The Persistence of Fee Dispersion among Mutual Funds

Michael Cooper, Michael Halling and Wenhao Yang*

May 2020

Keywords: mutual funds, fund fees, fund expenses, price dispersion, price persistence.

JEL Classifications: G10, G11, G23.

---

* Cooper is with the David Eccles School of Business, University of Utah. Halling is with the Stockholm School of Economics and the Swedish House of Finance. Yang is with the School of Management and Economics, Chinese University of Hong Kong, Shenzhen. We thank Yakov Amihud, Jack Bogle, Gabriela Camera, Darwin Choi, James Choi (WFA discussant), Susan Christoffersen, Magnus Dahlquist, Monika Gehde-Trapp (discussant), Amit Goyal, Joseph Halford, Micah Hauptman, Rachel Hayes, Derek Hoff, Keith Jacob, Stephan Jank (discussant), Greg Kadlec, E. Han Kim, Dong Lou (discussant), Veronika K. Pool (discussant), Jonathan Reuter, David Robinson, Barbara Roper, Norman Schuerhoff, Mark Seasholes, Henri Servaes, Paolo Sodini, Laura Starks, Paul Tetlock, Charles Trzcinka, and Sunil Wahal. We also thank participants of the CEPR ESSFM meetings (2011, asset pricing week), the Northern Finance Association Meetings (2011), the 2012 WFA meetings, the Austrian Working Group on Banking and Finance (2013), the John McConnell's Contributions to Finance Conference (2015), the 12th FBI Symposium in Karlsruhe (2014), the Consumer Federation of America's research briefing on this paper (2017), the Chapman Conference on the Experimental and Behavioral Aspects of Financial Markets (2018) and seminar participants at Arizona State, HEC Lausanne, the Helsinki Finance Seminar, the Institute of Advanced Studies in Vienna, KAIST, Nanyang Tech University, the Shanghai Advanced Institute of Finance, the University of Arizona, the University of Colorado, the University of Gothenburg, the University of Montana, the University of Luxembourg, the University of Virginia, the University of California, San Diego, and VU University Amsterdam for their comments. Earlier versions of this paper were titled "Violations of the Law of One Fee in the Mutual Fund Industry" and "The Mutual Fund Fee Puzzle." Particular thanks go to Michael Lemmon for his contributions to earlier versions of the paper.

The Persistence of Fee Dispersion among Mutual Funds

Abstract

Previous work shows large differences in fees for S&P 500 index funds and other funds, and suggests that investors suffer wealth losses investing in high-fee funds when similar low-fee funds are available. In contrast, the neoclassical model of mutual funds (Berk and van Binsbergen, 2015) argues that percentage fees are irrelevant, as fund size will adjust in equilibrium such that net alphas are equal to zero. We show that fees matter from an investor perspective. We document (a) a strong negative association between net-of-fee fund performance and fees in a sample of all US and international equity funds, (b) economically large, robust, persistent, and pervasive fee dispersion in the mutual fund industry, and (c) important economic effects for investors. During the sample period, the mutual fund industry has generated a total value lost (i.e., a negative net value added) of 125 billion USD, coming predominantly from high-fee funds.

1. Introduction

Two important papers, Elton, Gruber, and Busse (2004) and Hortacsu and Syverson (2004), document substantial price dispersion for essentially identical S&P 500 index funds.[1] These results are surprising because in competitive markets, prices for close to identical products should have similar prices. Elton, Gruber, and Busse (2004) conclude that the combination of the inability to arbitrage (i.e., one cannot short sell open-ended mutual funds) and uninformed investors is sufficient to have the law of one price fail in the S&P 500 index fund market. Hortacsu and Syverson (2004), in contrast, link the fee dispersion to a combination of nonfinancial fund differentiation and search frictions. Elton, Gruber, and Busse and Hortacsu and Syverson both show evidence of wealth losses for investors in high-fee funds relative to similar low-fee funds. Thus, these two papers suggest that mutual fund markets are not perfectly competitive and that fees do matter to investors.

In contrast, the neoclassical view of mutual funds ((Berk and Green, 2004; Berk and van Binsbergen, 2015; Pastor, Stambaugh and Taylor, 2019; Stambaugh, 2019; Zhu, 2018) implies that fees do not matter. In the neoclassical model, with competitive markets and rational investors, in equilibrium, as fund size adjusts due to investor flows, gross-of-fee alphas will be on average equal to fees, and average net alphas will be zero. As a result, the percentage fees that funds charge are irrelevant. In this setting, manager skill is measurable using the gross-of-fee value added, which is the dollar value of before-fee alpha. However, from an investor perspective, fund net alphas are zero because positive net alphas are competed away by investors. For example, Berk and van Binsbergen (2015) show, in a sample of US and International equity funds, that average gross

[1] See also Elton, Gruber, and Rentzler (1989) who find that public commodity funds exist that underperform the risk-free rate, Christoffersen and Musto (2002) who find a wide dispersion in expenses across similar money market funds, and Iannotta and Navone (2012) who document that fund characteristics can only explain 40% of the variation in management fees among US equity funds. Theories on optimal fund fees are rather scarce (see, for example, Nanda, Narayanan and Warther, 2000; Das and Sundaram, 2002; Berk and Green, 2004; Pastor and Stambaugh, 2012). Nanda, Wang and Zheng (2009) analyze the fund's choice to issue several share classes that differ in their cost and load structure.

value added is positive (consistent with fund manager skill) but net-of-fee alphas are zero, consistent with the neoclassical view that fund fees should not matter for investors.[2]

In this paper, we consider and emphasize the investors' perspective. We have two primary goals. The first is to determine which viewpoint — fees matter to investors or fees do not matter — is on average evident in the data. The second goal, conditional on finding that fees do matter, is to examine the competitiveness of the mutual fund markets via an examination of mutual fund pricing. Since the publication of Elton, Gruber, and Busse (2004) and Hortacsu and Syverson (2004), the mutual fund markets have experienced dramatic growth, and along with this growth has come the continued debate, on the part of academics, practitioners, the justice system, and regulators, concerning the competitiveness of the mutual fund markets and related questions concerning mutual fund fees and their impact on investor performance.[3] Thus, our two goals are relevant and timely, and have important policy implications. Note that if fees are irrelevant (i.e., net alphas are zero and there is no systematic link between fees and net alpha), then the literature on mutual fund fee dispersion and the related political and legal battles over fund fees are not warranted.[4]

---

[2] A notable exception from this framework is Garleanu and Pedersen (2018) who propose a model of the asset management industry that in contrast to the neoclassical benchmark model of Berk and Green (2004) does not rely on diseconomies of scale at the individual fund level. In their model, investors can invest directly or search for an asset manager, there are search costs, and fees are determined endogenously by managers, with an equilibrium where there is dispersion in after-fee fund performance.

[3] Haslem, Baker and Smith (2006), Gil-Bazo and Ruiz-Verdu (2009), and Barras, Scaillet and Wermers (2010) argue that there is lack of price competition among funds. In contrast, Khorana and Servaes (2009), Wahal and Wang (2011), Cremers et al. (2016) and Hoberg, Kumar and Prabhala (2018) find evidence that the mutual fund industry behaves like a competitive industry.

[4] In the legal arena, the New York Times (August 9, 2016) reports in an article titled "M.I.T., N.Y.U. and Yale Are Sued Over Retirement Plan Fees," on a class action suit accusing prominent academic institutions of allowing their employees to be charged excessive fees on their retirement savings. The Wall Street Journal (May 18, 2015) reports on a recent case at the US Supreme court related to mutual fund fees charged within a company 401(k) plan. The plaintiffs in the case of Tibble et al. v. Edison International allege that the "…California utility company Edison International breached its duty to plan participants by using high-cost shares of some mutual funds when lower-cost alternatives were available." The topic of mutual fund fees has also received attention from the Executive branch of the US government. In a New York Times (February 28, 2015) editorial, they write that "A new study by the White House Council of Economic Advisers has found that financial advisers seeking higher fees and commissions drain $17 billion a year from retirement accounts by steering savers into high-cost products and strategies rather than comparable lower-cost ones."

We first investigate the question of whether fees matter to investors. As mentioned above, the neoclassical model of mutual funds argues that percentage fees are irrelevant, as fund size will adjust in equilibrium such that net alphas are equal to zero. As a first look at fee relevancy, we estimate the average net alpha across funds in our sample of all US and International equity funds from 1980 to 2017. Using both CAPM alphas and alphas constructed using tradable Vanguard mutual funds (following Berk and van Binsbergen, 2015, we deem these "BvB alphas"), we find that the average net alphas (both CAPM and BvB) are indistinguishable from zero, consistent with the neoclassical framework that fees do not matter and consistent with empirical findings in Berk and van Binsbergen.

However, we also find that the average alphas hide important fee-related *cross-sectional variation* among funds. We analyze in a next step whether there exists any systematic relation between fund fees and net alpha. We regress net-of-fee alphas on lagged fees and controls. For both CAPM and BvB alphas, we find statistically significant negative coefficients on fees. Thus, funds with higher fees have lower CAPM and BvB net alphas. This implies a systematic cross-sectional relation between fees and net alphas and suggests that, from an investor standpoint, fees do matter, and, as we will show, have important investor wealth implications.

Having established that fees are important from an investor standpoint, we then investigate the pricing of mutual funds. We study whether the price dispersion effects documented in Elton, Gruber, and Busse (2004) and Hortacsu and Syverson (2004) — that similar index funds charge different fees — are present in all US and International equity funds on the CRSP Mutual Fund Database. We use an extensive time-series and cross-section of fund data ranging from 1980 to 2017. The key advantage of this rich empirical setup is that it allows us to document novel and interesting results regarding the determinants of fund fees in the cross-section and the dynamics of fees and fee dispersion over time for a broad set of funds. Thus, in contrast to these earlier studies, which use a narrowly defined set of funds and a shorter time period, we are able to evaluate the pervasiveness of fee dispersion in a much broader context.[5]

Methodologically, we make the fees charged by different funds comparable by following a widely used approach in the economics literature to standardize prices: we examine the residuals (which we label "residual expenses") from yearly, cross-sectional regressions of total annual expenses (i.e., annual operating expenses, including management fees and 12b-1 fees) on lagged fund characteristics, such as risk and performance characteristics, measures of fund manager skill from Berk and van Binsbergen (2015), the extent of active management, service levels, fund size and age.[6] The residual approach allows us to compare prices across "similar" funds, under the assumption that we have controlled for the relevant fund characteristics.

In our empirical analysis, we first extend the work of Elton, Gruber, and Busse (2004) and Hortacsu and Syverson (2004) and find large levels of fee dispersion across similar S&P 500 index funds in the most recent 15 years of data after their studies. Elton, Gruber, and Busse (2004), for example, report the value of 34 bps for the interquartile spread in the reported annual expense ratios for their S&P 500 sample.[7] Remarkably, we find that the average interquartile spread in reported annual expenses has only slightly decreased to 30 bps for S&P 500 index funds in the period after their studies. Further, the 10^{th}–90^{th} percentile spread in reported expenses over this more recent period is 61 bps.

We also extend the analysis to include all index funds, and determine that the average spread in residual expenses between the 25^{th} and the 75^{th} percentile (between the 10^{th} and 90^{th} percentile) across all index funds over the sample period is 23 bps (47 bps); showing that the dispersion in

---

[5] Elton, Gruber, and Busse (2004) and Hortacsu and Syverson (2004), for example, are limited to a sample period of 6 years between 1995/1996 and 2000/2001 and a set of less than 100 S&P index funds.

[6] See Bakos (2001), Brown and Goolsbee (2002), Brynjolfsson and Smith (2000), Lach (2002), Nakamura (1999), Pratt, Wise and Zeckhauser (1979), Scholten and Smith (2002), and Sorensen (2000).

[7] Note that we also find the value of 34 bps if we limit our sample to the 1996 to 2001 period of the Elton, Gruber, and Busse (2004) study.

pricing for S&P 500 index funds originally documented in Elton, Gruber, and Busse (2004) and Hortacsu and Syverson (2004) extends to all styles of index funds.

