}[m^{*2}] \right)}{\mathrm{E}[w^2]}}. \tag{15}$$

We call the solution $\bar{\alpha}_{\mathrm{MF}}$ the "best clientele alpha" to refer to our earlier discussion; it indicates whether mutual funds add value from the perspective of their best potential clienteles, the class of investors most favorable to the fund. Similarly, $\bar{m}$ represents the "best clientele SDF." In this solution, m^* is the SDF identified by Hansen and Jagannathan (1991) as having minimum volatility under the LOP condition. It is a linear function of the passive portfolio returns $\mathbf{R_K}$. The error term w represents the difference between the fund return R_{MF} and the best "hedging" or "replicating" payoff $\mathbf{c}'\mathbf{R_K}$ obtained from passive portfolio returns.

[4] It is also possible to impose many conditions simultaneously. Cochrane and Saá-Requejo (2000) consider the no-good-deal and NA conditions jointly. In Section V.A, we examine empirically if the NA condition represents a binding constraint for the SDFs considered in this paper. We find that SDFs are almost always positive. Hence, the NA condition does not provide a meaningful empirical restriction in our sample and we do not consider it further.

https://doi.org/10.1017/S002210901700045X Published online by Cambridge University Press

Hence, w is the part of the fund return that is not spanned by passive portfolio returns. Finally, v is the parameter that accounts for the no-good-deal restriction and is a function of the maximum Sharpe ratio $\bar{h}$.

We can further economically understand the solution by rewriting it as follows:

$$\bar{\alpha}_{\text{MF}} = \text{E}[m^* R_{\text{MF}}] - 1 + v\,\text{E}[w^2]. \tag{16}$$

This equation shows that the best clientele alpha can be decomposed into two parts. The first part, $\text{E}[m^* R_{\text{MF}}] - 1$, is the LOP alpha developed by Chen and Knez (1996), based on the minimum volatility SDF. Similar to the best clientele performance measure, the LOP measure gives zero performance to passive portfolios by construction and thus does not suffer from the benchmark choice problem. It has been used also by Dahlquist and Söderlind (1999), Farnsworth, Ferson, Jackson, and Todd (2002), and Ahn et al. (2009), among others.

The second part, $v\text{E}[w^2]$, can be viewed as the investor disagreement between the best clientele alpha and the LOP alpha. This disagreement can come from two sources: the replication error w and the maximum Sharpe ratio restriction $\bar{h}$. With regard to the first source, if the passive portfolios perfectly span the mutual fund return, so that $w = 0$, then there can be no disagreement in evaluation. Otherwise, as the expected squared replication error for a fund becomes larger, such that it becomes tougher for investors to get the "same type" of opportunities from passive portfolios, then the potential disagreement among investors becomes greater. With regard to the second source, if the maximum Sharpe ratio allowed corresponds to the highest Sharpe ratio attainable in the passive portfolios, such that $\bar{h} = h^*$ and $\text{E}[\bar{m}^2] = \text{E}[m^{*2}]$, then there is no valuation disagreement because $v = 0$. In this case, no opportunity better than the ones from passive portfolios is deemed reasonable by investors. Otherwise, as the additional opportunities that investors would find admissible (i.e., not consider good deals) increase, as specified by a larger difference $\bar{h} - h^*$, the potential disagreement among investors increases.[5]

III. Methodology

A. Estimation

The solution for the best clientele performance evaluation measure necessitates the estimation of $2K + 1$ parameters for $\bar{m}$ ($\mathbf{a}$, $\mathbf{c}$, v), along with the alpha ($\bar{\alpha}_{\text{MF}}$). These parameters can be estimated and tested for significance using Hansen's (1982) GMM. For a sample of T observations, we rely on the

[5] Specifically, we can rewrite the disagreement parameter as

$$v = \sqrt{(\bar{h} - h^*)(\bar{h} + h^*)/(R_F^2 \text{E}[w^2])}.$$

This expression shows that v increases with $\bar{h} - h^*$ but is also related to the passive portfolios. For a given $\bar{h} - h^*$, v increases with h^* and decreases with $\text{E}[w^2]$, so that passive portfolios with a lower optimal Sharpe ratio and a lower spanning ability generate a lower disagreement parameter.

https://doi.org/10.1017/S002210901700045X Published online by Cambridge University Press

following $2K+2$ moments:

$$\frac{1}{T}\sum_{t=1}^{T}[(\mathbf{a}'\mathbf{R_{Kt}})\mathbf{R_{Kt}}]-\mathbf{1} = 0, \tag{17}$$

$$\frac{1}{T}\sum_{t=1}^{T}[(R_{\mathrm{MF}t}-\mathbf{c}'\mathbf{R_{Kt}})\mathbf{R_{Kt}}] = 0, \tag{18}$$

$$\frac{1}{T}\sum_{t=1}^{T}[(\mathbf{a}'\mathbf{R_{Kt}})+v(R_{\mathrm{MF}t}-\mathbf{c}'\mathbf{R_{Kt}})]^2-\frac{(1+\overline{h}^2)}{R_F^2} = 0, \tag{19}$$

$$\frac{1}{T}\sum_{t=1}^{T}[(\mathbf{a}'\mathbf{R_{Kt}}+v(R_{\mathrm{MF}t}-\mathbf{c}'\mathbf{R_{Kt}}))R_{\mathrm{MF}t}]-1-\bar{\alpha}_{\mathrm{MF}} = 0. \tag{20}$$

The K moments in equation (17) allow for the estimation of the LOP SDF, $m_t^*=\mathbf{a}'\mathbf{R_{Kt}}$, by ensuring that it correctly prices K passive portfolio returns. The K moments in equation (18) represent orthogonality conditions between the replication error term, $w_t=R_{\mathrm{MF}t}-\mathbf{c}'\mathbf{R_{Kt}}$, and passive portfolio returns, which are needed to estimate the coefficients $\mathbf{c}$ in the best replicating payoff $\mathbf{c}'\mathbf{R_{Kt}}$. The moment in equation (19) imposes the no-good-deal condition to estimate the parameter v, which is restricted to be positive to obtain an upper bound. In this moment, R_F represents a risk-free rate equivalent and is simply set to 1 plus the average 1-month Treasury bill return in our sample, which is 0.3393%. For consistency, we also include this 1-month Treasury bill return as one of the passive portfolio returns so that the estimated mean SDF is similar to $1/R_F$. Finally, using the estimated best clientele SDF, $\overline{m}_t=m_t^*+vw_t$, we obtain the upper performance bound for a mutual fund using the moment specified by equation (20).

For comparison with the best clientele alpha, we also examine the LOP performance measure of Chen and Knez (1996), which is based on the SDF with the lowest volatility. Specifically, the LOP alpha can be estimated with the following additional moment:

$$\frac{1}{T}\sum_{t=1}^{T}[(\mathbf{a}'\mathbf{R_{Kt}})R_{\mathrm{MF}t}]-1-\alpha_{\mathrm{LOP}} = 0. \tag{21}$$

In all cases, the system is just identified because the number of parameters is equal to the number of moments. Hence, the estimates are not influenced by the choice of weighting matrix in the GMM. Furthermore, although the system is estimated for one fund at a time, Farnsworth et al. (2002) show that this strategy produces the same point estimates and standard errors for alpha as estimating a system that includes an arbitrary number of funds. Finally, statistical significance is assessed with standard errors adjusted for conditional heteroskedasticity and serial correlation using the method of Newey and West (1987) with 2 lags.[6]

[6]We choose 2 lags to control for the nonzero serial correlation in monthly returns of a significant fraction of funds, which might invest in thinly traded stocks. For example, FC ((2015), Table 1) report a median serial correlation of 0.12, with 25% of funds having a serial correlation greater than 0.19, potentially due to microstructure effects such as nonsynchronous trading (e.g., Lo and MacKinlay (1990)). As a robustness check, we also reproduce the results of Table 3 by using no lag or 4 lags and find similar results.

https://doi.org/10.1017/S002210901700045X Published online by Cambridge University Press

B. Maximum Sharpe Ratio Choice

To implement the best clientele performance measure, two choices are particularly important: passive portfolios and the maximum Sharpe ratio. In Section IV, we introduce 3 sets of passive portfolios to assess the sensitivity of the results to this choice. This section discusses the maximum Sharpe ratio choice, which we base on the existing literature.

In general, the literature shows that researchers impose a subjective constraint on the maximum Sharpe ratio. One early contribution is Ross (1976). To study deviations from the arbitrage pricing theory, he imposes a maximum Sharpe ratio of twice the market Sharpe ratio, leading to a value of 0.25. Since the market portfolio should have the highest Sharpe ratio according to the CAPM, he argues that this restriction should reasonably account for all attainable Sharpe ratios. With a related argument that high Sharpe ratios are unlikely from the perspective of risk-based models, MacKinlay (1995) considers that a squared annual Sharpe ratio higher than approximately 0.6 is implausibly high.

In applying their no-good-deal bounds to Standard & Poor's (S&P) 500 option pricing, Cochrane and Saá-Requejo (2000) rule out opportunities with a Sharpe ratio greater than twice that of the S&P 500 (or, equivalently, twice the Sharpe ratio of their basis asset). Pyo (2011) and Floroiu and Pelsser (2013) use the same assumption in their empirical studies. Huang and Zhou (2017) develop an upper R^2 bound for predictive regressions using the no-good-deal approach and pick four times the market Sharpe ratio as the maximum ratio. A few papers consider different values of the maximum Sharpe ratio. Kanamura and Ohashi (2009) use values ranging from 2 to 3 times the market Sharpe ratio to find bounds for summer day options. Martin (2013) provides upper bounds on risk aversion using the no-good-deal approach and calibrates his bounds with 3 values for the maximum annual Sharpe ratio (i.e., 0.75, 1.00, and 1.25).