Next, we address the important questions of whether the phenomenon of fee dispersion extends beyond index funds and how it has evolved over a more extended period of time. Here we turn our attention to the comprehensive sample of all US and international mutual funds investing in equities over 1980 to 2017 (we deem this the *All* funds sample). In yearly cross-sectional regressions of reported expenses on fund characteristics, we explain on average between 32% and 42% of the variation in fees across periods and models. Using the residuals from these regressions, we find that the average annual spread in residual expenses between the 25th and the 75th percentile (between the 10th and 90th percentile) across all funds over the sample is 47 bps (98 bps). These are economically large numbers.

Although we explicitly control for fund size in our regression framework, we also separately replicate our analysis for the funds in the largest TNA quintile to rule out the possibility that our results are driven by small funds or non-linearities associated with fund size. By the end of the sample, these large funds represent almost 85% of the market value of our fund universe. For these funds, we find an interquartile (10th–90th percentile) spread in residual expenses of 33 bps (70 bps). These numbers are smaller than the *All* funds sample, but they are still economically large given the substantial size of the funds included in the *Large* sample.

In a next step, we analyze the dynamics of fee dispersion. Despite the enormous growth in the mutual fund industry, the dispersion in residual expenses has stayed about the same. If we split the sample into halves, we find an interquartile (10th–90th percentile) spread in residual fees of 48 bps (104 bps) in the first and 47 bps (92 bps) in the second half for the *All* funds sample. Surprisingly, this pattern is true even for the sub-sample of large funds. In this case, we find again that spreads are similar across the first and second halves of the sample – the interquartile (10th–90th percentile) spread in residual fees is 31 bps (70 bps) in the first and 34 bps (69 bps) in the second half. Thus, our results on fee dispersion are not driven by early years in the sample. Instead, we find that the phenomenon of fee dispersion is very persistent across time.

Our results on fee dispersion are robust to a host of alternative specifications of the regression residual approach including variations on the fund characteristics used to generate the residual expenses — such as expanding the fund performance measures to include the gross value added skill measure and the gross alpha of Berk and van Binsbergen (2015). Furthermore, we add control variables that potentially help us understand better the role of percentage fees in the neoclassical setup such as an explicit measure of manager skill — the *skill ratio* — from Berk and van Binsbergen, an estimate of the alpha earned on the first dollar of a fund (parameter *a* from Zhu, 2018) and an estimate of the diseconomies of scale for a given fund (parameter *b* from Zhu, 2018). Across all of these extensions to our basic fee pricing regressions, we find only small changes to the average residual fee spreads and to the patterns of the residuals over time.

Our findings of economically large fee dispersion across arguably similar funds carries important investor implications, especially over the long run. To illustrate this, we estimate the returns to two hypothetical investors. One investor purchases low fee funds and the other purchases similar but higher fee funds. Based on residual expenses of the *All* funds sample of funds, the investor purchasing the lowest expense funds (i.e., the bottom quintile) would have earned compounded CAPM, Carhart (1997) four-factor model (FFC), and BvB net-of-fee alphas of 30%, 58%, and 24% higher, respectively, than the investor purchasing the most expensive (i.e., the top quintile) funds over our study period.[8] We find similarly results for the *Large* and *Index* funds samples.

Next, using measures from Berk and van Binsbergen (2015), we evaluate the distribution of gross value added and net value added (i.e., the product of gross or net-of-fee alpha and fund size) for our sample of funds as well as across quintiles of funds sorted by expense ratios. For the *All* funds sample, we find that gross value added is positive, consistent with the existence of skilled

[8] In many of our tests, we report abnormal returns using the CAPM, the Carhart four factor model (FFC), and the BvB set of tradable indexes. Berk and van Binsbergen (2015) show biases in the alphas computed using the FFC model due to the non-traded nature of the premiums in that model. In contrast, Berk and van Binsbergen use tradable index funds in the BvB model, which then result in alpha estimates that are based on factors that incorporate trading costs. To facilitate comparisons with earlier research, we report CAPM and FFC alphas, as well as BvB alphas.

managers. Next, we turn to the investor viewpoint and evaluate net value added. The equilibrium view from Berk and van Binsbergen (2015) posits that net value added should be on average zero, as funds with positive (negative) net value added grow (shrink) until net value added equals zero due to diseconomies of scale.

Interestingly, in our sample of *All* funds, we find that the average net value added is significantly negative whereas, for *Large* and *Index* funds, it is significantly positive. These patterns are consistent with an over-(under-) allocation of capital into the *All* (*Large* and *Index*) fund samples during the sample period. However, as we previously found for net alphas, there is important cross-sectional variation for net value added across fees levels: low fee funds tend to have positive net value added while high-fee funds have negative net value added, suggesting again a systematic relation between net value added and percentage fees.

In the final step of our analysis, we quantify the industry-wide economic importance of fee dispersion by summing net value added across all funds and years in the sample period. This analysis shows that the total net value added for the *All* funds sample amounts to negative 125 billion USD. This number can be interpreted as a measure of the total economic value added or, in this case, lost by the mutual fund industry for its investors. For the *Large* and *Index* funds sample, the same number is, in contrast, positive — 54 billion USD and 21 billion USD, respectively.

Most importantly, we again find that the total net value added number varies by percentage fees. While cheapest funds (i.e., fee quintile 1) earn positive total net value added across all samples of funds, expensive funds (i.e., fee quintiles 3 to 5) earn consistently negative total net value added. For example, the most expensive funds in the *All* funds sample generated a value lost for investors (i.e., a negative net value added) of 42 billion USD during the sample period.

In addition to the total net value added, we also provide a quantification of misallocated capital and an assessment of its relation to percentage fees. The notion of misallocated capital is based on Zhu (2018) and captures the notion that funds with positive (negative) net value added are too small (large), i.e., capital should flow into (away from) these funds. We follow Zhu (2018) to determine the optimal size of each fund based on its alpha on the first dollar invested as well as the diseconomies of scale associated with the fund's strategy. This analysis shows that 70% of the *All* funds sample is overinvested and, thus, should shrink. The amount of capital overinvested amounts to 1.4 trillion USD and represents an aggregate, life-time measure for all the funds in the sample. Interestingly, while the fraction of overinvested funds does not vary significantly across fee quintiles, the total amount of capital overinvested decreases substantially from 665 billion USD of funds in fee quintile 1 to 80 billion USD in fee quintile 5. This later pattern is consistent with the observation that, on average, high-fee funds are much smaller than low-fee funds; it is also consistent with the notion that, in terms of cumulative capital, investors do understand that high fee funds often times do not earn their fees (i.e., generate negative net value added, as discussed before).[9]

The remainder of the paper is organized as follows. In Section 2 we describe the data used in our analysis. In Section 3 we present results based on the regression-based fund pricing approach, including our standard fee models and extensions related to the neoclassical model of mutual funds. In section 4, we discuss the economic effects of fee dispersion, including investor wealth implications, value added effects, and misallocated capital. Section 5 concludes.

2. Data

2.1 Sample Construction

Our sample is based on the CRSP Mutual Fund Database. We start by selecting all the mutual funds whose CRSP Object Code starts with "E." These represent all the equity funds in the database, including both domestic equity funds and international equity funds. We include international funds since they are an increasingly important group of funds based on the fraction of all fund assets under management (Berk and van Binsbergen, 2015). Share classes are not

[9] Note that if we normalize the amount of overinvested capital by fund size – i.e., express it in relative terms – we do find that high-fee funds show the largest amounts of misallocated capital. For example, the median overinvested fund in the top-fee-quintile should shrink by 72% while the median overinvested fund in the bottom-fee-quintile should shrink by 64%.

automatically identified within the CRSP Mutual Fund Database. We use the MFLINKS tables provided by WRDS for this purpose. The original idea of these tables is to link the funds in the CRSP Mutual Fund Database with the ones covered in the Thomson Reuters Mutual Fund Ownership Database. In the time-series dimension, our analysis thus begins in March 1980, because that is when the share class data starts, and ends at the end of 2017. After identifying the individual share classes for a given fund, we aggregate the share classes (i.e., the expenses, returns, and other characteristics) into a common fund using value weighting (using the total net asset values to determine weights).[10]

2.2 Descriptive Statistics

Table I reports summary statistics of our fund sample. Details of the variable construction can be found in Table A in the Data Appendix. In most tables, we distinguish between a pre-1999 (up to and including 1998) and a post-1999 (including 1999) sample because several important variables such as fund family information (i.e., information on management companies) and flags for institutional funds only became available in the CRSP Mutual Fund Database in 1999.

The descriptive statistics show the dramatic increase in mutual funds over the past 30 years. In Table I, Panel A, the pre-1999 sample, the mean number of funds per year is 723, which increases to 2499 in the post-1999 sample. Note that the mean fund size (*TNA*) also increases from 668 Million USD pre-1999 to 1481 Million USD post-1999. Obviously the mutual fund industry has experienced a considerable increase in assets under management.

Intuitively, given more funds and thus presumably increased competition, we would have expected to find that the rapid expansion of the mutual fund industry was also accompanied by a substantial decrease in average expense ratios. This is not the case, however. Average annual

---

[10] We impose the following additional sample filters: A fund's inflation adjusted TNA (in 2014 USD) must be greater than $5 million. Once the fund TNA exceeds $5 million, we keep the fund in the sample regardless of TNA; Funds must have at least 24 months of return data; and Funds with negative TNA are set to missing.

expense ratios (*Expense ratio*) decreased from 136 basis points (bps) to 129 bps.[11] When we examine various measures of average abnormal returns to mutual funds, we see evidence consistent with the equilibrium hypothesized in the neoclassical model of mutual funds (Berk and van Binsbergen, 2015) in that average gross alphas obtain economically important point estimates (and in some cases statistically significant estimates) and net of fee alphas are essentially zero.[12] For example, the average TNA-weighted performance of our sample of funds, as measured by the annual CAPM gross alpha, is 0.61% (t-statistic = 0.80) pre-1999 and 0.90% (t-statistic = 1.32) post-1999. The BvB gross alpha is 1.14% (t-statistic = 1.37) pre-1999 and 0.77% (t-statistic = 2.45) post-1999. The CAPM net alpha is –0.28% (t-statistic = –0.36) pre-1999 and 0.04% (t-statistic = 0.05) post-1999, and the BvB net alpha is 0.17% (t-statistic = 0.21) pre-1999 and –.02% (t-statistic = –0.08) post-1999.[13]

Consistent with the differences in the gross and net alphas, at least in the post-1999 sample, we find that the average annual BvB gross value added (equal to gross alpha times TNA) is 15.75 Million USD, indicating that managers exhibit skill. From an investor viewpoint — that is, after fees — the average BvB net value added is much smaller, at 0.87 Million USD, and is consistent with Berk and van Binsbergen's (2015) conclusion that fund size will adjust in equilibrium such that net value added is equal to zero. In the pre-1999 period, things look somewhat different, as both gross value added and net value added are negative.

The average fund, over both time periods, has a market beta (*beta_mkt*) close to 1, a negligible exposure to HML (*beta_hml*), a small positive exposure to SMB (*beta_smb*), and a tiny positive exposure to UMD (*beta_umd*). After 1998, funds load slightly more on the market, and less on SMB and UMD, consistent with an aggregate strategy shift to market indexing. The four-factor

[11] The average expenses reported in Table 1 are equal-weighted. If we value weight the expenses, we find a slight decrease from pre to post-1999. In this case, the corresponding values are 102 bps for pre-1999 expenses and 92 for post-1999 expenses.

[12] All alphas reported in Table 1 are value weighted. We estimate them by first computing a value weighted average of fund returns for each month, then we regress the monthly excess returns on the corresponding factor models to obtain alpha estimates and t-statistics.

[13] For the FFC alphas, the pre-1999 gross alpha is 1.33% (t-statistic = 1.8), and the post-1999 gross alpha, pre-1999 net alpha, and post-1999 net alphas are all statistically insignificant.

model works well on average in explaining fund returns, yielding R^2 of 75% and 82%, in the pre and post-1999 periods, respectively.

Panel B reports full period pooled correlations across contemporaneous values of the variables. In this panel, we observe the first evidence that fees matter to investors: the correlations between fund expenses and net-of-fee performance measures are negative and statistically significant. For example, the correlation among expenses and net CAPM, FFC, and BvB alphas are all negative (albeit small in magnitude) and statistically significant. Thus, at least univariately, as fees go up, net performance goes down. We note that gross performance variables tend to increase as fees increase — the correlations among expenses and CAPM and FFC gross alphas are positive and statistically significant (albeit again small in absolute terms), a finding consistent with mechanisms in Berk and van Binsbergen (2015). The correlation between expenses and BvB gross alpha is insignificant. We also find a negative and significant correlation between expenses and fund size (*TNA*). Combining the results on size and performance leads to the results for BvB value added. We find a negative correlation between fees and both BVB gross value added and BVB net value added.

The most important limitation of this univariate analysis, of course, is that it ignores how expense ratios may reflect different fund strategies and characteristics. We will explore this matter in more detail in later sections of the paper. These simple summary statistics, however, already suggest that, to some extent, expense ratios can be explained by economic determinants. For example, funds' risk characteristics seem to be correlated with expense ratios: more expensive funds tend to exhibit higher absolute loadings on standard risk factors (i.e., on MKT, SMB, and UMD). Similarly, the average R^2 of the Carhart four-factor model decreases as fees increase, suggesting that managers of higher expense funds may be following "unique" strategies, likely in an attempt to outperform. However, these managers also trade much more (i.e., the *turnover* is positively correlated with expenses), which may contribute to their lower net alphas. Overall, these patterns between risk characteristics and expense ratios are intuitive and suggest that expensive funds do follow, at least to some extent, more active strategies, load more aggressively on individual risk factors, and implement strategies that go beyond the standard risk factors.

2.3 Relevancy of Fees

The correlations described above between fund expenses and net alphas suggest that fees do matter to investors. In this section, we more formally investigate the relevancy of fees from an investor standpoint by expanding the analysis to a regression framework. This is an important step in our analysis as the neoclassical fund framework implies the irrelevance of expense ratios. In Table II, we regress net-of-fee annual alphas on fees and controls. All independent variables except for fees in year t ($Expense\ ratio_{t}$) are lagged one year relative to the dependent variables. In columns 1 through 4, we present models for the pre-1999 period using BvB net alphas (Columns 1 and 2) and CAPM net alphas (columns 3 and 4). In columns 5 through 8, we present models for the post-1999 period, again using BvB net alphas (Columns 5 and 6) and CAPM net alphas (columns 7 and 8).

Across both periods, whether measured contemporaneously or lagged one year relative to the dependent variables, annual fund expenses are negatively and statistically significantly related to performance. These results are similar across periods, so, for brevity, we discuss the post-1999 period coefficients. The t-statistics on the expense ratio coefficients across the BvB and CAPM alpha models in columns 5 through 8 range from –3.06 to –4.52. The estimated coefficients are also economically significant given that estimates are, broadly speaking, around –1; this implies that any change in fees leads to a one-to-one change in the opposite direction in net alphas. Thus, in models with other important fund characteristics, net alphas are statistically significantly and economically negatively related to fund fees. Overall, these regressions show that fund fees matter for investors and are, accordingly, a relevant topic of study. Therefore, we next investigate the pricing of mutual funds.

3. The Pricing of Mutual Funds

Our goal is to compare prices (total expense ratios including management expenses and 12b-1 fees) across funds. Of course, not all funds are the same, and differences in fund characteristics might justify price differences. In the regression-based approach, we follow Lach (2002) and Sorensen (2000) to control for fund heterogeneity using cross-sectional regressions. Given that theory provides little guidance on the relevant characteristics to control for, [14] we use a comprehensive set of fund characteristics shown to be important in determining fund expenses (e.g., Gil-Bazo and Ruiz-Verdu, 2009; Wahal and Wang, 2011).[15]

3.1 Residual Expense Estimation and the Pricing of Individual Fund Characteristics

We regress fund expenses on lagged fund characteristics including performance and risk characteristics. As our set of explanatory variables changes over time (e.g., fund family information is only available after 1998), we estimate a cross-sectional regression each year. Another advantage of this specification is that it allows for changing relationships (i.e., time-varying coefficients) between fund characteristics and expenses. The residuals of these regressions can be interpreted as deviations of fund expenses from expected expenses given the set of characteristics used in the regression. Using the residuals, we can compare prices across "similar" funds under the assumption that we have controlled for the correct fund characteristics.[16]

In Panels A through D of Table III, we present the details of the yearly cross-sectional regressions used to estimate the residuals. The reported coefficients are time series averages of

[14] Most recent theory papers follow the neoclassical framework of Berk and Green (2004), in which percentage fees do not play any role and, as a consequence, are not chosen in any optimal sense. In Garleanu and Pedersen (2018), in contrast, percentage fees are determined endogenously through Nash bargaining. The equilibrium fee in their model would naturally have to be zero if asset markets were perfectly efficient, so that no benefit of information existed. If markets are not perfectly efficient, fees depend on the size of the market inefficiency as well as of the search costs that investors have to incur to find skillful managers.

[15] We have also used an alternative approach of matching funds based on holdings data. That approach yields qualitatively very similar results. The results are available from the authors.

[16] We are careful to include fund characteristics that we believe should matter to the average investor. Many of these characteristics are related to fund performance – items that should be the first order determinants of fund expenses. We also perform extensive robustness tests by varying our fund performance measures, controlling for retail and institutional funds, distribution channels, and other non-performance related characteristics.

cross-sectional regression coefficients obtained from annual cross-sectional regressions. We estimate these models separately for four samples: (i) S&P 500 index funds (Panel A),[17] (ii) all *Index* funds (Panel B),[18] (iii) the *All* funds sample of all active US and International equity mutual funds (excluding index funds, Panel C), and (iv) for the largest quintile of annually ranked TNA active funds (*Large* funds, Panel D). We standardize the independent variables to have a mean of zero and a standard deviation of one. The standardized coefficients thus allow us to discuss a fund fee price estimate for a one standard deviation change in each independent variable, and they also allow us to rank the fund characteristics by economic importance.

We start with S&P 500 index funds. Panel A of Table III summarizes coefficient estimates for this category. Elton, Gruber, and Busse (2004) and Hortacsu and Syverson (2004) claim that S&P 500 index funds are very homogeneous and, consequently, can directly be compared in terms of fees. In the period of 1999 to 2017, we find that our set of fund characteristics is still able to explain, on average, 65% of the variation in expense ratios across these funds. According to these fund characteristics, they are not as homogeneous as argued in previous studies. The most important characteristic in terms of fund pricing is *Family100 Dummy,* i.e., whether a fund is part of large management company (with more than 100 funds). These funds charge on average 14 to 15 basis points more. We observe that, as the total size of the fund management company (*MgmtCompTNA*) increases, funds charge about 9 basis points less for a one standard deviation change. Similarly, funds reduce fees by 7 basis points for a one standard deviation increase in the fund size (*ln(TNA)*). Both fund size and management company size likely indicate economies of scale effects.

Funds with institutional share classes (*Institutional*) charge 7 basis points less than non-institutional funds. When we examine the coefficients on various gross-of-fee performance

[17] We identify S&P 500 index funds by the following steps: (1) we first extract all index funds using CRSP's index fund flag; (2) for each index fund identified in step 1, we regress the full-sample monthly returns of the fund on the S&P 500 index monthly returns, and (3) using the beta of each fund from the regression we retain the funds that have a beta within the range [0.95, 1.05] and a R-squared within the range [0.95 and 1]. This results in a total of 105 S&P 500 index funds from 1999 to 2017.

[18] We identify all equity index funds using the CRSP index fund flag.

measures (lagged *Annual return*, lagged *CAPM gross alpha*, and lagged *CAPM gross* v*alue added*) and factor loadings (on MKT, SMB, HML, and UMD), the economic magnitudes are small, with little to no significance, consistent with expectations for a fund designed to track the S&P 500 index. Overall, we conclude that the subset of S&P 500 index funds is certainly homogenous in terms of strategy but not in terms of fee-relevant fund characteristics.

In Panel B of Table III, we expand the sample to all *Index* funds. We include all index funds in the analysis in order to gain a more comprehensive view on this segment of the mutual fund industry. Our empirical results here are very similar to the ones observed for the S&P 500 index fund sample. The same variables and patterns emerge as important drivers of index funds' expense ratios. Notably, *Turnover* plays a particularly important role in this sample, and a one-standard-deviation increase in turnover results in extra fees of around 14 basis points.

Next, we examine the fee regressions for the *All* funds sample of active mutual funds in Panel C of Table III (i.e., excluding index funds). Across all specifications and time periods, the pricing model explains approximately 33% of the variation in expenses for this sample. The signs of the coefficients are mostly consistent with the literature: e.g., we observe that less volatile funds (*Sdmret*), larger funds (*ln(TNA)*), younger funds (*Fund age*), lower turnover funds (*Turnover*), funds with large institutional share classes (*Institutional*), and funds with higher R^2 from the Carhart four-factor model have lower expenses. Across the pre- and post-1999 periods, we see that the negative relationships among fund fees, R^2 and *ln(TNA)* are similar, but links among fees, fund volatility, and fund size are only evident in the post-1999 period.

Zooming into the "pricing" of individual fund characteristics, and focusing on the post-1999 period to increase the number of variables and observations, we observe substantial variation across variables. Investors, for example, pay an additional 24 bps to be invested in a fund that belongs to a management company with more than 100 funds (*Family100 dummy*). Similarly, investors are willing to pay 11 bps to decrease the fund's r-squared (R^2) with the 4-factor model by one standard deviation, which can probably be viewed as the price of buying a more active and less diversified fund. A one standard deviation smaller fund *ln(TNA)* charges, on average, 28 bps more pre-1999 but only 14 bps more post-1999.

Regarding the price that investors pay for gross-of-expenses fund performance (in terms of lagged *Annual return*), the coefficients are negative and small, and only statistically significant in the pre-1999 period. The negative sign is consistent with Christoffersen and Musto (2002) and Gil-Bazo and Ruiz-Verdu (2008), who argue that, as fund returns go down, performance sensitive investors exit the fund, leaving a majority of performance-insensitive investors, for whom fund management then raises the fees. Note that we explicitly control for the flow-performance sensitivity of investors in a given fund in the empirical pricing models. In the case of the *All* funds sample in Panel C, we find a significantly positive coefficient, which that, in economic terms, implies that a one-standard-deviation increase in flow-performance sensitivity is associated with a 5 bps increase in fund expense ratios. This positive association is inconsistent with the above explanation that high-fee funds are predominantly populated by investors insensitive to performance or fees.

Coming back to the link between a fund's gross performance and its expense ratio, it is important to emphasize that past fund performance does not seem to be of first order importance for fund fees. In terms of the absolute magnitude of its price, at least, the lagged returns measure is far away from the impact of the most important variables. In addition to lagged gross returns, we also estimate specifications that use CAPM gross alphas as well as lagged CAPM gross value added as performance or managerial skill measures. In both cases, the results are mixed. Whereas Berk and van Binsbergen (2016) argue that the CAPM best explains investor behavior in the selection of mutual funds, our analysis suggests that CAPM gross alphas are not strongly associated with expense ratios. Gross value added (computed using gross CAPM alphas multiplied by funds inflation adjusted TNAs) — the measure of managerial skill proposed in Berk and van Binsbergen (2015) — shows a significant, positive coefficient in the post-1999 sample but is insignificant in the pre-1999 sample. In terms of economic magnitude, the point estimates of the coefficients are all tiny. We provide an expanded fee analysis based on the neoclassical fund framework in section 3.3.1.