Overall, although the maximum Sharpe ratio is somewhat subjectively specified, the literature offers guidance on its selection. The most common choice is a maximum Sharpe ratio of twice the one of the underlying basis assets (which oftentimes include only an equity index). This choice can also be seen as adding the market Sharpe ratio to the maximum Sharpe ratio of the basis assets. In this paper, we follow this guidance by adding to the attainable Sharpe ratio of our passive portfolios a value of 0.1262, corresponding to the monthly Sharpe ratio of the market index (the Center for Research in Security Prices (CRSP) value-weighted index) in our sample. We denote this choice by $\bar{h} = h^* + h\text{MKT}$. More conservatively, we also consider adding half of this value as additional allowable opportunities, so that $\bar{h} = h^* + 0.5h\text{MKT}$. Although these are our two basic choices, we also examine the effects on the results of other sensible choices, such as directly doubling the attainable Sharpe ratio of the passive portfolios.

C. Cross-Sectional Performance Statistics

To summarize the alpha estimates for mutual funds in our sample, we provide numerous cross-sectional statistics. First, we provide the mean, standard deviation, and selected percentiles of the distributions of estimated alphas and their corresponding t-statistics, computed as $t = \hat{\alpha}/\hat{\sigma}_{\hat{\alpha}}$, where $\hat{\alpha}$ is the estimated alpha and $\hat{\sigma}_{\hat{\alpha}}$ is its Newey–West (1987) standard error. Second, we present t-statistics

https://doi.org/10.1017/S002210901700045X Published online by Cambridge University Press

https://doi.org/10.1017/S002210901700045X Published online by Cambridge University Press

to test the hypothesis that the cross-sectional mean of estimated alphas is equal to 0. To perform this test, we assume that the cross-sectional distribution of alphas is multivariate normal with a mean of 0, a standard deviation equal to the observed cross-sectional standard deviation, and a correlation between any two alphas of 0.044. This last value matches the average correlation between fund residuals, adjusted for data overlap, reported by BSW ((2010), p. 193) and by FC ((2015), Appendix, p. 36) when discussing the cross-sectional dependence in alpha among funds in their samples (which are similar to ours).

Third, we compute the proportions of estimated alphas that are positive, negative, significantly positive, and significantly negative (at the 2.5% level), and we report the p-values on the significance of these proportions using the following likelihood ratio test proposed by Christoffersen (1998) based on a binomial distribution:[7]

$$\text{LR} = 2\log\left[\frac{\left(1-\frac{n}{N}\right)^{N-n}\left(\frac{n}{N}\right)^{n}}{(1-\text{pr})^{N-n}(\text{pr})^{n}}\right] \sim \chi^2(1), \tag{22}$$

where n is the number of funds that respects a given criterion (i.e., being positive, negative, significantly positive, or significantly negative), N is the total number of funds, n/N is the empirical proportion tested, and pr is the expected probability under the null.

Fourth, to control for mutual funds that exhibit significant alphas by luck or "false discoveries," we apply the technique of BSW (2010) and its extension by FC (2015). Their ideas consist of computing the proportion of funds with t-statistics outside the thresholds implied by a significance level (γ), denoted by t^- and t^+, then removing from it the fraction of funds that exhibit large estimated alphas by luck, and finally adjusting the result to account for the power of the tests for detecting skilled or unskilled funds. The approach thus provides proportions adjusted for false discoveries.

Using the generalization of FC (2015), the proportions of unskilled ($\hat{\pi}^-$), skilled ($\hat{\pi}^+$), and zero alpha ($\hat{\pi}^0$) funds can be found by solving the following system of equations:

$$\widehat{\text{pr}}(t > t^+) = \frac{\gamma}{2}\times\hat{\pi}^0 + \delta^+\times\hat{\pi}^- + \beta^+\times\hat{\pi}^+, \tag{23}$$

$$\widehat{\text{pr}}(t < t^-) = \frac{\gamma}{2}\times\hat{\pi}^0 + \beta^-\times\hat{\pi}^- + \delta^-\times\hat{\pi}^+, \tag{24}$$

$$\hat{\pi}^0 = 1-\hat{\pi}^+-\hat{\pi}^-, \tag{25}$$

where $\widehat{\text{pr}}(t > t^+)$ ($\widehat{\text{pr}}(t < t^-)$) is the proportion of funds with t-statistics greater than t^+ (less than t^-), $\gamma/2$ is the test size, δ^+ (δ^-) is the probability that an unskilled (a skilled) fund is erroneously classified as a skilled (an unskilled) fund, and β^+ (β^-)

[7] A first test examines whether the proportions of positive or negative alphas are equal to 50%. A second test examines whether the proportions of significantly positive or negative alphas are equal to 2.5%.

is the probability that a skilled (an unskilled) fund is correctly classified. Hence, the proportion of skilled funds $\hat{\pi}^+$ is equal to the sample proportion of significantly positive funds $\widehat{\text{pr}}(t > t^+)$ less the proportions of funds with significantly positive performance by luck ($\gamma/2 \times \hat{\pi}^0$ for lucky zero alpha funds and $\delta^+ \times \hat{\pi}^-$ for very lucky unskilled funds), adjusted for the test power β^+. The proportion of unskilled funds $\hat{\pi}^-$ has a similar interpretation with respect to significantly negative funds, and the proportion of zero alpha funds $\hat{\pi}^0$ is the fraction of funds not classified as skilled or unskilled.

We follow BSW (2010) and FC (2015) in our implementation. As shown by FC, BSW use a special case where $\delta^+ = \delta^- = 0$ and $\beta^+ = \beta^- = 1$, which can be solved as $\hat{\pi}^0 = \widehat{\text{pr}}(t^- < t < t^+)/(1-\gamma)$, $\hat{\pi}^+ = \widehat{\text{pr}}(t > t^+) - \gamma/2 \times \hat{\pi}^0$, and $\hat{\pi}^- = \widehat{\text{pr}}(t < t^-) - \gamma/2 \times \hat{\pi}^0$. They advocate values of $t^- = -0.5$ and $t^+ = 0.5$ as efficient thresholds to adequately classify funds, which corresponds to a size of 0.3085. These choices result in our BSW classification.[8] FC argue that the BSW special case leads to biases because their assumed values for δ^+, δ^-, β^+, and β^- are unrealistic; instead, they use simulations to calibrate them. Advocating a size of 0.05, which leads to asymptotic values of $t^- = -1.645$ and $t^+ = 1.645$, they obtain (see their Panel A of Table 4) $\delta^+ = 0.005$, $\delta^- = 0.004$, $\beta^+ = 0.604$, and $\beta^- = 0.512$. Finally, they solve the equations numerically by minimizing the sum of squared errors of equations (23) and (24) subject to the Kuhn–Tucker conditions for constraints that $\hat{\pi}^+ \geq 0$, $\hat{\pi}^- \geq 0$, and $\hat{\pi}^+ + \hat{\pi}^- \leq 1$. These choices result in our FC classification.[9]

IV. Data

A. Mutual Fund Returns

Our fund data consist of monthly net-of-fees returns on actively managed open-ended U.S. equity mutual funds from Jan. 1984 to Dec. 2012. Our data source is the CRSP Survivor-Bias-Free U.S. Mutual Fund Database. Following Kacperczyk, Sialm, and Zheng (2008), we exclude bond funds, balanced funds, money market funds, international funds, and funds that are not strongly invested in common stocks to focus on U.S. equity funds.[10] To focus on actively managed

[8] The false discovery adjustment of BSW (2010) can lead to a negative proportion of unskilled or skilled funds when the unadjusted observed proportion is close to 0. In such instances, we follow BSW by setting the adjusted proportion to 0 and readjusting the proportion of zero alpha funds so that the proportions sum to 1.

[9] We do not recalibrate δ^+, δ^-, β^+, and β^- with simulations, as our sample is very similar to the FC (2015) sample. In Section V.C.4, we examine the robustness of our results to using simulated critical values at a size of 0.05 for t^- and t^+ and find that our conclusions are robust. Following BSW (2010) and the base case proposed by FC, our classifications assume that the true mean alphas are equal to 0.317% for skilled funds and −0.267% for unskilled funds. FC refine the approach further by estimating simultaneously the true mean alphas and proportions of skilled and unskilled funds and by estimating simpler models with only two alpha groups. We do not pursue these refinements.

[10] As in Kacperczyk et al. (2008), we identify U.S. equity funds by policy codes: CS; Strategic Insight objective codes AGC, GMC, GRI, GRO, ING, and SCG; Weisenberger objective codes G, G-I, AGG, GCI, GRO, LTG, MCG, and SCG; and Lipper objective codes EIEI, EMN, LCCE, LCGE, LCVE, MATC, MATD, MATH, MCCE, MCGE, MCVE, MLCE, MLGE, MLVE, SCCE, SCGE, and SCVE.

https://doi.org/10.1017/S002210901700045X Published online by Cambridge University Press

https://doi.org/10.1017/S002210901700045X Published online by Cambridge University Press

funds, we exclude index funds identified by the Lipper objective codes SP and SPSP and funds with a name that includes the word "index." We also exclude mutual funds that are not open-ended by consulting the variable "open to investors" in the database. Finally, we keep the funds only if they hold, on average, between 80% and 105% in common stocks.