Finally, we look at the annually ranked largest (top TNA quintile) funds separately (Table III Panel D). Even though we control for fund size in our *All* funds sample estimates, our results might still be driven by small funds if, for example, the relationship between size and fees is nonlinear. Furthermore, from an economic point of view, this is an important group since they represent, at the end of our sample period, more than 80% of the capital invested in our sample of funds. Our regression-based approach explains 32% pre-1999 and 41% post-1999 for the *Large* funds. In terms of coefficient estimates and relative importance of fund characteristics for fund pricing, large funds behave similarly to all funds, but some characteristics, such as fund size (*ln(TNA)*), return volatility (*Sdmret*), and *Fund age* lose statistical significance, especially post-1999. Similar to the *All* funds sample, past fund performance is not economically important compared to other characteristics, with the main difference relative to the *All* funds sample being that for the *Large* funds, the coefficients on lagged returns (*Annual return*) and *CAPM gross alpha* are positive (and statistically significant in the post-1999 period). The positive coefficient on the gross-of-fee CAPM alpha is consistent with the equilibrating mechanism proposed in Berk and van Binsbergen (2015); fees are positively associated with gross alphas.

Relative to the *All* funds sample, investors in *Large* funds pay more for fund style; the coefficients on the factor loadings (on MKT, SMB, HML, and UMD) are larger than for the *All* funds sample and are often statistically significant in the post-1999 period. Characteristics of fund management companies again play a dominant role in terms of economic impact on the pricing of the largest funds. Finally, other variables drop somewhat in terms of economic significance. For example, investors in the largest funds are only willing to pay an extra 9 bps to reduce the r-square (R^2) with the 4-factor model by one standard deviation, compared to 11 bps in the case of the *All* funds sample of active funds.

3.2 Detailed Analysis of Fee Dispersion

Our main point of interest, the distribution of residual expenses from the Table III expense regressions, is summarized in Table IV and visualized over time in Figure 1. In the figure, each year we plot the reported (left column) and the residual (right column) expense spread between the 25th and 75th, 10th and 90th, and 1st and 99th percentile points of the distribution (note that the mean residual is zero by construction).

For the sample of all *Index* funds, as reported in Table IV Panel A, we find that the average interquartile residual expense spread is 23 bps and that the average 10th–90th percentile residual expense spread is 47 bps. (For S&P 500 index funds, the interquartile spread in residual expenses is 18 bps and the 10th–90th spread is 36 bps.) For reported expenses, the average interquartile spread is 39 bps and the 10th–90th percentile spread amounts to 67 bps (30 and 61 bps for S&P 500 index funds, respectively). These results are consistent with the numbers reported in Elton, Gruber and Busse (2004) for the subset of S&P 500 index funds.

Considering the relative homogeneity of these funds in terms of strategy and the large number of fund characteristics controlled for in the fund pricing regressions, we consider these values to be stunningly large. Even more revealingly, Figure 1 (for all *Index* Funds[19]) shows that spreads in reported and residual expenses have not decreased much over time — they are nearly as big at the end of the sample as they were towards the beginning (ignoring the 1st to 99th percentile spread), despite large increases in the assets under management (from under 10% to over 35% by the end of the sample relative to the total size of our sample of active mutual funds that experience a substantial increase in assets under management themselves).

When we extend our analysis to the *All* funds sample of active equity funds, we find an average interquartile residual expense spread of 47 bps and an average 10th–90th percentile spread of 98 bps. Figure 1 (*All* Funds) illustrates that these large numbers are not driven by the early years of our sample. Rather, the opposite is the case, as we observe a gradual increase in these

[19] Given the relatively small size of the S&P 500 index fund sample as well as the similarity to the sample of all index funds, we focus on the larger all index funds sample in the remainder of the paper.

spreads up until the early 1990s, followed by largely constant levels until the mid-2000s, and a slight decrease during the most recent years. If we split the *All* funds sample into the first and second halves using 1999 as the midpoint, we find that the residual expenses are slightly higher in the first half. Specifically, we find an interquartile (10^{th}–90^{th} percentile) spread in residual expenses of 48 bps (104 bps) in the first and 47 bps (92 bps) in the second half (the first and second half numbers are not reported in the tables). These patterns seem surprising and rather counter-intuitive given the immense growth — in terms of dollars invested and funds competing for capital — experienced by the fund industry and given the increase in transparency due to improved regulation and easier information dissemination.

Even more surprisingly, we find similar patterns for the sub-sample of *Large* funds. On average, we document an interquartile residual expense spread of 33 bps and a 10^{th}–90^{th} percentile spread of 70 bps for residual expenses in this case. If we split the sample into halves, the interquartile (10^{th}–90^{th} percentile) spread in residual expenses is 31 bps (70 bps) for the first half and 34 bps (69 bps) for the second half. Thus, the residual expenses for largest TNA funds are mostly unchanged over time. Figure 1 shows the detailed dynamics of raw and residual expenses over time for the sample of *Large* funds. It is notable that, between 1985 and 2000, the subset of *Large* funds increased their share of TNA considerably. While the largest funds represent, on average, 78.6% of the market value of our sample, we see a pronounced increase starting from around 65.5% in 1984 to more than 85% around 2000. From 1985 to 2000 period, levels and spreads of reported expenses (left column) as well as spreads of residual expenses (right column) increased also noticeably. After 2000, these dynamics were somewhat reversed: expense levels of *Large* funds as well as their market share dropped while the tails of the reported and residual expense spread distributions slightly decreased, and the interior points (10^{th}–90^{th} and the interquartile spread) stayed more or less stable. Given that this sample only considers the 20% largest funds in a given year, we find these results to be economically relevant and surprising. Importantly, they show that the phenomenon of fee dispersion is not driven by or limited to small funds.

3.3 Fee Dispersion and the Neoclassical Model of Mutual Funds

Next, we extend the basic fund expense regression models used in Table III to consider other variables motivated by the neoclassical model of mutual funds. The neoclassical model has received a lot of attention as well as empirical support in the literature (see Berk and van Binsbergen, 2017, for a recent review), and it may help explain some of the unexplained residual fee dispersion. Specifically, we adjust the Table III expense regressions in the following important ways: (a) we use alphas and value added skill measures computed using a set of traded mutual funds (Berk and van Binsbergen, 2015) instead of alphas based on non-traded factor returns; (b) we add the *Skill ratio* — the t-static of the gross value added estimate measured over the entire history of the fund until a given point in time — also from Berk and van Binsbergen (2015)[20]; and (c) we add control variables that potentially help us understand the role of percentage fees in the neoclassical setup. Specifically, we include parameters *a* (an estimate of the alpha earned on the first dollar of a fund) and *b* (an estimate of the diseconomies of scale for a given fund) based on Zhu (2018). The Data Appendix provides further details regarding the calculation of those variables.[21]

Table V shows the results. Most noteworthy, except for the sample of *Large* active funds before 1999, there is no statistically significant relation between the BvB gross alphas and the expense ratios. This result challenges the equilibrium proposed in Berk and van Binsbergen (2015), as gross performance and fees should be positively related to each other, on average. We do find positive and significant coefficients on the *Skill ratio* for *All* funds and *Large* funds, at least in the post-1999 period, consistent with the notion that more skilled managers charge higher fees.

[20] Berk and van Binsbergen (2015) report strong evidence of long-lived performance persistence in their sample of active funds, thus we add the skill ratio as another measure to capture longer-term manager skill. See the Appendix in Berk and van Binsbergen for details on how they construct their mutual fund based traded asset benchmark.

[21] The calculation of parameters *a* and *b* requires monthly TNA values that are frequently missing in the CRSP mutual fund database before 1991. Thus, those parameters and all related analyses are calculated for the sample period 1991 to 2017. The *All* funds sample average value of *a* (*b*) is 0.54% (0.09%) with a standard deviation of 0.24% (0.04%). Those values are qualitatively comparable, albeit somewhat smaller, to the values reported in Zhu (2018).

We also find that parameter *a* of Zhu's model (i.e., the alpha on the first dollar) shows surprising and counter-intuitive results. First, the coefficient is frequently estimated to be significantly negative; in the case of the sample of *Large* funds, it even flips signs across time; in the post-1999 period it has the expected positive coefficient, but in the pre-1999 sample it shows a significantly negative coefficient. Parameter *b* of Zhu's model (i.e., the measure for diseconomies of scale — a higher *b* means more diseconomies of scale) receives significantly positive coefficients in 4 out of the 5 regressions, and these are frequently also large in economic terms. These positive coefficients imply that funds with strategies more susceptible to diseconomies of scale charge higher fees. This seems reasonable if we assume that those funds are also funds that, broadly speaking, are more active and, as a consequence, charge higher fees.

Next, we investigate the residual fee distributions after controlling for the new variables from the Table V expense regressions. Table IV Panel B shows the corresponding empirical distributions of these residual fees. The distributions of residual fees using the expanded set of neoclassical measures look very similar to the basic fee model residuals reported in Panel A. For the *All* funds sample of active funds, the Panel B time-series averages of the 25^{th}–75^{th} (10^{th}–90^{th}) percentile spreads are 49 (101) basis points. Across the other fund groups (*Large* and *Index* funds), the spreads are very similar across the two fee models and remain economically large. Note further that, if we plot the residuals from these regressions across time, as we did before, we find similar levels and dynamics in fee dispersion using this pricing framework. We interpret these new results to highlight that performance measures (calculated in the way suggested by Berk and van Binsbergen, 2015) as well as measures of diseconomies of scale do not materially reduce the variation in expense ratios across funds.[22]

[22] Following Christoffersen and Musto (2002), Bris et al. (2007), Bergstresser, Chalmers and Tufano (2009) we also evaluate fee dispersion separately for institutional and retail, both directly-sold and broker-sold, funds. Most importantly, our main results about the magnitude and the persistence of fee dispersion among mutual funds do not seem to be driven by distribution channels. Spreads, both in terms of reported expense ratios as well as residual expense ratios, are smallest for institutional funds followed by directly-sold retail funds.

4. Economic Magnitude of Fee Dispersion

In this section, we estimate the economic effects to investors arising from mutual fund fee dispersion using two approaches. First, we estimate returns to simple trading strategies that invest in high- and low-fee funds. Second, we analyze in detail the relation between value added of funds and annual expenses.

4.1 Trading Strategy

To implement the fee-based trading strategies, we consider a simple ex-ante trading strategy that trades funds based on the residual expense distribution, illustrating the negative wealth effects of investing in similar but higher expense funds. We assume no taxes. For comparison purposes, we also report a similar strategy using reported instead of residual expenses. More specifically, we compute the returns to a trading strategy that buys funds in the bottom quintile and sells funds in the top quintile of expenses. We rebalance these portfolios every year and compute the cumulative net CAPM, Fama-French-Carhart (FFC), and BvB alphas over the 38-year sample period to equally weighted portfolios.[23] Thus, this analysis can be viewed as a comparison between two hypothetical investors; one who invests in a portfolio of low fee funds and another one who invests in similar but higher fee funds.[24]

We report the results in Figure 2. Across all three fund groups (*All*, *Large*, and *Index*), we find that a low-fee investor strongly outperforms a high-fee investor. For *Index* funds (bottom row), the cumulative performance of our residual fee trading strategy over the 18-year sample is 7.55% (BvB net alpha), 7.81% (CAPM net alpha), and 9.84% (FFC net alpha). For the *All* funds sample

[23] We estimate the cumulative CAPM and FFC model alphas as follows. Using monthly net-of-fee returns from the annually rebalanced low-fee minus high-fee portfolio, we estimate the CAPM or FFC alpha each year by subtracting the model implied expected return from the realized return. The model implied expected return is computed using the beta estimated from that year. We then multiply the monthly alphas by 12 and compound over the years and report the cumulative alphas. To estimate the BvB alphas, we follow the Berk and van Binsbergen (2015) approach by regressing the full time series of monthly returns to the low-high fee portfolio on the set of traded mutual funds following their exact approach to obtain full period betas. We then estimate monthly alphas using the betas and realized monthly fund returns. We then compound the monthly alphas over time and report the cumulative alphas.

[24] This strategy illustrates the difference in long-term return performance across the high and low fee portfolios. It is not an implementable strategy since one cannot short sell open-ended mutual funds.

(Figure 2, top row), the cumulative alphas range from 24.33% (BvB net alpha) to 58.27% (FFC net alpha), showing strong performance for the low fee investor relative to the high fee investor. For the subset of *Large* funds, the pattern is somewhat different, as the hypothetical strategy underperforms in the first 10 years losing a cumulative abnormal return of approximately –15% during that period. It seems that in these early years, *Large* funds with high residual fees were generating abnormal performance. After 1990, however, the same pattern as in the other cases emerges, and the *Large* fund trading strategy increases monotonically over time — at speed similar to the cumulative spread in residual fees (the dashed lines).

In the left column of Figure 2, we report results for a trading strategy that conditions on reported expenses, rather than residual expenses. As one would expect, the patterns are similar across columns in Figure 2, but the magnitudes are considerably larger for the reported fee strategies. Interestingly, in the case of the *Large* sample, we also find that the cumulative returns from the hypothetical trading strategy outperform the cumulative spread in the expense ratios by a wide margin across all three measures of abnormal returns, indicating that high-fee funds underperform low-fee funds by more than just the fee difference.[25]

In sum, across the fund categories, these are all sizeable values in terms of performance disadvantage for those investors invested in high-fee funds. Even in the case of the *Large* active funds and the *Index* funds, these cumulated differences in reported fees amount to economically substantial values. Thus, while a difference in expense ratios of a few basis points might appear small on an individual fund-year level, it quickly accumulates to economically large effects.

4.2 Value Added and Capital Allocation

In this section, we evaluate the distribution of gross value added and net value added (i.e., gross alpha multiplied by fund size, and net alpha multiplied by fund size, respectively) for our sample of funds as well as across quintiles of funds sorted by expense ratios. If net value added is

[25] In contrast to our results, Ramadorai and Streatfield (2011) find little difference in performance across high and low management fees (i.e., the non-performance fee part of hedge fund expenses) for hedge funds. They conclude that high management fees are "money for nothing" in the hedge fund industry.

negative for a given fund — i.e., if investors earn negative net alphas — then the fund destroys value relative to the chosen sample of passive benchmarks; similarly, if a fund earns positive net value added, it creates positive value for the investors. The estimation follows Berk and van Binsbergen (2015) and represents simple (ex-ante distribution) and time-weighted (ex-post distribution) averages of funds' life-time value added.

Table VI reports the results. We first report annual before-fee value added estimates in Panel A. Consistent with Berk and van Binsbergen (2015), we find that the average value-added — measured ex-ante as well as ex-post (i.e., using a time-weighted average) — is significantly positive across all fund samples. For example, the average annual BvB gross value-added estimated ex-post is equal to 11.57 million USD in 2014dollars for the sample of *All* active funds and increases to an impressive average value-added of 56.91 million USD for the sample of *Large* funds. (Of course, conditioning on fund size implicitly conditions on success and, thus, the high value-added may not be surprising.) The overall takeaway, from the standpoint of gross performance measured with the tradable BvB indexes, is that fund managers do indeed appear to have skill.

When we switch to the investors' viewpoint, and examine net value added in Panel B, the overall picture across fund categories is largely consistent with the equilibrium hypothesized in the neoclassical view of mutual funds, which holds that average net value added should not be significantly different from zero. The point estimates for the annual BvB net value added are relatively close to zero, and they vary in sign across fund groups. For the *All* funds sample, the average yearly BvB net value added is –1.92 million USD (ex-post) and –2.77 million USD (ex-ante). For the *Large* funds, the average yearly BvB net value added is 4.17 million USD (ex-post) and 0.59 million USD (ex-ante).

The mean-value-added numbers, however, hide a lot of cross-sectional variation among funds. In Figure 3, we plot the yearly distribution of the BvB net value added for *All*, *Large*, and *Index* funds and report for each figure the spread between the 25th–75th, 10th–90th, and 5th–95th percentile points of the distribution.[26] Focusing on the *All* funds sample, we see large variation over time across the distribution of yearly net value added. There are noticeably large positive and negative spikes in the years immediately leading up to and after the 2000 period, and again near the end of the sample, consistent with the idea that investors are less than efficient in allocating monies between positive and negative alpha funds. We observe similar time series patterns for the *Large* funds. The *Index* fund graphs also illustrate relatively large yearly spreads in positive and negative net value added, but the time series occurrence of high and low spikes is different from the *All* and *Large* funds samples. The general pattern across the fund groups is one of increasing dispersion over the sample, especially relative to the pre-1995 period for the *All* and *Large* funds and the pre-2004 period for the *Index* funds.

Given the focus of our paper on expense ratios, we analyze in an additional step whether a relationship exists between expense ratios and value added. For this purpose, we split each fund group into quintiles based on annual expense ratios and then recalculate average gross and net-value-added measures separately for each sample. For the *All* funds sample, we find in Panel A of Table VI that the BvB gross value added is significantly positive in all fee quintiles (except for the ex-ante approach for the top fee quintile), suggesting that managerial skill can be found in each group. Interestingly, gross value added drops monotonically across fee quintiles. Looking next at the BvB net value added in Panel B for the *All* funds sample, we find that means are significantly negative for all fee quintiles except for the cheapest funds (i.e., the bottom quintile). Thus, consistently across fee quintiles, we find that charged expense ratios, on average, exceed the value created by fund management.

For the sample of *Large* funds, we observe a similar picture. Gross valued added (Panel A) is substantial and significantly positive but net value added (Panel B) is only significantly positive (and large in economic terms) for the bottom two quintiles of the fee distribution (it is insignificant for the cheapest funds in the ex-ante distribution case). Net value added is negative and

[26] We use the 5^{th} and 95^{th} points of the value added distribution since the 1^{st} and 99^{th} points contain more extreme values and create scaling issues for the figures.

economically large (and statistically significant) in the top two fee quintiles for the ex-post distribution (and negative and significant in the top fee quintile for the ex-ante distribution). We think it is particularly notable that such a strong association still holds between percentage fees and net value added after limiting the sample to funds in the top quintile of the size distribution.

Finally, for *Index* funds, gross value added (Panel A) is positive and statistically significant for the lowest fee quintile, and then is much lower and flips signs across fee quintiles 2 to 5. Thus, these funds seem to lack managerial skill except for those in the lowest fee quintile. For net value added, the *Index* funds have large positive and significant estimates for the low fee quintiles.[27] For quintiles 2 to 5, the net value-added means are negative and significant for the ex-post distribution and negative (but only significant in quintiles 4 and 5) for the ex-ante distribution. In other words, *Index* funds are very good investments with positive and significant net value added *only* in the cheapest fee quintile; as soon as these funds become expensive (i.e., quintile 2-5), they lack managerial skill and generate significantly negative net value added to an extent similar to the average active fund.[28]

In Panel C, we sum across *All* funds and report the sample funds' total lifetime net value added as well as the total lifetime net value added of *Large* and *Index* funds. This table offers another quantification of the economic magnitudes.[29] For *All* funds, the lifetime net value added is –124.65 billion, 53.61 billion for *Large* funds, and 20.97 billion for *Index* funds. Within expense quintiles, we see that for *All* funds, the aggregate negative lifetime net value added comes from the funds in the top four fee quintiles, which are all negative. Only the lowest fee quintile has lifetime positive net value added. We observe similar patterns for the *Large* and *Index* funds expense quintiles. Overall, only the very lowest expense quintile for all three categories of funds experience positive

[27] If the low fee quintile index funds on average have lower fees than the benchmark funds (i.e., the Vanguard funds), then they will have positive net alpha and positive net value added.
[28] We also estimate gross and net value added using the CAPM. Across all fund groups and fee quintiles, the values are uniformly lower than the BvB value added means and are almost always negative and statistically significant.
[29] This analysis is mechanically related to the ex-post distribution of net value added. Thus, the patterns in terms of signs are identical, as expected. It adds, however, an industry-wide, cumulative economic quantification in dollar terms.

lifetime net value added (and the second expense quintile for *Large* funds), but all other expense quintile funds have negative lifetime net value added. Most importantly, these cumulative net value added numbers are economically important and relevant.

As a last step in this analysis, we quantify the amount of capital invested in mutual funds that is misallocated such that a given fund's size deviates from its optimal size. In the context of Berk and van Binsbergen (2015) and Zhu (2018), funds with negative lifetime net value added are considered funds with an over-allocation of investment — potentially due to some investors behaving irrationally and committing too much capital to negative net alpha managers. Similarly, funds with positive net value added are under-allocated. As Berk and Green (2004) and Berk and van Binsbergen (2015) discuss, in a rational market for mutual fund capital, the provision of capital by investors to mutual funds should be competitive and should, across time, offset differences in net alphas.

To explicitly address this aspect, we follow Zhu (2018) and estimate each fund's optimal size. The Data Appendix explains the estimation in detail (variable name *Optimal Fund Size*). Once we have estimated optimal size, we calculate misallocated capital as the difference between optimal and actual fund size. We use the life-time average assets under management as our proxy for the actual size of the fund. Panel D summarizes the corresponding results. Note that we perform this analysis only for the *All* sample, as the estimation procedure for optimal fund sizes requires a large cross-sectional dimension and becomes unreliable in the substantially smaller subsets of *Large* and *Index* funds. In the case of Index funds, it is also questionable whether the framework, in general, applies, as these funds supposedly do not suffer from diseconomies of scale to an extent comparable to actively managed funds.

Panel D first reports the fraction of overinvested funds for each sample. We find that, out of the *All* funds sample, 70% of the funds are bigger than their optimal fund size. The fraction of overinvested funds shows a u-shaped pattern across fee quintiles; quintiles 1 and 5 feature the largest fractions, with 74% and 72% respectively, while quintile 3 has the lowest fraction, with 65%.

In a further step, we quantify the amount of capital overinvested in those funds by accumulating it across all funds. This step reveals that around 1.4 trillion USD are misallocated in underperforming funds in our sample, an economically relevant number indicating that a substantial amount of capital seems may be misallocated to negative-net-value-added funds. Across fee quintiles, the sum of overinvested capital decreases substantially. Although the value is around 665 billion USD for funds in fee quintile 1, it drops to 80 billion USD for funds in fee quintile 5. This substantial decrease is driven by the fact that high-fee funds are much smaller than low-fee funds.

In Panel D we also report the median of relative misallocation, where relative misallocation is calculated as the ratio of overinvested capital to fund TNA. We find that, across the *All* sample, 61% of the size of the median fund represents overinvested capital that should be withdrawn from the fund and invested elsewhere. This measure again shows a u-shaped pattern across fee quintiles. As one would expect, it reaches a maximum of 72% for funds in the highest fee quintile. This illustrates that, even though funds in fee quintile 5 contribute little to the total amount of misallocated capital, due to their small size, they are still the most inefficient funds in terms of relative misallocation.

An important question that remains to be discussed is how investors should alternatively invest the capital that is misallocated to funds in excess of their optimal sizes. The simplest strategy would be to reallocate it within the universe of actively managed funds if the industry offered enough capacity; i.e., if there are enough positive-net-value-added funds that are sufficiently below their optimal sizes. To assess the feasibility of this approach, we determine the capacity among all funds as well as among funds within a given fee quintile by summing across all negative amounts of misallocated capital.

The last column of Panel D, underinvested fund capacity, shows our estimates of those fund capacities. We find that in the *All* funds sample there is a capacity of around 2.2 trillion USD available; i.e., positive-net-value-added funds are too small by this aggregate amount. This value exceeds the 1.4 trillion USD discussed before that represent the capital that is inefficiently invested in negative-net-value-added funds. Thus, the sample of active funds offers enough capacity for the misallocated capital to be invested in an efficient way.

If we compare estimates of misallocated capital and capacity across fee quintiles, we find that reallocations within the same fee quintile would be enough to lead to an efficient allocation with the exception of the lowest fee quintile. In the latter case, the available capacity is not sufficient to cover the entire amount of misallocated capital. In this case, some reallocation to funds in higher fee quintiles such as 2 and 3 would be needed.