From this initial sample of funds, we make further sampling decisions to alleviate known biases in the CRSP mutual fund database. Survivorship bias occurs when only surviving funds are sampled out of a population in which some funds enter and leave. Following Fama and French (2010), we select 1984 as our starting year because the CRSP mutual fund database is free from this bias from then on. This starting year also eliminates a related selection bias in the early years of the database, as discussed by Elton, Gruber, and Blake (2001) and Fama and French (2010). Back-fill and incubation biases are studied by Evans (2010). Back-fill bias arises because the database includes fund returns that are realized prior to the fund database entry. Incubation bias refers to a situation where only funds that perform well in an incubation period are eventually open to the public and included in the database. To address these biases, we follow Elton et al. (2001) and Kacperczyk et al. (2008). We eliminate observations before the organization date of the funds, funds that do not report their organization date, and funds without a name because they tend to be incubated funds. We also exclude funds that have total net assets (TNA) inferior to $15 million in the first year of entering the database.

As a last sampling choice, following BSW (2010) and others, we consider a minimum fund return requirement of 60 months, given that our GMM estimation system calls for enough observations to obtain reliable statistical estimates. Fama and French (2010) argue that this requirement introduces a survivorship bias and use instead an 8-month survival screen to estimate their alphas based on regressions. Estimating SDF alphas in our GMM setup is obviously not feasible with only 8 observations. Furthermore, an important part of our analysis consists of comparing best clientele alphas with LOP alphas, and survivorship bias should affect them similarly. Nevertheless, as a robustness check, we estimate our main performance measures for funds with at least 36 months of returns. We find that our results are similar and our conclusions are not altered by this choice.[11]

Considering all previous steps, we have a final sample of 2,786 actively managed open-ended U.S. equity mutual funds with returns for at least 60 months between 1984 and 2012.

B. Passive Portfolio Returns

The choice of basis assets imposes a trade-off between economic power (i.e., in theory, all assets available to investors should be included) and statistical power (i.e., an econometric estimation imposes limitations on the number of assets). We select 3 sets of basis assets to represent passive opportunities available to investors. Our basis assets always include the risk-free rate plus 1 of 3 sets: i) 10 industry portfolios, ii) 6 style portfolios, or iii) the market portfolio. These assets

[11] Specifically, we reproduce the results of Table 3 by using a 36-month screen rather than a 60-month screen. We find that with a 36-month screen, the mean LOP alpha is reduced by 0.0056% and the mean best clientele alphas are reduced by 0.0037% and 0.0031% for maximum Sharpe ratios of $h^* + 0.5h\text{MKT}$ and $h^* + h\text{MKT}$, respectively.

https://doi.org/10.1017/S002210901700045X Published online by Cambridge University Press

have been widely used in the empirical asset pricing literature and mutual fund performance literature to capture the cross section of stock returns. Classifications based on industry, style, or market sensitivities are also common in practice to categorize equity investments for investors. The inclusion of the risk-free rate accounts for cash positions in equity mutual funds and fixes the mean of the SDF to a relevant value (Dahlquist and Söderlind (1999)). By varying the number and type of assets included, we aim to examine the sensitivity of our results to these choices in light of the aforementioned trade-off.

Ten industry portfolios are used for our main results, and the data are from Kenneth French's Web page (http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html). The portfolios consist of consumer nondurables, consumer durables, manufacturing, energy, high technology, telecommunication, shops, healthcare, utilities, and other industries. The 6 style portfolios are also obtained from Kenneth French's Web page. The portfolios are constructed from 2 market equity capitalization (size) sorts (large and small) and 3 book-to-market (value) sorts (low, medium, and high). The market portfolio is the CRSP value-weighted index. Its returns and the risk-free rate returns are taken from the CRSP database.

C. Summary Statistics

Table 1 presents the summary statistics for the monthly net-of-fees returns of our sample of actively managed U.S. equity mutual funds (Panel A) and of the basis assets (Panel B). Panel A also includes summary statistics for the SDF alphas from the model of Carhart (1997) (hereafter Carhart alphas) and their corresponding *t*-statistics, which will serve as a basis for comparison for the LOP and best clientele alphas.[12]

In Panel A of Table 1, monthly equity fund average returns (net of fees) have a mean of 0.73% and a standard deviation of 0.3% across funds. Average returns range from −4.83% to 2.09%, and standard deviations range from 0.92% to 16.92%. Monthly Sharpe ratios vary from −0.464 to 0.379, with a mean of 0.086 and a standard deviation of 0.053. The cross-sectional distribution of Carhart alphas has a mean of −0.12% (*t*-stat. = −2.32) and a standard deviation of 0.246%, and 73% of funds have a negative Carhart alpha. At the 5% significance level, approximately 15% (2%) of funds have significantly negative (positive) Carhart alphas. These results are comparable to the typical findings in the literature. In Panel B, industry portfolios have mean returns from 0.83% to 1.17% and standard deviations from 3.99% to 7.22%. Style portfolios have mean returns from 0.80% to 1.22% and standard deviations from 4.58% to 6.76%. Sharpe ratios vary from 0.070 to 0.196 for industry portfolios and from 0.069 to 0.149 for style portfolios. Statistics for the market portfolio and the risk-free asset are as expected.

[12]The Carhart (1997) SDF is a linear function of the market, size (SMB), value (HML), and momentum factors available on Kenneth French's Web page. For each fund, we estimate jointly the Carhart SDF parameters and the alpha with GMM by requiring the SDF to price correctly the 1-month Treasury bill return and the 4 Carhart factors and by using a moment similar to equation (21) for the alpha, but with the Carhart SDF replacing the LOP SDF.

TABLE 1
Summary Statistics

Table 1 presents summary statistics for the monthly data from Jan. 1984 to Dec. 2012. Panel A reports the cross-sectional summary statistics (average (Mean), standard deviation (Std. Dev.), and selected percentiles) on the distributions of the average (Mean), standard deviation (Std. Dev.), minimum (Min.), maximum (Max.), Sharpe ratio (h), and Carhart alphas with their corresponding t-statistics for the net-of-fees returns on 2,786 actively managed open-ended U.S. equity mutual funds. It also reports the t-statistics (t-stat.) on the significance of the mean of estimated Carhart alphas (see test description in Section III.C). Panel B gives the average (Mean), standard deviation (Std. Dev.), minimum (Min.), maximum (Max.), and Sharpe ratio (h) for passive portfolio returns. Passive portfolios include 10 industry portfolios (consumer nondurables (NO_DUR), consumer durables (DURBL), manufacturing (MANUF), energy (ENRGY), high technology (HI_TEC), telecommunication (TELCM), shops (SHOPS), healthcare (HLTH), utilities (UTILS), and other industries (OTHER)), 6 style portfolios based on 2 market equity capitalization (size) sorts (big (B) or small (S)) and 3 book-to-market (value) sorts (low (L), medium (M), or high (H)), the market portfolio (MKT) based on the CRSP value-weighted index, and the risk-free asset (RF) based on the 1-month Treasury bills. All statistics are in percentages except for Sharpe ratios and the t-statistics. ** indicates statistical significance at the 5% level.

Panel A. Mutual Fund Returns and Carhart Alphas

Returns	Mean	Std. Dev.	Min.	Max.	h	Carhart α_{MF}	t-Statistics
Mean	0.7338	5.3400	−19.9976	16.4080	0.0857	−0.1200	−0.7398
Std. dev. (t-stat.)	0.3008	1.5632	5.6560	7.8772	0.0532	0.2456 (−2.321)**	1.2534
99th	1.3677	10.3594	−5.5453	41.5517	0.1990	0.4231	2.2289
95th	1.1405	8.2353	−12.8253	32.5704	0.1593	0.2120	1.3169
90th	1.0453	7.2002	−14.4430	27.0565	0.1427	0.1271	0.8321
75th	0.9044	6.0719	−16.5652	18.5645	0.1178	0.0109	0.0731
50th	0.7464	5.0272	−19.4020	14.1103	0.0900	−0.1016	−0.7466
25th	0.5946	4.3865	−22.9302	11.4669	0.0613	−0.2214	−1.5149
10th	0.4232	3.8968	−26.3190	9.9918	0.0229	−0.3811	−2.2580
5th	0.2943	3.4811	−29.0886	9.0829	−0.0033	−0.5037	−2.7996
1st	−0.1139	1.6197	−36.9313	5.3715	−0.0827	−0.8329	−4.1538

Panel B. Passive Portfolio Returns

Returns	Mean	Std. Dev.	Min.	Max.	h
Industry Portfolios					
NO_DUR	1.1713	4.3493	−21.0300	14.7400	0.1962
DURBL	0.8311	7.0347	−32.8900	42.9200	0.0698
MANUF	1.0667	5.1203	−27.3200	17.7800	0.1420
ENRGY	1.1223	5.3691	−18.3900	19.1300	0.1459
HI_TEC	0.9338	7.2260	−26.1500	20.4600	0.0822
TELCM	0.9689	5.2612	−15.5600	22.1200	0.1199
SHOPS	1.0394	5.0898	−28.3100	13.3800	0.1375
HLTH	1.1140	4.7552	−20.4700	16.5400	0.1636
UTILS	0.9444	3.9952	−12.6500	11.7600	0.1521
OTHER	0.8829	5.3165	−23.6800	16.1100	0.1024
Style Portfolios					
B_L	0.9403	4.7004	−23.1900	14.4500	0.1281
B_M	0.9705	4.5832	−20.3200	14.8500	0.1378
B_H	0.9279	5.2367	−24.4700	22.1600	0.1126
S_L	0.8037	6.7657	−32.3400	27.0200	0.0685
S_M	1.1221	5.2548	−27.5700	18.8700	0.1487
S_H	1.2282	6.2180	−28.0500	38.3900	0.1426
Market Portfolio					
MKT	0.9174	4.5814	−22.5363	12.8496	0.1262
Risk-Free Asset					
RF	0.3393	0.2166	0.0000	1.0000	—