Overall, the asset management industry shows several characteristics consistent with the neoclassical framework. Above all, high-fee funds are much smaller than low-fee funds, reflecting their poorer performance. The results also highlight that capital is potentially misallocated in high- as well as low-fee funds. According to the neoclassical model of optimal fund size, a sizeable fraction of low-fee funds are too big and should shrink. Importantly, however, we do find, as one would expect, that the relative fraction of misallocated capital is largest for funds in the highest fee quintile. Our results on the dispersion of net alpha and net value added are potentially consistent with the theoretical work in Garleanu and Pedersen (2018) concerning the relation between fund fees and net performance. In their model, asset markets are inefficient, and active management is able to outperform passive strategies due to information acquisition costs. Uninformed investors face search costs and thus, in equilibrium, need to be compensated with positive expected net-of-fees alphas (otherwise, they would not search). Optimal percentage fees are determined by the fund's ability to create alpha and the target investors' search costs. Active funds that underperform their passive benchmarks survive because some uninformed investors pick funds randomly to avoid search costs.

5. Conclusion

In this paper, we have examined how mutual funds price their services based on a large cross-section of funds (i.e., all mutual funds that focus on investing in US and international equities) and a long time-series of 37 years. Surprisingly, after we control for a variety of fund characteristics related to performance, service, and other features that investors are likely to care about, we find that the unexplained portion of fund expenses exhibits considerable dispersion. Fee dispersion in this context may imply some degree of pricing inefficiency, in the sense that funds with similar characteristics charge different fees. We find this result among all index funds, the subset of S&P 500 index funds, and the subset of the largest funds.

In the neoclassical framework of Berk and Green (2004) and Berk and van Binsbergen (2015), percentage fees should not matter. In their framework, assuming competitive markets, fund flows respond to differences in fund performance and, together with diseconomies of scale, result in net alphas that are zero in equilibrium. An important contribution of our paper is to focus on the investor perspective (i.e., on net alphas and net value added) and to clearly document that percentage fees do matter.

We do not believe that our results conflict with the neoclassical paradigm. Indeed, we view them as complementing the existing empirical evidence. Consistent with Berk and van Binsbergen (2015), we also find that, on average, the mutual fund industry is well-described by the neoclassical framework. On the other hand, as we show, the average values of neoclassical framework fund manager performance measures like gross and net value added hide ample cross-sectional variation. One important cross-sectional fund dimension is percentage fees, as we document in the paper using various empirical tests. Although the proposed equilibrating mechanism of capital flows seems to work to some extent, it fails to completely offset the levels of fee dispersion documented in this paper.

Some potential limitations of our approach are that we do not control for the complete set of fund characteristics that affect fund fees, and that we do not accurately capture some relevant characteristics of the fund industry, such as frictions. While we do our utmost to include a comprehensive set of fund characteristics in our tests, and the robustness tests include reasonable variations in fund characteristics, we are limited to available fund-related databases with quality historical data. Accordingly, our analysis excludes variables not included in these data sources, such as potentially important measures like manager trust, as well as other related measures of investor satisfaction with a fund and its manager.[30]

Explanations potentially consistent with our results on fee dispersion are the existence of multiple investor clienteles with varying levels of sophistication, access to information, and exposure to market frictions. For example, retail investors often have limited knowledge of financial products. Thus, issues such as financial literacy and advising of households should be of first order importance for regulators. Of course, it is not obvious that enabling (retail) investors with the basic tools to select funds would solve the issue of fee dispersion.[31] As pointed out by Carlin and Manso (2011), funds may optimally react to investor learning by increasing the level of obfuscation (i.e., by making it harder for investors to learn). They argue, however, that an increase in competition should lower the incentives for obfuscation and should thus enable investors to learn more quickly.[32] From a regulator's perspective, therefore, it is also important to increase transparency and comparability in the industry.

[30]Gennaioli, Shleifer, and Vishny (2015) develop a model in which investors pick portfolio managers on performance and trust. In their model, they find that investor trust in the manager lowers an investor's perception of the portfolio's risk, and allows managers to charge higher expenses to investors who trust them more.

[31]Perhaps smarter investors are more fee aware. In a study of investor heterogeneity based on intelligence, Grinblatt et al. (2016) find that high IQ investors in Finland prefer low fee funds.

[32] Ellison and Wolitzky (2012) develop a static model of obfuscation and find that competition might actually lead to more obfuscation, increased search costs and more price dispersion.

Data Appendix

Table A. Variable Construction and Definitions

Variable Name	Variable Definition
12b-1 fees	Ratio of the total assets attributed to marketing and distribution costs. Available since 1992.
a	Parameter that captures the fund's alpha on the first dollar invested based on Equation (23) of the Zhu (2018), $a_i = \frac{1}{T_i}\sum_{t=1}^{T_i}(r_{it} + b_i \log(q_{it-1}))$, where T_i is the number of the years for fund i in the sample, b_i is defined below, r_{it} is the fund's gross return, and q is fund size.
Annual flow	Annual fund flow is estimated as $Flow_t = (TNA_t - TNA_{t-12}(1 + return_t))/TNA_{t-12}$ and winsorized at the 1% level. $Return_t$ is the return over the prior 12 months.
Annual return	Annual fund return, gross of expenses. We first compound monthly after-expense returns within the previous 12 months. This 12-month return is then added to the annual total expense ratio. Monthly return values are calculated as a change in NAV, including reinvested dividends from one period to the next. NAVs are net of all management expenses and 12b-fees. Front and rear load fees are excluded. Annual return is in decimal form, that is, 0.01 is 1%.
AUM	Inflation adjusted TNA (2014 US dollars).
b	A parameter capturing the extent of diseconomies of scale. Specifically, it is the regression coefficient of gross alpha on the natural log of the fund size, using the recursive demeaning method described in Zhu (2018). The regression is run for each fund decile formed based on the life-time average of fund size.
Back-end load	The load is a fee charged by the fund when an investor withdraws funds. The load typically varies by investment level and duration of the investment. The load value per fund is the equal weighted average across all back-end load values reported in the database across these dimensions.
B*eta_mkt* *Beta_hml* *Beta_smb* *Beta_umd*	Fund betas from the four-factor model. Each December, and for each fund, we estimate the monthly four-factor model betas using 3 years of monthly after-expense excess return.
BvB (net/gross) alpha	A fund's alpha computed using the method described in Berk and van Binsbergen (2015). Specifically, we use a set of Vanguard index funds to estimate a benchmark return using the life-time return for each fund, and then obtain the monthly alphas by subtracting the benchmark return from realized (net/gross) returns. Annual alphas are compounded from monthly alphas.
BvB (net/gross) value added	The product between a fund's (net/gross) annual BvB alpha and lagged TNA. See Berk and van Binsbergen (2015).
CAPM (net/gross) alpha	Each month, we estimate market beta using the fund's net returns with a rolling window of past 36 months. Then we compute the annual (net/gross) CAPM alpha using the realized fund (net/gross) return minus the CAPM implied expected return. Annual alphas are compounded from monthly alphas.
CAPM (net/gross) value added	The product between fund's (net/gross) annual CAPM alpha and lagged TNA.
Expense ratio	Annual ratio of total investment that shareholders pay for the fund's operating expenses, which include 12b-1 fees.

Variable Name	Variable Definition
Family100 dummy	A dummy variable equal to 1, 0 else, if a fund is part of a management company with more than 100 funds associated with it.
FFC (net/gross) alpha	We use the same procedure as in the case of CAPM alphas but add the value (HML), size (SMB), and momentum (MOM) factors to the regressions.
Flow-performance sensitivity	For each fund and each year, we estimate the fund's flow-performance sensitivity as the coefficient on lagged monthly performance in a regression that explains monthly flows. The regression starts with 1 year of monthly data and uses an expanding window. It is abbreviated as *Perf. sens.* in the tables.
Fund age	Age of fund calculated as the difference between current year and year of fund initiation.
Institutional	TNA weighted average of the share class level dummy that equals 1 if an institutional share class, 0 otherwise.
Life-time average AUM	The lifetime average of the fund-inflation-adjusted total net assets.
ln(MgmtComp TNA)	The natural log of each December's sum of total net assets of all funds belonging to the same management company.
Open	TNA weighted average of the share class level dummy that equals to 1 if open to new investors, 0 otherwise.
Optimal fund size	Optimal amount the fund manager can actively manage. $q^* = e^{\frac{a}{b}-1}$; see equation (21) in Zhu (2018). Parameters *a* and *b* are also defined in this Data Appendix.
R^2	For each December and each fund, we estimate the four-factor model using 3 years of monthly fund returns. Then we collect the adjusted R^2 of these models.
Sdmret	Standard deviation of monthly net-of-fee returns calculated from 3 years of monthly fund returns. Sdmret is in decimal form, that is, 0.01 is 1%.
Skill ratio	The skill ratio at time τ is the t-statistic of the *BvB gross value* added estimated from the period $t = 1$ to τ.
TNA	Total net assets as of end of December in millions of USD.
ln(TNA)	The natural log of total net assets per fund as of end of December.
Turnover	Annual fund turnover is calculated as the minimum of aggregated sales or aggregated purchases of securities, divided by the average 12-month total net assets of the fund.

References

Bakos, Y. (2001) The emerging landscape of retail e-commerce, *Journal of Economic Perspectives* 15, 69–80.

Barras, L., Scaillet, O., and Wermers, R. (2010) False discoveries in mutual fund performance: measuring luck in estimated alphas, *The Journal of Finance* LXV, No. 1, 179–216.

Bergstresser, D., Chalmers, J. M. R., and Tufano, P. (2009) Assessing the costs and benefits of brokers in the mutual fund industry, *The Review of Financial Studies* 22, 4129–4156.

Berk, J. B., and Green, R. C. (2004) Mutual fund flows and performance in rational markets, *Journal of Political Economy* 112, 1269–1295.

Berk, J. B., and van Binsbergen, J. H. (2015) Measuring skill in the mutual fund industry, *Journal of Financial Economics* 118, 1–20.

Berk, J. B., and van Binsbergen, J. H. (2016) Assessing asset pricing models using revealed preference, *Journal of Financial Economics* 119, 1–23.

Berk, J. B., and van Binsbergen, J. H. (2017) Mutual funds in equilibrium, *Annual Review of Financial Economics* 9, 147–167.

Bravin, J., and Moyer, L. (2015) High court ruling adds protections for investors in 401(k) plans, *Wall Street Journal* May 18.

Bris, A., Gulen, H., Kadiyala, P., and Rau, P. R. (2007) Good stewards, cheap talkers, or family men? The impact of mutual fund closures on fund managers, flows, fees, and performance, *The Review of Financial Studies* 20, 953–982.

Brown, J. and Goolsbee, A. (2002) Does the internet make markets more competitive? Evidence from the life insurance industry, *Journal of Political Economy* 110, 481–507.

Brynjolfsson, E., and Smith, M. (2000) Frictionless commerce? A comparison of internet and conventional retailers, *Management Science* 46, 563–585.

Carhart, M. (1997) On persistence in mutual fund performance. *Journal of Finance* 52, 57–82.

Carlin, B. I., and Manso, G. (2011) Obfuscation, learning, and the evolution of investor sophistication, *The Review of Financial Studies* 24, 754–785.

Christoffersen, S., and Musto, D. (2002) Demand curves and the pricing of money management, *Review of Financial Studies* 15, 1499–1524.

Cremers, M., Ferreira, M., Matos, P., and Starks, L. (2016) Indexing and active fund management: International evidence, *Journal of Financial Economics* 120, 539–560.

Das, S. R., and Sundaram, R. K. (2002) Fee speech: signaling, risk-sharing, and the impact of fee structures on investor welfare, *The Review of Financial Studies* 15, 1465–1497.

Editorial Board (2015) Protecting fragile retirement nest eggs, *New York Times* February 28.

Ellison, G., and Wolitzky, A. (2012) A search cost model of obfuscation, *Rand Journal of Economics* 3, 1756–2171.

Elton, E., Gruber, M., and Busse, J. (2004) Are investors rational? Choices among index funds, *Journal of Finance* 59, 261–288.

Elton, E., Gruber, M., and Rentzler, J. C. (1989) New public offerings, information, and investor rationality: The case of publicly offered commodity funds, *Journal of Business* 62, 1–15.

Garleanu, N., and Pedersen, L. H. (2018) Efficiently inefficient markets for assets and asset management, *The Journal of Finance* 73, 1663–1712.

Gennaioli, N., Shleifer, A., and Vishny, R. W. (2015) Money doctors, The *Journal of Finance* 70, 91–114.

Gil-Bazo, J., and Ruiz-Verdu, P. (2008) When cheaper is better: Fee determination in the market for equity mutual funds, *Journal of Economic Behavior and Organization* 67, 871–885.

Gil-Bazo, J., and Ruiz-Verdu, P. (2009) The relation between price and performance in the mutual fund industry, *Journal of Finance* 64, 2153–2183.

Grinblatt, M., Seppo, I., Keloharju, M., and Knüpfer, S. (2016) IQ and mutual fund choice, *Management Science* 62, 924–944.

Haslem, J. A., Baker, H. K., and Smith, D. M. (2006) Are retail S&P 500 index funds a financial commodity? Insights for investors, *Financial Services Review* 15, 99–116.

Hoberg, G., Kumar, N., and Prabhala, N. (2018) Mutual fund competition, managerial skill, and alpha persistence, *Review of Financial Studies* 31, 1896–1929.

Hortacsu, A., and Syverson, C. (2004) Product differentiation, search costs, and competition in the mutual fund industry: A case study of S&P 500 index funds, *Quarterly Journal of Economics* 119, 403–456.

Iannotta, G., and Navone, M. (2012) The cross-section of mutual fund fee dispersion, *Journal of Banking and Finance* 36, 846–856.

Khorana, A., Servaes, H., and Tufano, P. (2009) Mutual fund fees around the world, *Review of Financial Studies* 22, 1279–1310.

Lach, S. (2002) Existence and persistence of price dispersion: an empirical analysis, *The Review of Economics and Statistics* 84, 433–444.

Nakamura, L. (1999) The measurement of retail output and the retail revolution, *Canadian Journal of Economics* 32, 408–425.

Nanda, V., Narayanan, M. P., and Warther, V. A. (2000) Liquidity, investment ability, and mutual fund structure, *Journal of Financial Economics* 57, 417–443.

Nanda, V., Wang, Z. J., and Zheng, L. (2009) The ABCs of mutual funds: On the introduction of multiple share classes, *Journal of Financial Intermediation* 18, 329–361.

Pastor, L., and Stambaugh, R. F. (2012) On the size of the active management industry, *Journal of Political Economy* 120, 740–781.

Pastor, L., Stambaugh, R. F., and Taylor, L. A. (2019) Fund tradeoffs, working paper.

Porter, E. (2015) Americans aren't saving enough for retirement, but one change could help, *New York Times* March 3.

Pratt, J., Wise, D., and Zeckhauser, R. (1979) Price differences in almost competitive markets, *Quarterly Journal of Economics* 939, 189–211.

Ramadorai, T., and Streatfield, M. (2011) Money for nothing? Understanding variation in reported hedge fund fees, Working Paper.

Scholten, P., and Smith, S. (2002) Price dispersion then and now: Evidence from retail and e-tail markets, *Advances in Microeconomics: Economics of the Internet and e-Commerce* 11, 63–88.

Sorensen, A. (2000) Equilibrium price dispersion in retail markets for prescription drugs, *Journal of Political Economy* 108, 833–862.

Stambaugh, R. F. (2019), Skill and fees in active management, working paper.

Wahal, S., and Wang, A. (2011) Competition among mutual funds, *Journal of Financial Economics* 99, 40–59.

Zhu, M. (2018) Informative fund size, managerial skill, and investor rationality, *Journal of Financial Economics* 130, 114–134.

Table I. Summary Statistics

This table reports summary statistics and a correlation table of our sample of all equity mutual funds. The data covers the period from 1980 to 2017 and is a yearly panel. The sample is from the CRSP Mutual Fund Database and includes all mutual funds whose CRSP Object Code starts with "E," which represents all equity funds in the database, including both domestic equity and international equity funds. Variables are defined in the Data Appendix. Some information is only available after 1999 (e.g., information on management companies), so we split the sample into pre-1999 and a post-1999 subsets. Panel A presents full sample summary statistics. Panel B contains correlations. Stars indicate significance at the 1% (*) and 5% (+) level. *Expense ratio* is in basis points. Unless otherwise specified, all other numbers are in decimal form, that is, 0.01 is 1%. TNA is total net assets as of the end of December in millions of USD.

Panel A. All Funds Sample

	Pre-1999		**Post-1999**	
	Mean	**SD *(t-stat)***	**Mean**	**SD *(t-stat)***
Number of funds per year	723		2499	
Expense ratio	136.04	79.29	129.20	69.40
Annual return	0.16	0.18	0.10	0.24
CAPM gross alpha	0.61%	*(0.80)*	0.90%	*(1.32)*
CAPM net alpha	–0.28%	*(–0.36)*	0.04%	*(0.05)*
BvB gross alpha	1.14%	*(1.37)*	0.77%	*(2.45)*
BvB net alpha	0.17%	*(0.21)*	–0.02%	*(–0.08)*
BvB gross value added	–3.47	171.21	15.75	285.59
BvB net value added	–12.23	176.42	0.87	277.54
Beta_mkt	0.91	0.25	0.99	0.27
Beta_smb	0.27	0.42	0.17	0.36
Beta_hml	0.00	0.44	–0.01	0.38
Beta_umd	0.04	0.28	0.00	0.23
R^2	0.75	0.25	0.82	0.18
TNA	667.59	2414.14	1481.14	5599.03
ln(TNA)	4.92	1.75	5.56	1.90
Sdmret	0.05	0.02	0.05	0.02
Fund age	12.36	13.44	14.06	12.01
Perf. Sens.			0.12	0.22
Turnover			0.90	1.81
ln(MgmtComp TNA)			8.95	2.57
Institutional			0.29	0.38
Family100 dummy			0.94	0.22
Open			0.46	0.50

Panel B. Pooled Correlations

		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20
(1)	**Expense ratio**	1.00																			
(2)	**FFC net alpha**	-0.05*	1.00																		
(3)	**FFC gross alpha**	0.01+	1.00*	1.00																	
(4)	**CAPM net alpha**	-0.03*	0.77*	0.77*	1.00																
(5)	**CAPM gross alpha**	0.02*	0.77*	0.77*	1.00*	1.00															
(6)	**BvB net alpha**	-0.09*	0.59*	0.59*	0.65*	0.65*	1.00														
(7)	**BvB gross alpha**	-0.01	0.59*	0.59*	0.66*	0.66*	1.00*	1.00													
(8)	**BvB net value added**	-0.01+	0.15*	0.15*	0.17*	0.17*	0.27*	0.27*	1.00												
(9)	**BvB gross value added**	-0.02*	0.15*	0.15*	0.17*	0.17*	0.27*	0.26*	0.99*	1.00											
(10)	**Annual return**	-0.00	0.48*	0.48*	0.53*	0.53*	0.43*	0.43*	0.11*	0.11*	1.00										
(11)	**Beta_mkt**	0.02*	-0.09*	-0.09*	-0.08*	-0.08*	-0.05*	-0.05*	-0.03*	-0.02*	-0.05*	1.00									
(12)	**Beta_smb**	0.13*	-0.01	0.00	0.00	0.01+	0.02*	0.03*	0.01	-0.01	0.07*	0.08*	1.00								
(13)	**Beta_hml**	-0.02*	-0.04*	-0.04*	0.05*	0.05*	-0.03*	-0.04*	-0.02*	-0.02*	-0.04*	-0.15*	0.07*	1.00							
(14)	**Beta_umd**	0.02*	-0.08*	-0.08*	0.11*	0.11*	0.05*	0.05*	0.02*	0.02*	-0.01	0.11*	0.08*	0.04*	1.00						
(15)	***R^2***	-0.21*	-0.02*	-0.03*	-0.06*	-0.07*	-0.02*	-0.04*	-0.01+	-0.00	-0.02*	0.37*	0.02*	-0.13*	0.02*	1.00					
(16)	**ln(TNA)**	-0.35*	0.06*	0.04*	0.06*	0.05*	0.11*	0.08*	0.02*	0.10*	0.08*	0.03*	-0.12*	-0.01*	0.01*	0.11*	1.00				
(17)	**Sdmret**	0.18*	0.01+	0.02*	0.03*	0.04*	0.01*	0.03*	-0.00	-0.01*	-0.08*	0.43*	0.26*	-0.23*	-0.04*	0.00	-0.13*	1.00			
(18)	**Fund age**	-0.09*	-0.02*	-0.03*	-0.04*	-0.04*	-0.02*	-0.03*	0.00	0.04*	0.01*	-0.01*	-0.11*	-0.07*	-0.03*	0.11*	0.35*	-0.11*	1.00		
(19)	**Perf. sens.**	0.09*	0.00	0.01+	0.01+	0.01*	0.03*	0.04*	0.00	0.00	-0.02*	-0.07*	0.08*	0.06*	0.01*	-0.19*	0.04*	-0.01+	-0.12*	1.00	
(20)	**Turnover**	0.13*	-0.01+	-0.00	-0.00	0.01	-0.03*	-0.02*	-0.00	-0.01*	-0.03*	0.01*	0.07*	-0.06*	0.05*	-0.06*	-0.14*	0.14*	-0.06*	-0.01*	1.0
(21)	**ln(MgmtComp TNA)**	-0.26*	0.06*	0.04*	0.09*	0.08*	0.08*	0.05*	0.02*	0.06*	-0.02*	0.10*	-0.14*	-0.04*	-0.03*	0.11*	0.58*	0.01*	0.18*	0.03*	-0.0

Table II. Alpha Regression

This table reports the results of Fama-Macbeth regressions in which yearly net alphas are regressed on contemporaneous and lagged expense ratios, lagged net alphas (which correspond to a model's dependent variable), and the following lagged fund characteristics (see Table A in the Data Appendix for a detailed description of the variables): market beta, SMB beta, HML beta, UMD beta, R2 of the Fama-French-Carhart four factor model, TNA, standard deviation of monthly fund returns, fund age, flow-performance sensitivity, turnover, size of the management company, the relative size of the institutional share class, and the relative size of open share classes. For brevity, coefficients associated with the additional controls are not reported in the table. The data covers the period from 1980 to 2017 and is a yearly panel. We split the sample into pre- and post-1999 sub-periods because several variables are only available after 1999. Coefficient estimates are time series averages of cross-sectional regression coefficients obtained from annual cross-sectional regressions. Values in parentheses are t-statistics.

	Pre-1999				Post-1999			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	BvB Net Alpha		CAPM Net Alpha		BvB Net Alpha		CAPM Net Alpha	
Expense ratio $_t$	–1.310**		–1.676***		–0.783***		–1.213***	
	(–2.883)		(–3.206)		(–4.524)		(–3.623)	
Expense ratio $_{t-1}$		–0.955*		–1.046**		–0.631***		–1.041***
		(–2.007)		(–2.426)		(–4.070)		(–3.065)
Alpha $_{t-1}$	0.161***	0.169***	0.133**	0.143***	0.115***	0.118***	0.098**	0.102***
	(4.047)	(4.183)	(2.888)	(3.062)	(3.393)	(3.499)	(2.870)	(3.018)
Additional Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N. Obs.	11,341	11,341	11,340	11,340	47,459	47,477	47,459	47,477
Avg. Rsq	0.260	0.256	0.414	0.409	0.191	0.189	0.438	0.437
N. of Years	18	18	18	18	19	19	19	19

Table III. Base-Case Fund Expense Regressions

This table reports the results of Fama-MacBeth regressions in which yearly expense ratios are regressed on lagged fund characteristics (see Table A in the Data Appendix for a detailed description of the variables). We standardize all the independent variables to mean 0 and standard deviation 1 using each variable's yearly cross-sectional mean and standard deviation. The data covers the period from 1980 to 2017 and is a yearly panel. All variables are lagged by one year. Expense ratio (the dependent variable in all regressions) is in basis points. We split the sample into pre-and post-1999 sub-periods since several variables are only available after 1999. The specifications reported in this table represent the base-case specifications. Coefficient estimates are time series averages of cross-sectional regression coefficients obtained from annual cross-sectional regressions. Values in parentheses are t-statistics. We perform the regressions on S&P 500 index funds (post-1999 only, Panel A), all *Index* funds (post-1999 only, Panel B), the *All* funds sample of active mutual funds (Panel C), and on the largest quintile (*Large*) of annually ranked *TNA* funds (Panel D).