V. Empirical Results

A. Best Clientele SDFs

We first examine the estimated best clientele SDFs and the sources of disagreement. As discussed in Section II.C, the best clientele SDF $\bar{m}$ depends on the error term w, related to the spanning ability of passive portfolios, and the disagreement parameter v, which accounts for the no-good-deal restriction. Table 2 reports cross-sectional statistics on R^2 from regressions of fund returns on different sets

https://doi.org/10.1017/S002210901700045X Published online by Cambridge University Press

TABLE 2

Spanning Regression R^2, Disagreement Parameter Estimates, and Stochastic Discount Factors for Performance Evaluation

Table 2 reports statistics on the cross-sectional distributions of R^2 values from the regressions of mutual fund returns on passive portfolio returns (Panel A), estimates of the disagreement parameter v for two best clientele performance measures (denoted by $h^*+0.5h$MKT and h^*+hMKT) (Panel B), and estimates of the stochastic discount factors (SDFs) for two best clientele performance measures and for the LOP measure (denoted by h^*) (Panel C). The passive portfolios in Panel A are the risk-free rate and 10 industry portfolios, 6 style portfolios, the market portfolio, or the factors from the model of Carhart (1997). The best clientele measures in Panels B and C consider maximum Sharpe ratios of $h^*+0.5h$MKT and h^*+hMKT (see definition in Section III.B), and all measures in Panels B and C use the risk-free rate and 10 industry portfolios (RF + 10I) as basis assets. Panel A provides the mean, standard deviation (Std. dev.), and 99th and 1st percentiles. Panel B provides the mean, standard deviation (Std. dev.), and proportions in percentages of estimated v that are significantly positive at the 5% (%v signif > 0 at 5%) and 10% (%v signif > 0 at 10%) levels. Panel C provides the mean, standard deviation (Std. dev.), 99th and 1st percentiles, and proportions in percentages of SDF estimates that are positive (%m > 0) and negative (%m < 0). The data (see description in Table 1) cover the period Jan. 1984–Dec. 2012.

Panel A. R^2 from Regressions of Mutual Fund Returns on Passive Portfolio Returns

	10 Industry Portfolios	6 Style Portfolios	Market Portfolio	Carhart Factors
Mean	0.8157	0.8526	0.7475	0.8352
Std. dev.	0.1432	0.1436	0.1652	0.1450
99th	0.9723	0.9773	0.9650	0.9744
1st	0.1741	0.1391	0.0570	0.1152

Panel B. Estimates of Disagreement Parameter v Using the RF + 10I Passive Portfolio Set

	v		t-Statistics	
	$h^*+0.5h$MKT	h^*+hMKT	$h^*+0.5h$MKT	h^*+hMKT
Mean	11.0036	13.7640	1.8285	3.9927
Std. dev.	5.5602	7.6070	0.3309	0.7389
%v signif > 0 at 5%	28.07	100.00		
%v signif > 0 at 10%	67.37	100.00		

Panel C. SDFs Using the RF + 10I Passive Portfolio Set

	h^*	$h^*+0.5h$MKT	h^*+hMKT
Mean	0.9971	0.9971	0.9971
Std. dev.	0.2571	0.3186	0.3819
99th	1.6621	1.8281	1.9967
1st	0.4538	0.2614	0.0777
%m > 0	99.98	99.81	99.39
%m < 0	0.02	0.19	0.61

of passive portfolio returns and, for comparison, on factors from the model of Carhart (1997) (Panel A), on estimates of the disagreement parameter v (Panel B), and on SDFs used for performance evaluation (Panel C). Panels B and C show the results using the 10 industry portfolios and the risk-free asset as passive portfolios.

Panel A of Table 2 is informative on the ability of passive portfolios to replicate fund payoffs. The panel shows that the 6 style portfolios provide the highest mean R^2 at 85.3%, followed by the factors from the model of Carhart (1997) at 83.5% and the industry portfolios at 81.6%. The market portfolio provides the lowest mean R^2 at 74.8%. Although the 3 sets are similar in their 99th percentile R^2, industry portfolios are better at spanning the fund returns that are most difficult to replicate, as indicated by the R^2 at the 1st percentile. Panel B shows that best clientele SDFs include a significant disagreement parameter v in most cases. When the maximum Sharpe ratio $\bar{h}=h^*+0.5h$MKT, the estimates of v have a mean of 11.0, with t-statistics averaging 1.8, and are significant at the 10% level for 67.4% of the funds. When $\bar{h}=h^*+h$MKT, the estimates have a mean of 13.8 and are statistically significant for all funds. Panel C looks at the empirical SDFs

https://doi.org/10.1017/S002210901700045X Published online by Cambridge University Press

in the best clientele measures and in the LOP measure (denoted by h^*). Given that the risk-free rate is part of the basis assets, the 3 SDFs have the same mean. Consistent with the increased estimates of v, the SDF standard deviations increase with the maximum Sharpe ratio. Finally, the fractions of negative SDFs are well below 1% in all cases, suggesting that imposing the NA condition to ensure the positivity of SDFs would not materially affect our results.

B. Best Clientele Performance Results

Table 3 presents the main empirical results. Using the risk-free rate and 10 industry portfolios as basis assets, it shows statistics on the cross-sectional

TABLE 3

Best Clientele Alphas Using the RF + 10I Passive Portfolio Set

Table 3 reports statistics on the cross-sectional distribution of monthly SDF alphas estimated with two best clientele performance measures, allowing for maximum Sharpe ratios of $h^* + 0.5h$MKT and $h^* + h$MKT (see definition in Section III.B), and with the LOP measure (denoted by h^*), using the risk-free rate and 10 industry portfolios (RF + 10I) as basis assets. Panel A provides the mean, standard deviation (Std. dev.), and selected percentiles of the distributions of estimated alphas (columns under Performance) and their corresponding t-statistics (columns under t-statistics). It also reports the t-statistics (t-stat.) on the significance of the mean of the estimated alphas (see test description in Section III.C). Panel B gives proportions of estimated alphas that are positive ($\%\tilde{\alpha}_{MF} > 0$), negative ($\%\tilde{\alpha}_{MF} < 0$), significantly positive ($\%\tilde{\alpha}_{MF}$signif > 0), and significantly negative ($\%\tilde{\alpha}_{MF}$signif < 0). It also provides proportions adjusted for false discoveries according to the BSW and FC classifications (see description in Section III.C), that is, proportions of zero alpha, unskilled, and skilled funds. It finally presents the p-values (in parentheses) for the likelihood ratio tests (see description in Section III.C) that the proportions of positive estimated alphas are equal to 50%, and the proportions of significantly positive and significantly negative estimated alphas are equal to 2.5%. The data (see description in Table 1) cover the period Jan. 1984–Dec. 2012. All statistics are in percentages except the t-statistics. *** indicates statistical significance at the 1% level.

Panel A. Performance and t-Statistics of Individual Mutual Funds

	Performance			t-Statistics		
	h^*	$h^* + 0.5h$MKT	$h^* + h$MKT	h^*	$h^* + 0.5h$MKT	$h^* + h$MKT
Mean	−0.1789	0.2360	0.4436	−1.2291	1.0220	1.9033
Std. dev.	0.2707	0.3344	0.4183	1.4743	1.4333	1.4009
(t-stat.)	(−3.138)***	(3.351)***	(5.036)***			
99th	0.3580	1.2708	1.8383	2.0211	4.4955	5.4554
95th	0.1787	0.7851	1.1586	1.0473	3.3544	4.2584
90th	0.0967	0.6440	0.9764	0.5435	2.7535	3.5898
75th	−0.0364	0.4225	0.6801	−0.2258	1.9417	2.7639
50th	−0.1630	0.1841	0.3575	−1.1398	1.0652	1.8940
25th	−0.2854	0.0223	0.1494	−2.1715	0.1720	1.0671
10th	−0.4501	−0.1029	0.0182	−3.1627	−0.8232	0.1570
5th	−0.5978	−0.1909	−0.0708	−3.6814	−1.4254	−0.4560
1st	−0.9247	−0.4353	−0.2388	−4.8763	−2.6052	−1.4975

Panel B. Performance Proportions

	h^*	$h^* + 0.5h$MKT	$h^* + h$MKT
Performance Sign			
$\%\tilde{\alpha}_{MF} > 0$	20.32	78.36	91.49
	(0.00)***	(0.00)***	(0.00)***
$\%\tilde{\alpha}_{MF} < 0$	79.68	24.41	8.51
Performance Significance			
$\%\tilde{\alpha}_{MF}$signif > 0	1.04	21.64	47.52
	(0.00)***	(0.00)***	(0.00)***
$\%\tilde{\alpha}_{MF}$signif < 0	29.54	2.26	0.47
	(0.00)***	(40.96)	(0.00)***
BSW Classification Adjusted for False Discoveries			
Zero alpha	49.49	49.19	22.06
Unskilled	50.51	0.00	0.00
Skilled	0.00	50.81	77.94
FC Classification Adjusted for False Discoveries			
Zero alpha	67.16	64.77	34.66
Unskilled	32.84	0.00	0.00
Skilled	0.00	35.23	65.34

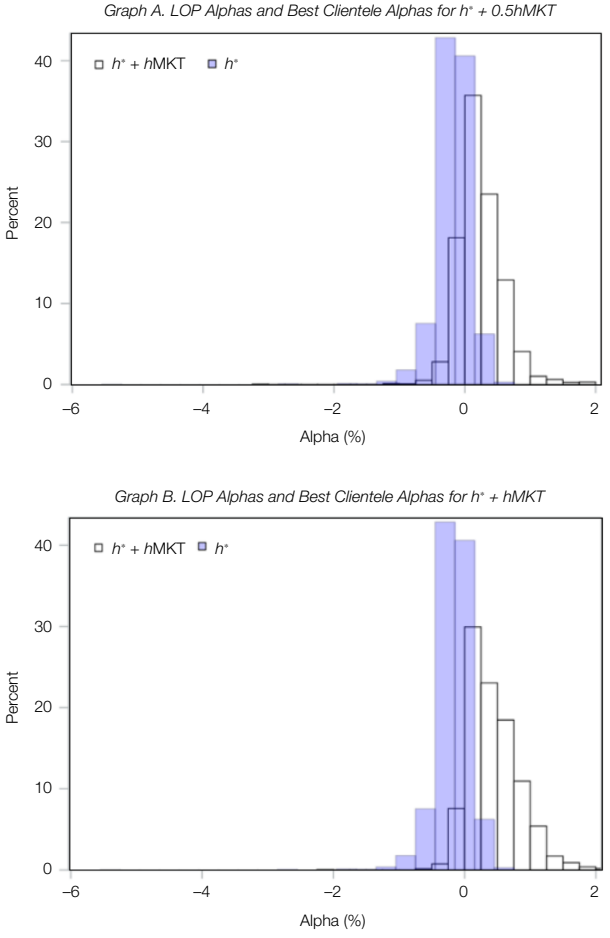
https://doi.org/10.1017/S002210901700045X Published online by Cambridge University Press

distribution of SDF alphas estimated with two best clientele performance measures, allowing for maximum Sharpe ratios of $\bar{h} = h^* + 0.5h\text{MKT}$ and $\bar{h} = h^* + h\text{MKT}$. The results for the LOP measure of Chen and Knez (1996) (denoted by h^*) are also reported. Figure 1 illustrates these results by presenting histograms on the distributions of LOP alphas and best clientele alphas for either $\bar{h} = h^* + 0.5h\text{MKT}$ (Graph A) or $\bar{h} = h^* + h\text{MKT}$ (Graph B).

In Panel A of Table 3, we provide the mean, standard deviation, and selected percentiles of the distributions of the estimated alphas (columns under Performance) and their corresponding *t*-statistics (columns under *t*-statistics). We also report the *t*-statistics discussed in Section III.C on the significance of the

FIGURE 1
Histograms of the Best Clientele and LOP Alphas

Figure 1 displays histograms illustrating the distributions of the best clientele and LOP alpha estimates, using the risk-free rate and 10 industry portfolios as basis assets. Graph A illustrates LOP alphas (denoted by h^*) and best clientele alphas allowing for a maximum Sharpe ratio of $h^* + 0.5h\text{MKT}$. Graph B illustrates LOP alphas and best clientele alphas allowing for a maximum Sharpe ratio of $h^* + h\text{MKT}$.



https://doi.org/10.1017/S002210901700045X Published online by Cambridge University Press

cross-sectional mean of the estimated alphas (see t-stat). The distribution of SDF alphas from the best clientele measure with $\bar{h} = h^* + h\text{MKT}$ has a mean of 0.444% and a standard deviation of 0.418%. When $\bar{h} = h^* + 0.5h\text{MKT}$, the mean and standard deviation decrease to 0.236% and 0.334%, respectively. Both means are statistically significant, with respective t-statistics of 5.04 and 3.35. Comparatively, the mean alpha from the LOP measure, which has no investor disagreement (as $v = 0$), is -0.179% (t-stat. $= -3.14$), a value similar to the mean Carhart alpha presented in Table 1, which also does not consider disagreement.

As in Ferson and Lin (2014), these results support an economically important divergence in performance evaluation between clienteles. For example, the magnitude of average disagreement between the LOP alpha and the best clientele alpha with $\bar{h} = h^* + 0.5h\text{MKT}$ is 0.415%. This value is comparable to the magnitude of investor disagreement documented by Ferson and Lin ((2014), Table III), who obtain bounds between 0.21% and 0.38% for the expected disagreement with traditional regression alphas for various benchmark returns. The divergence in alphas is well illustrated by the alpha distributions in Figure 1. The distributions of the t-statistics confirm that the increased performance associated with the best clientele measures results in more significantly positive alphas and fewer significantly negative alphas.

Panel B of Table 3 studies this issue further. It gives proportions of estimated alphas that are positive ($\%\bar{\alpha}_{\text{MF}} > 0$), negative ($\%\bar{\alpha}_{\text{MF}} < 0$), significantly positive ($\%\bar{\alpha}_{\text{MF}}\text{signif} > 0$), and significantly negative ($\%\bar{\alpha}_{\text{MF}}\text{signif} < 0$). It also presents the p-values for the likelihood ratio tests (described in Section III.C) on the significance of the proportions. It finally provides proportions adjusted for false discoveries with the techniques of BSW (2010) (denoted by the BSW classification) and of FC (2015) (denoted by the FC classification), that is, proportions of unskilled ($\hat{\pi}^-$), skilled ($\hat{\pi}^+$), and zero alpha ($\hat{\pi}^0$) funds.

The results show that the proportions of positive alphas and significantly positive alphas increase when considering best clienteles. They go from 20.3% to 91.5% for positive alphas and from 1.0% to 47.5% for significantly positive alphas. Accordingly, the proportions of negative and significantly negative alphas decrease, from 79.7% to 8.5% for negative alphas and from 29.5% to 0.5% for significantly negative alphas. The p-values confirm the significance of these results. Furthermore, the proportions of skilled funds increase considerably for best clienteles (going from 0.0% to 65.3% for the FC classification and from 0.0% to 77.9% for the BSW classification). Inversely, the proportions of unskilled funds for best clienteles disappear.[13] Again, accounting for investor disagreement and focusing on the best potential clienteles are keys to understanding the difference between our results and the existing literature. For example, the findings from the LOP measure are comparable to the results from the model of Carhart (1997) and typical of the literature, with less than 30% (more than 70%) of funds having

[13]For the FC classification, results for the LOP measure are similar to those of FC (2015) based on the Fama and French (1993) 3-factor model. For the BSW classification, they show fewer zero alpha funds and more unskilled funds compared to the results of BSW (2010) based on the Carhart (1997) model. This finding is consistent with the distribution of t-statistics being more to the left for LOP alphas than Carhart alphas, and with the analysis of FC that finds that the BSW classification is sensitive to the choice of test size.

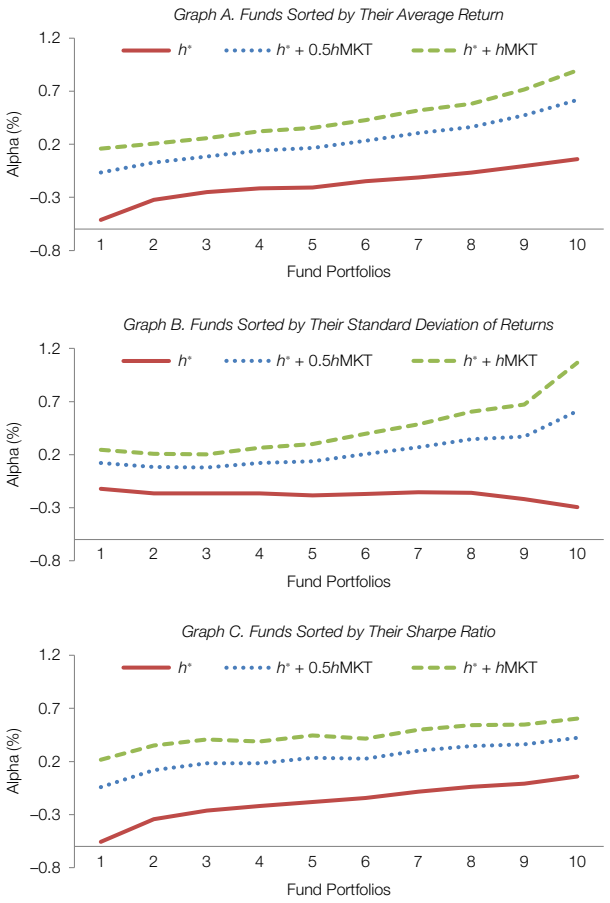
https://doi.org/10.1017/S002210901700045X Published online by Cambridge University Press

positive (negative) values. We find that increasing allowable opportunities by half the market Sharpe ratio is sufficient to obtain approximately the opposite result.

To gain more insight into the best clientele performance evaluation, Figure 2 presents the best clientele and LOP alphas for decile portfolios of the 2,786 mutual funds sorted in increasing order of their average return (Graph A), their standard deviation of returns (Graph B), and their Sharpe ratio (Graph C). Graphs A and C show that, not surprisingly, funds with higher average returns or Sharpe ratios are generally given higher best clientele and LOP alphas. Graph B reveals that best clientele alphas also increase with the standard deviation of returns, particularly for the 5 decile portfolios with the highest volatility, a relation not observed for LOP alphas. Hence, the difference between best clientele and LOP alphas

FIGURE 2
Best Clientele and LOP Alphas for Decile Fund Portfolios

Figure 2 displays best clientele alpha estimates (denoted by $h^* + 0.5h$MKT and $h^* + h$MKT) and LOP alpha estimates (denoted by h^*) for mutual funds grouped in decile portfolios. In Graph A, funds are sorted in increasing order of their average return. In Graph B, funds are sorted in increasing order of their standard deviation of returns. In Graph C, funds are sorted in increasing order of their Sharpe ratio.



https://doi.org/10.1017/S002210901700045X Published online by Cambridge University Press

(a disagreement measure in the spirit of the one proposed by Ferson and Lin (2014)) increases across portfolios formed on the standard deviation of returns. Intuitively, volatile funds represent somewhat "unique" opportunities for investors, because they cannot be easily replicated by passive portfolio returns. This "uniqueness" allows for a greater valuation disagreement between investors.

Overall, we find that an increase in admissible investment opportunities equivalent to half the Sharpe ratio of the market index leads to a generally positive performance for best clienteles. As stipulated by Chen and Knez (1996) and Ferson and Lin (2014), if the SDF alpha is positive, then there are some investors who would want to buy the fund. Our best clientele alpha results suggest that there are clienteles who would want to buy a majority of mutual funds, consistent with the real-life continued investments in these vehicles.

C. Additional Results

In this section, we conduct several robustness checks to the main analysis in Section V.B. The Internet Appendix (available at www.jfqa.org) provides a detailed description of the checks and displays the full results.

1. Conditional Best Clientele Performance Results

Ferson and Schadt (1996) and Christopherson, Ferson, and Glassman (1998) first argue that using public information improves performance measures and that alpha varies across the business cycle. In particular, Glode (2011) shows that mutual funds could be valuable to their clienteles by providing positive alphas during recessions, when their marginal utility (or SDF) is high. Moskowitz (2000), Kosowski (2011) and Kacperczyk, Van Niwuwerburgh, and Veldkamp (2014) also find evidence of better fund performance in recessions. We implement a conditional best clientele performance measure by following the scaled payoffs strategy of Chen and Knez (1996) and others. The Internet Appendix provides the details. We find a generally positive performance for best clienteles, with evidence more favorable in recessions than in expansions, consistent with Glode (2011) and others. Overall, the inclusion of conditioning information does not alter our conclusion on the importance of investor disagreement and best clienteles.

2. Sensitivity to Passive Portfolio Choice

We study the sensitivity of the unconditional results to the choice of passive portfolios by using basis assets based on the 6 style portfolios or the market portfolio. The results reported in the Internet Appendix confirm the previous findings. An increase in admissible investment opportunities equivalent to half the market Sharpe ratio leads to generally positive best clientele performance values and more skilled funds than unskilled funds, for all sets of basis assets. For example, the best clientele alphas estimated with $\bar{h} = h^* + 0.5h\text{MKT}$ have a mean of 0.289% (t-stat. = 3.98) for the style portfolio set and 0.270% (t-stat. = 4.23) for the market portfolio set.

3. Alternative Maximum Sharpe Ratios

We examine the effects of selecting other sensible choices of maximum Sharpe ratio $\bar{h}$ using all 3 sets of passive portfolios. The Internet Appendix details three additional approaches for setting $\bar{h}$ that can be justified from the

https://doi.org/10.1017/S002210901700045X Published online by Cambridge University Press

literature. We find that all $\bar{h}$ investigated lead to best clientele performance values that have significantly positive means and increase with the importance of additional opportunities allowed. Furthermore, the results show that the choice of $\bar{h}=h^*+0.5h\text{MKT}$, favored in previous sections, is a relatively conservative one. It adds fewer investment opportunities than the other sensible maximum Sharpe ratios.

4. Finite Sample Properties of Best Clientele Alphas

All previous results use the asymptotic GMM theory of Hansen (1982) to make inferences on estimated alphas. However, the finite sample properties of GMM estimators can deviate from their asymptotic properties. To obtain finite sample evidence on the estimates reported in Table 3, we conduct a bootstrap experiment that imposes the null hypothesis that alpha is 0 by adapting the procedure proposed in Fama and French (2010) to the case of SDF alpha. The Internet Appendix provides details on this experiment and reports the finite sample evidence. The simulations confirm that our findings on the generally positive performance for best clienteles are robust to finite sample issues. For example, the proportions of skilled funds for the FC classification are 31.3% with $\bar{h}=h^*+0.5h\text{MKT}$ and 61.5% with $\bar{h}=h^*+h\text{MKT}$ when computed from the simulated empirical distributions of the t-statistics, compared to 35.2% and 65.3%, respectively, in Table 3, when significance is assessed with the asymptotic distribution.

D. Zero-Alpha Implied Maximum Sharpe Ratios

Our analysis has thus far relied on an exogenous choice for the maximum Sharpe ratio $\bar{h}$. Previously, we showed that although this choice is somewhat subjectively specified, the literature offers guidance, providing a justification for our selections. This section investigates an alternative estimation strategy that does not require fixing $\bar{h}$ and leads to an evaluation of the investor disagreement needed for funds to be fairly priced by their potentially best clienteles.

To understand this strategy, notice that the selection of $\bar{h}$ allows for the estimation of the disagreement parameter v in equation (19). Then, v is needed for the estimation of $\bar{\alpha}_{\text{MF}}$ with equation (20). In this section, we proceed in reverse. We set a value for alpha that implies that the best clienteles give zero value to a mutual fund, $\bar{\alpha}_{\text{MF}}=0$. This choice leads to the estimation of v with equation (20), which then allows for the estimation of $\bar{h}$ with equation (19). We call the resulting estimated $\bar{h}$ the "zero-alpha implied maximum Sharpe ratio." Finally, the difference between this value for a fund and its sample optimal basis asset Sharpe ratio, $\bar{h}-h^*$, gives the increase in admissible investment opportunities sufficient to achieve zero alpha. As discussed in Section II.C, a greater difference leads to a wider potential valuation disagreement among investors (with no disagreement when $\bar{h}-h^*=0$). Hence, a small value for $\bar{h}-h^*$ indicates that little disagreement is needed to find clienteles who give a nonnegative value to a fund.

Table 4 shows the distributions of the implied Sharpe ratios estimated when fixing the best clientele alpha at 0 (under $\bar{\alpha}_{\text{MF}}=0$), the attainable optimal Sharpe ratios of the passive portfolios (under Basis Assets), and the fund-by-fund differences between the Sharpe ratios. For all 3 sets of basis assets, the average

https://doi.org/10.1017/S002210901700045X Published online by Cambridge University Press

TABLE 4

Zero-Alpha Implied Maximum Sharpe Ratios

Table 4 reports statistics on the cross-sectional distribution of monthly maximum Sharpe ratios implied by fixing the best clientele alpha at 0 (denoted by $\tilde{\alpha}_{MF}=0$), attainable monthly optimal Sharpe ratios of the passive portfolios (denoted by Basis Assets), and differences between both Sharpe ratios (denoted by Difference), using the risk-free rate and 10 industry portfolios, 6 style portfolios, or the market portfolio as basis assets. It provides the mean, standard deviation (Std. dev.), and selected percentiles of the distributions of the values. The data (see description in Table 1) cover the period Jan. 1984–Dec. 2012.

	10 Industry Portfolios			6 Style Portfolios			Market Portfolio		
	$\tilde{\alpha}_{MF}=0$	Basis Assets	Difference	$\tilde{\alpha}_{MF}=0$	Basis Assets	Difference	$\tilde{\alpha}_{MF}=0$	Basis Assets	Difference
Mean	0.2983	0.2571	0.0412	0.3087	0.2790	0.0296	0.1480	0.1146	0.0335
Std. dev.	0.0718	0.0408	0.0576	0.0556	0.0275	0.0449	0.0588	0.0383	0.0485
99th	0.5746	0.4628	0.2591	0.5384	0.3643	0.2186	0.3510	0.2713	0.2549
95th	0.4527	0.3002	0.1622	0.4119	0.3267	0.1135	0.2648	0.1626	0.1305
90th	0.3868	0.2838	0.1136	0.3676	0.3103	0.0763	0.2158	0.1484	0.0869
75th	0.3169	0.2692	0.0538	0.3223	0.2847	0.0373	0.1665	0.1270	0.0428
50th	0.2765	0.2528	0.0182	0.2924	0.2761	0.0134	0.1334	0.1171	0.0150
25th	0.2539	0.2364	0.0038	0.2777	0.2677	0.0033	0.1170	0.0965	0.0034
10th	0.2397	0.2214	0.0006	0.2674	0.2515	0.0005	0.0944	0.0700	0.0006
5th	0.2293	0.2147	0.0002	0.2540	0.2367	0.0001	0.0788	0.0593	0.0001
1st	0.2135	0.1895	0.0000	0.2356	0.2276	0.0000	0.0522	0.0100	0.0000

differences are small, so only a small increase in admissible opportunities is needed to change the negative mean LOP alphas into zero mean best clientele alphas. With 10 industry portfolios as basis assets, augmenting Sharpe ratio opportunities by only 0.041 (approximately one third of the sample market Sharpe ratio) is sufficient for fund evaluation to become 0, on average. Even fewer additional opportunities are needed for the passive portfolios based on 6 style portfolios (0.030) and the market portfolio (0.034). The percentiles show that funds require different levels of investor disagreement to be valued fairly. For example, for the industry portfolio basis assets and at the 1st and 99th percentiles, the zero-alpha implied Sharpe ratios vary from 0.214 to 0.575 and the Sharpe ratio differences vary from 0.000 to 0.259. Overall, these findings suggest that our conclusion on the generally positive performance values for best potential clienteles would hold unless an unreasonably low value for the maximum Sharpe ratio is selected.

E. Worst Clientele Performance Results and Total Disagreement

In all previous sections, the results focus on the best clientele alpha, a useful measure for managers who hope to cater to the right clienteles and for researchers who wish to understand whether there are some investors who would want to buy a fund. In this section, we examine the worst clientele alpha by estimating the lower bound on the expected alpha in our setup, and we investigate the total performance disagreement found by comparing the best clientele alpha with the worst clientele alpha. The worst clientele alpha is helpful in understanding whether there are clienteles that value funds more negatively than previous evidence shows. More optimistically, a positive worst clientele alpha indicates that all investors favorably value a fund. The measure of total performance disagreement extends the results of Ferson and Lin (2014), who argue that there is an economically important divergence in performance evaluation between clienteles.

https://doi.org/10.1017/S002210901700045X Published online by Cambridge University Press

It is straightforward to obtain the worst performance alpha. Cochrane and Saá-Requejo (2000) show that the solution to the lower-bound problem can be found using the SDF given by $\underline{m} = m^* - vw$, which we call the "worst clientele SDF." Then, the worst performance alpha is given by $\underline{\alpha}_{\text{MF}} = \text{E}[\underline{m} R_{\text{MF}}] - 1$ and can be estimated empirically by the GMM using the estimation strategy presented previously. Table 5 presents the cross-sectional performance statistics for the estimates of worst clientele alphas. The results in Panel A of Table 5 show that the distribution of worst clientele SDF alphas with $\bar{h} = h^* + h\text{MKT}$

TABLE 5

Worst Clientele Alphas Using the RF + 10I Passive Portfolio Set

Table 5 reports statistics on the cross-sectional distribution of monthly SDF alphas estimated with two worst clientele performance measures, allowing for maximum Sharpe ratios of $h^* + 0.5h\text{MKT}$ and $h^* + h\text{MKT}$ (see definition in Section III.B), and with the LOP measure (denoted by h^*), using the risk-free rate and 10 industry portfolios (RF + 10I) as basis assets. Panel A provides the mean, standard deviation (Std. dev.), and selected percentiles of the distributions of estimated alphas (columns under Performance) and their corresponding *t*-statistics (columns under *t*-statistics). It also reports the *t*-statistics (*t*-stat.) on the significance of the mean of the estimated alphas (see test description in Section III.C). Panel B gives proportions of estimated alphas that are positive ($\%\tilde{\alpha}_{\text{MF}} > 0$), negative ($\%\tilde{\alpha}_{\text{MF}} < 0$), significantly positive ($\%\tilde{\alpha}_{\text{MF}}\text{signif} > 0$), and significantly negative ($\%\tilde{\alpha}_{\text{MF}}\text{signif} < 0$). It also provides proportions adjusted for false discoveries according to the BSW and FC classifications (see description in Section III.C), that is, proportions of zero alpha, unskilled, and skilled funds. It finally presents the *p*-values (in parentheses) for the likelihood ratio tests (see description in Section III.C) that the proportions of positive estimated alphas are equal to 50%, and the proportions of significantly positive and significantly negative estimated alphas are equal to 2.5%. The data (see description in Table 1) cover the period Jan. 1984–Dec. 2012. All statistics are in percentages except the *t*-statistics. *** indicates statistical significance at the 1% level.

Panel A. Performance and t-Statistics of Individual Mutual Funds

	Performance			*t*-Statistics		
	h^*	$h^* + 0.5h\text{MKT}$	$h^* + h\text{MKT}$	h^*	$h^* + 0.5h\text{MKT}$	$h^* + h\text{MKT}$
Mean	−0.1789	−0.5938	−0.8014	−1.2291	−3.5688	−4.4825
Std. dev.	0.2707	0.3883	0.4796	1.4743	1.4261	1.4545
(*t*-stat.)	(−3.138)***	(−7.261)***	(−7.936)***			
99th	0.3580	−0.0667	−0.1978	2.0211	−0.5195	−1.5281
95th	0.1787	−0.1978	−0.3185	1.0473	−1.3712	−2.2725
90th	0.0967	−0.2572	−0.3861	0.5435	−1.8549	−2.7444
75th	−0.0364	−0.3625	−0.5015	−0.2258	−2.5970	−3.4673
50th	−0.1630	−0.5140	−0.6903	−1.1398	−3.4628	−4.3675
25th	−0.2854	−0.7159	−0.9670	−2.1715	−4.4925	−5.3940
10th	−0.4501	−1.0216	−1.3558	−3.1627	−5.3940	−6.3130
5th	−0.5978	−1.2511	−1.6315	−3.6814	−6.0399	−7.0541
1st	−0.9247	−1.8508	−2.3907	−4.8763	−7.2135	−8.2666

Panel B. Performance Proportions

	h^*	$h^* + 0.5h\text{MKT}$	$h^* + h\text{MKT}$
Performance Sign			
$\%\tilde{\alpha}_{\text{MF}} > 0$	20.32	0.39	0.04
	(0.00)***	(0.00)***	(0.00)***
$\%\tilde{\alpha}_{\text{MF}} < 0$	79.68	99.61	99.96
Performance Significance			
$\%\tilde{\alpha}_{\text{MF}}\text{signif} > 0$	1.04	0.00	0.00
	(0.00)***	(0.00)***	(0.00)***
$\%\tilde{\alpha}_{\text{MF}}\text{signif} < 0$	29.54	88.69	97.38
	(0.00)***	(0.00)***	(0.00)***
BSW Classification Adjusted for False Discoveries			
Zero alpha	49.49	1.56	0.00
Unskilled	50.51	98.44	99.96
Skilled	0.00	0.00	0.04
FC Classification Adjusted for False Discoveries			
Zero alpha	67.16	8.13	1.37
Unskilled	32.84	91.87	98.63
Skilled	0.00	0.00	0.00

https://doi.org/10.1017/S002210901700045X Published online by Cambridge University Press

has a mean of −0.801% (t-stat. = −7.94) and a standard deviation of 0.480%. When $\bar{h} = h^* + 0.5h\text{MKT}$, the mean and standard deviation decrease to −0.594% (t-stat. = −7.26) and 0.388%, respectively. Both distributions clearly show more underperforming funds compared to LOP alphas or Carhart alphas (reported in Panel A of Table 1). In fact, the proportions in Panel B of Table 5 indicate that less than 1% of funds have positive worst clientele alphas, suggesting that almost all funds are negatively evaluated by some investors. For worst clienteles, more than 90% of mutual funds appear unskilled. Thus, the evidence from worst clienteles is more negative than that from standard models.

With estimates of best and worst clientele alphas, it is straightforward to obtain the total performance disagreement. Cochrane and Saá-Requejo (2000) analyze the difference between their upper and lower bounds. In our context, we can interpret the difference between best and worst clientele alphas, $\bar{\alpha}_{\text{MF}} - \underline{\alpha}_{\text{MF}} = 2v\text{E}(w^2)$, as a measure of total performance disagreement between mutual fund investors. This measure is different from the bound on the expected disagreement with a traditional alpha of Ferson and Lin (2014). Their bound provides an estimate of the disagreement that best or worst clienteles could have with a traditional alpha but not the larger total disagreement that they have with each other. Focusing on the best and worst clientele measures with a maximum Sharpe ratio of $\bar{h} = h^* + 0.5h\text{MKT}$, Table 6 shows statistics on the cross-sectional distributions of total disagreement values and their t-statistics for the 3 sets of basis assets. The results extend the evidence of Ferson and Lin (2014) on the economically and statistically significant performance disagreement in mutual fund evaluation. For the 3 sets of basis assets, total disagreement values are significantly positive for more than 99% of funds, with mean values of 0.830% (t-stat. = 8.20) for the industry portfolio set, 0.746% (t-stat. = 8.48) for the style portfolio set, and 0.677%

TABLE 6

Total Performance Disagreement

Table 6 reports statistics on the cross-sectional distribution of monthly total performance disagreement from the best and worst clientele performance measures with a maximum Sharpe ratio of $h^* + 0.5h\text{MKT}$ (see definition in Section III.B), using the risk-free rate and either 10 industry portfolios, 6 style portfolios, or the market portfolio as basis assets. It provides the mean, standard deviation (Std. dev.), and selected percentiles of the distributions of estimated total disagreement values (columns under Value) and their corresponding t-statistics (columns under t-statistics). It also reports the t-statistics (t-stat.) on the significance of the mean of the estimated disagreement values. Total performance disagreement is defined as the difference between best and worst clientele alphas. The data (see description in Table 1) cover the period Jan. 1984–Dec. 2012. All statistics are in percentages except the t-statistics. *** indicates statistical significance at the 1% level.

	10 Industry Portfolios		6 Style Portfolios		Market Portfolio	
	Value	t-Statistics	Value	t-Statistics	Value	t-Statistics
Mean	0.8297	6.3779	0.7455	6.5374	0.6770	5.4709
Std. dev.	0.4818	1.8945	0.4177	1.7208	0.3732	1.8587
(t-stat.)	(8.179)***		(8.477)***		(8.615)***	
99th	2.6483	10.9180	2.7855	10.7928	2.2666	10.1666
95th	1.7199	9.6564	1.3909	9.4813	1.3370	8.8657
90th	1.4243	8.9265	1.1522	8.8512	1.1325	8.0961
75th	1.0413	7.6173	0.8576	7.5729	0.8104	6.7388
50th	0.7015	6.2729	0.6630	6.4852	0.5982	5.2337
25th	0.4761	5.0391	0.4985	5.3896	0.4384	4.0360
10th	0.3779	3.9845	0.3901	4.4984	0.3112	3.2524
5th	0.3387	3.3579	0.3458	3.8592	0.2657	2.8132
1st	0.2736	2.3733	0.2690	2.1814	0.1755	2.2428

https://doi.org/10.1017/S002210901700045X Published online by Cambridge University Press

(t-stat. = 8.62) for the market portfolio set.[14] These results are comparable with twice the values of investor disagreement documented by Ferson and Lin ((2014), Table III), consistent with interpreting their bound as the expected disagreement from a traditional alpha applicable to both best and worst clienteles.[15]

VI. Conclusion

In this paper, we apply the no-good-deal approach of Cochrane and Saá-Requejo (2000) to measure mutual fund performance from the point of view of the most favorable clienteles. This approach allows us to consider investor disagreement in mutual fund performance measurement. We use a large cross section of actively managed open-ended U.S. equity mutual funds to provide the first comprehensive performance evaluation exercise from the point of view of the potentially most favorable clienteles of each fund.

Our empirical results suggest that the long-standing issue of actively managed mutual fund underperformance might be due to the use of representative investors in standard measures. Considering investor disagreement and focusing on best clienteles cause funds to perform better, with the cross-sectional average of estimated alphas increasing with additional admissible investment opportunities in an incomplete market. These results are robust to adjustments for false discoveries, finite sample issues, and the use of different passive portfolios, different maximum Sharpe ratios (which control the investor disagreement parameter), and conditioning information. Overall, they support the findings of Ahn et al. (2009) and Ferson and Lin (2014) on the importance of investor disagreement in mutual fund evaluation.

The best clientele performance approach can be extended in numerous ways. First, it can be interesting to identify and characterize valuable investor clienteles by studying the implied marginal preferences reflected in the best clientele SDFs. Second, it is possible to investigate how the best clientele evaluations differ from the evaluations of representative investors implicit in commonly used performance models. Such a comparison, by examining how a representative investor can misrepresent the value of funds for some of their clienteles, could form the basis of a diagnostic of standard measures for the purpose of performance evaluation. Finally, additional or different restrictions can be imposed on the set of investor SDFs (e.g., the maximum gain–loss ratio condition of Bernardo and Ledoit (2000)) to adapt the approach to measuring the performance of portfolios with nonlinear payoffs (e.g., hedge funds).

[14] As explained in Section II of the Internet Appendix, the mean disagreement is smaller for the market portfolio set because its much smaller h^* leads to smaller v. Compared to the industry portfolio set, the mean disagreement is smaller for the style portfolio set because its slightly better spanning ability tends to produce smaller $E(w^2)$.

[15] In unreported estimations, following Ferson and Lin (2014), we separate out retail from institutional funds on the assumption that individuals are more heterogeneous than institutional investors. We do not find conclusive evidence of differences in best clientele alpha, worst clientele alpha, or total disagreement between the subsamples. In their analysis of fund flows and disagreement, Ferson and Lin (2014) also do not find significant differences.

https://doi.org/10.1017/S002210901700045X Published online by Cambridge University Press

References

Ahn, D.; H. H. Cao; and S. Chrétien. "Portfolio Performance Measurement: A No Arbitrage Bounds Approach." *European Financial Management*, 15 (2009), 298–339.

Bailey, W.; A. Kumar; and D. Ng. "Behavioral Biases and Mutual Fund Investors." *Journal of Financial Economics*, 102 (2011), 1–27.

Barras, L.; O. Scaillet; and R. Wermers. "False Discoveries in Mutual Fund Performance: Measuring Luck in Estimated Alphas." *Journal of Finance*, 65 (2010), 179–216.

Bernardo, A. E., and O. Ledoit. "Gain, Loss, and Asset Pricing." *Journal of Political Economy*, 108 (2000), 144–172.

Carhart, M. M. "On Persistence in Mutual Fund Performance." *Journal of Finance*, 52 (1997), 57–82.

Chen, Z., and P. J. Knez. "Portfolio Performance Measurement: Theory and Applications." *Review of Financial Studies*, 9 (1996), 511–555.

Chrétien, S. "Bounds on the Autocorrelation of Admissible Stochastic Discount Factors." *Journal of Banking and Finance*, 36 (2012), 1943–1962.

Christoffersen, P. "Evaluating Interval Forecasts." *International Economic Review*, 39 (1998), 841–862.

Christopherson, J. A.; W. E. Ferson; and D. A. Glassman. "Conditioning Manager Alpha on Economic Information: Another Look at the Persistence of Performance." *Review of Financial Studies*, 11 (1998), 111–142.

Cochrane, J., and J. Saá-Requejo. "Beyond Arbitrage: Good-Deal Asset Price Bounds in Incomplete Markets." *Journal of Political Economy*, 108 (2000), 79–119.

Cremers, M.; A. Petajisto; and E. Zitzewitz. "Should Benchmark Indices Have Alpha? Revisiting Performance Evaluation." *Critical Finance Review*, 2 (2013), 1–48.

Dahlquist, M., and P. Söderlind. "Evaluating Portfolio Performance with Stochastic Discount Factors." *Journal of Business*, 72 (1999), 347–383.

Del Guercio, D., and J. Reuter. "Mutual Fund Performance and the Incentive to Generate Alpha." *Journal of Finance*, 69 (2014), 1673–1704.

Elton, E. J.; M. J. Gruber; and C. R. Blake. "A First Look at the Accuracy of the CRSP Mutual Fund Database and a Comparison of the CRSP and Morningstar Mutual Fund Databases." *Journal of Finance*, 56 (2001), 2415–2430.

Evans, R. B. "Mutual Fund Incubation." *Journal of Finance*, 65 (2010), 1581–1611.

Fama, E. F., and K. R. French. "Common Risk Factors in the Returns on Stocks and Bonds." *Journal of Financial Economics*, 33 (1993), 3–56.

Fama, E. F., and K. R. French. "Luck versus Skill in the Cross-Section of Mutual Fund Returns." *Journal of Finance*, 65 (2010), 1915–1948.

Farnsworth, H.; W. E. Ferson; D. Jackson; and S. Todd. "Performance Evaluation with Stochastic Discount Factors." *Journal of Business*, 75 (2002), 473–503.

Ferson, W. E. "Investment Performance Evaluation." *Annual Review of Financial Economics*, 2 (2010), 207–234.

Ferson, W. E., and Y. Chen. "How Many Good and Bad Fund Managers Are There, Really?" Working Paper, University of Southern California (2015).

Ferson, W. E., and J. Lin. "Alpha and Performance Measurement: The Effect of Investor Disagreement and Heterogeneity." *Journal of Finance*, 69 (2014), 1565–1596.

Ferson, W. E., and R. Schadt. "Measuring Fund Strategy and Performance in Changing Economic Conditions." *Journal of Finance*, 51 (1996), 425–462.

Floroiu, O., and A. Pelsser. "Closed-Form Solutions for Options in Incomplete Markets." Working Paper, Maastricht University (2013).

Glode, V. "Why Mutual Funds Underperform." *Journal of Financial Economics*, 93 (2011), 546–559.

Glosten, L., and R. Jagannathan. "A Contingent Claims Approach to Performance Evaluation." *Journal of Empirical Finance*, 1 (1994), 133–166.

Gruber, M. J. "Another Puzzle: The Growth in Actively Managed Mutual Funds." *Journal of Finance*, 51 (1996), 783–810.

Hansen, L. P. "Large Sample Properties of Generalized Method of Moments Estimators." *Econometrica*, 50 (1982), 1029–1054.

Hansen, L. P., and R. Jagannathan. "Implication of Security Market Data for Models of Dynamic Economies." *Journal of Political Economy*, 99 (1991), 225–262.

Huang, D., and G. Zhou. "Upper Bounds on Return Predictability." *Journal of Financial and Quantitative Analysis*, 52 (2017), 401–425.

Kacperczyk, M.; C. Sialm; and L. Zheng. "Unobserved Actions of Mutual Funds." *Review of Financial Studies*, 21 (2008), 2379–2416.

https://doi.org/10.1017/S002210901700045X Published online by Cambridge University Press

Kacperczyk, M.; S. Van Niwuwerburgh; and L. Veldkamp. "Time-Varying Fund Manager Skill." *Journal of Finance*, 69 (2014), 1455–1484.
Kanamura, T., and K. Ohashi. "Pricing Summer Day Options by Good-Deal Bounds." *Energy Economics*, 31 (2009), 289–297.
Koski, J. L., and J. Pontiff. "How Are Derivatives Used? Evidence from the Mutual Fund Industry." *Journal of Finance*, 54 (1999), 791–816.
Kosowski, R. "Do Mutual Funds Perform When It Matters Most to Investors? U.S. Mutual Fund Performance and Risk in Recessions and Expansions." *Quarterly Journal of Finance*, 1 (2011), 607–664.
Lo, A., and A. C. MacKinlay. "An Econometric Analysis of Nonsynchronous Trading." *Journal of Econometrics*, 45 (1990), 181–212.
MacKinlay, A. C. "Multifactor Models Do Not Explain Deviations from the CAPM." *Journal of Financial Economics*, 38 (1995), 3–28.
Martin, I. W. R. "Consumption-Based Asset Pricing with Higher Cumulants." *Review of Economics Studies*, 80 (2013), 745–773.
Moskowitz, T. "Mutual Fund Performance: An Empirical Decomposition into Stock-Picking Talent, Style, Transaction Costs, and Expenses: Discussion." *Journal of Finance*, 55 (2000), 1695–1703.
Newey, W. K., and K. D. West. "A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix." *Econometrica*, 55 (1987), 703–708.
Pyo, U. "Minimax Price Bounds in Incomplete Markets." *Journal of Economics and Finance*, 35 (2011), 274–295.
Ross, S. A. "The Arbitrage Theory of Capital Asset Pricing." *Journal of Economics Theory*, 13 (1976), 341–360.
Ross, S. A. *Neoclassical Finance*. Princeton, NJ: Princeton University Press (2005).
Sharpe, W. F. "Mutual Fund Performance." *Journal of Business*, 1 (1966), 119–138.
Snow, K. N. "Diagnosing Asset Pricing Models using the Distribution of Asset Returns." *Journal of Finance*, 46 (1991), 955–983.

https://doi.org/10.1017/S002210901700045X Published online by Cambridge University Press