	Panel A. S&P 500 Index Funds			Panel B. All Index Funds		
	1999–2017			1999–2017		
	(1)	(2)	(3)	(1)	(2)	(3)
Annual return	–0.240			1.704		
	(–0.207)			(1.158)		
CAPM gross alpha		2.055			1.721	
		(1.232)			(1.096)	
CAPM gross value added			–0.197			0.127
			(–0.452)			(0.291)
Beta_mkt	2.036	1.844	3.642	2.666*	3.951***	3.539**
	(1.064)	(0.908)	(1.234)	(1.778)	(2.976)	(2.389)
Beta_smb	2.862	2.897	3.572**	2.640**	3.060**	2.635**
	(1.464)	(1.357)	(2.125)	(2.140)	(2.302)	(2.338)
Beta_hml	–1.245	–1.263	1.234	–1.037	–1.724	–2.873**
	(–0.538)	(–0.506)	(0.634)	(–0.789)	(–1.225)	(–2.332)
Beta_umd	3.806*	3.400*	2.910	–1.854**	–2.045***	–1.510**
	(1.939)	(1.840)	(1.387)	(–2.846)	(–3.460)	(–2.362)
R^2	–8.234***	–7.958***	–8.210***	–9.583***	–9.712***	–10.238***
	(–3.974)	(–3.773)	(–3.962)	(–10.464)	(–10.442)	(–12.687)
ln(TNA)	–6.654***	–6.880***	–7.073***	–5.445***	–5.379***	–5.484***
	(–7.990)	(–9.328)	(–8.245)	(–3.464)	(–3.383)	(–3.350)
Sdmret	0.380	0.927	–1.241	–1.030	–1.416	–0.636
	(0.161)	(0.396)	(–0.554)	(–0.564)	(–0.756)	(–0.336)
Fund age	2.288***	2.229***	2.151***	1.317***	1.284***	1.266***
	(7.603)	(7.778)	(5.569)	(3.340)	(3.264)	(3.412)
Perf. sens.	1.194	1.339	1.308	0.538	0.602	0.531
	(1.184)	(1.404)	(1.251)	(0.565)	(0.623)	(0.565)
Turnover	5.836***	5.951***	5.886***	13.936***	13.971***	13.874***
	(3.283)	(3.350)	(3.261)	(12.719)	(13.155)	(12.958)
ln(MgmtCompTNA)	–9.255***	–8.878***	–9.206***	–17.853***	–17.899***	–18.143***
	(–11.415)	(–12.110)	(–11.484)	(–11.668)	(–11.680)	(–10.601)
Family100 dummy	14.524***	14.429***	14.466***	16.483***	16.609***	16.488***
	(10.100)	(10.184)	(10.278)	(11.623)	(11.547)	(11.798)
Institutional	–6.863***	–6.896***	–6.831***	–7.376***	–7.402***	–7.396***
	(–10.805)	(–10.911)	(–11.252)	(–9.967)	(–10.016)	(–9.677)
Open	–0.203	0.790	0.027	0.931**	0.976**	0.787***
	(–0.419)	(0.740)	(0.054)	(2.840)	(2.776)	(3.271)
Intercept	27.885***	27.927***	27.905***	40.831***	40.801***	40.817***
	(17.019)	(17.034)	(16.898)	(26.288)	(26.288)	(26.232)
N. Obs.	1,292	1,292	1,292	6,658	6,657	6,657
Avg. Rsq	0.652	0.654	0.648	0.632	0.633	0.629
N. of Years	19	19	19	19	19	19

Panel C. All Funds Sample of Active Mutual Funds

	1981–1998			1999–2017		
	(1)	(2)	(3)	(4)	(5)	(6)
Annual return	–3.157*			–1.826		
	(–1.784)			(–1.442)		
CAPM gross alpha		–3.149			–1.251	
		(–1.698)			(–1.067)	
CAPM gross value added			0.281			0.625*
			(0.383)			(2.005)
Beta_mkt	5.701***	5.473***	8.339***	2.532	2.912*	3.223*
	(2.920)	(3.175)	(4.393)	(1.604)	(1.755)	(1.996)
Beta_smb	1.539	1.242	2.317	2.994**	2.946**	2.928**
	(0.862)	(0.668)	(1.384)	(2.545)	(2.519)	(2.700)
Beta_hml	–1.556	–1.582	–2.834**	–0.960	–1.031	–2.669**
	(–1.708)	(–1.629)	(–2.127)	(–0.998)	(–1.124)	(–2.770)
Beta_umd	2.903*	2.960*	3.697*	–2.536**	–2.331**	–2.210**
	(1.825)	(1.759)	(1.970)	(–2.623)	(–2.296)	(–2.230)
R^2	–13.853***	–13.751***	–15.659***	–11.389***	–11.329***	–11.097***
	(–7.325)	(–7.091)	(–7.929)	(–12.919)	(–12.797)	(–11.692)
ln(TNA)	–27.802***	–27.816***	–28.814***	–13.536***	–13.564***	–13.981***
	(–20.750)	(–20.791)	(–20.757)	(–5.485)	(–5.484)	(–5.553)
Sdmret	1.040	1.114	–1.630	7.418***	7.407***	6.899***
	(0.436)	(0.457)	(–0.776)	(3.362)	(3.358)	(3.296)
Fund age	–1.219	–1.091	–0.558	4.592***	4.616***	4.666***
	(–0.773)	(–0.684)	(–0.340)	(5.517)	(5.589)	(5.651)
Perf. sens.				5.171***	5.150***	5.242***
				(12.670)	(12.541)	(12.460)
Turnover				3.779***	3.802***	3.810***
				(8.767)	(8.745)	(8.555)
ln(MgmtCompTNA)				–21.290***	–21.266***	–21.233***
				(–9.037)	(–9.043)	(–9.065)
Family100 dummy				23.563***	23.565***	23.562***
				(9.025)	(9.014)	(8.937)
Institutional				–11.079***	–11.077***	–11.049***
				(–10.728)	(–10.761)	(–10.775)
Open				–1.835***	–1.831***	–1.792***
				(–3.562)	(–3.548)	(–3.485)
Intercept	129.475***	129.475***	129.419***	118.005***	118.006***	118.007***
	(38.301)	(38.301)	(38.164)	(41.455)	(41.485)	(41.633)
N. Obs.	11,442	11,442	11,293	47,485	47,485	47,484
Avg. Rsq	0.330	0.331	0.330	0.334	0.334	0.332
N. of Years	18	18	18	19	19	19

Panel D. Sample of Large Mutual Funds

	1981–1998			1999–2017		
	(1)	(2)	(3)	(4)	(5)	(6)
Annual return	1.763			2.127**		
	(1.269)			(2.438)		
CAPM gross alpha		1.322			2.308***	
		(0.955)			(3.364)	
CAPM gross value added			1.357			0.734
			(1.203)			(1.484)
Beta_mkt	3.759	3.218	3.582	6.033***	7.056***	6.912***
	(1.289)	(1.101)	(1.130)	(4.220)	(5.294)	(4.994)
Beta_smb	2.214	1.976	1.934	5.037***	5.125***	5.354***
	(1.702)	(1.522)	(1.330)	(4.927)	(5.015)	(5.362)
Beta_hml	2.299	2.665	1.200	–2.202**	–2.255**	–2.660***
	(1.152)	(1.250)	(0.499)	(–2.307)	(–2.217)	(–2.919)
Beta_umd	6.526***	6.250***	5.263***	–0.556	–0.711	–0.473
	(4.949)	(4.877)	(3.598)	(–0.743)	(–0.866)	(–0.591)
R^2	–6.613***	–6.589**	–6.653**	–8.534***	–8.507***	–8.683***
	(–2.928)	(–2.759)	(–2.752)	(–7.530)	(–7.445)	(–7.732)
ln(TNA)	–5.607***	–5.608***	–5.870***	0.567	0.569	0.120
	(–6.210)	(–6.331)	(–5.450)	(0.393)	(0.394)	(0.082)
Sdmret	0.822	1.687	0.206	–1.635	–1.582	–1.612
	(0.251)	(0.494)	(0.058)	(–1.096)	(–1.062)	(–1.043)
Fund age	–7.333***	–7.280***	–6.939***	–1.361	–1.376	–1.398
	(–8.937)	(–8.728)	(–7.688)	(–1.603)	(–1.624)	(–1.625)
Perf. sens.				5.324***	5.335***	5.401***
				(6.255)	(6.296)	(6.235)
Turnover				3.825***	3.850***	3.856***
				(9.660)	(9.681)	(10.139)
ln(MgmtCompTNA)				–21.413***	–21.396***	–21.307***
				(–21.047)	(–21.104)	(–20.962)
Family100 dummy				27.358***	27.327***	27.121***
				(14.019)	(13.982)	(13.905)
Institutional				–5.608***	–5.590***	–5.609***
				(–5.629)	(–5.628)	(–5.705)
Open				–0.204	–0.188	–0.210
				(–0.562)	(–0.519)	(–0.570)
Intercept	93.689***	93.689***	93.500***	82.735***	82.755***	82.885***
	(27.556)	(27.556)	(27.271)	(28.837)	(28.872)	(28.983)
N. Obs.	2,282	2,282	2,272	9,491	9,491	9,491
Avg. Rsq	0.320	0.321	0.310	0.415	0.415	0.413
N. of Years	18	18	18	19	19	19

Table IV. Residual Expense Spreads

This table reports time-series averages of the cross-sectional residual and reported expense spreads, reported in basis points (bps). The data covers the period of 1980 to 2017 and is a yearly panel (for S&P 500 index funds and all index funds, the sample period is 1999 to 2017). The variables are defined in the Data Appendix. We report these statistics separately for the sample of S&P 500 index funds, the sample of all *Index* funds, and the *All* funds sample of active US-equity mutual funds, and for the *Large* funds (top TNA quintile). Panels A reports the residual spreads using the base-case pricing models of Table III. Panel B reports the residual spreads using the extended pricing models of Table V.

Panel A. Distribution of Observed and Residual Expenses (based on Table III expense models)

	25th to 75th Percentile	**10th to 90th Percentile**	**1st to 99th Percentile**
	Mean Reported Expense Spread (bps)		
S&P 500 (post-1999 only)	30	61	143
Index Funds (post-1999 only)	39	67	193
All Funds	59	117	286
Largest-TNA-Funds	42	87	172
	Mean Residual Expense Spread (bps)		
S&P500 (post-1999 only)	18	36	87
Index Funds (post-1999 only)	23	47	123
All Funds	47	98	225
Largest-TNA-Funds	33	70	143

Panel B. Distribution of Residual Expenses (based on extended fund expense models of Table V)

	25th to 75th Percentile	**10th to 90th Percentile**	**1st to 99th Percentile**
	Mean Residual Expense Spread (bps)		
Index Funds (post-1999 only)	23	47	119
All Funds	49	101	223
Largest-TNA-Funds	35	72	146

Table V. Extended Fund Expense Regressions

This table reports the results of Fama-MacBeth regressions in which yearly expense ratios are regressed on lagged fund characteristics (see Table A in the Data Appendix for a detailed description of the variables). We standardize all the independent variables to mean 0 and standard deviation 1 using each variable's yearly cross-sectional mean and standard deviation. The data covers the period from 1991 (we lose the first few years due to data limitations when calculating parameters "a" and "b" below) to 2017 and is a yearly panel. All variables are lagged by one year. Expense ratio (the dependent variable in all regressions) is in basis points. We estimate the regressions for the *All* funds sample of active US-equity mutual funds, the *Large* funds (top TNA quintile), and *Index* funds. We split the sample into pre-and post-1999 sub-periods because several variables are only available after 1999. Coefficient estimates are time series averages of cross-sectional regression coefficients obtained from annual cross-sectional regressions. Values in parentheses are t-statistics. The specifications extend the base case fund pricing regressions discussed in Table III by adding several fund characteristics following Berk and van Binsbergen (2015).

	All Funds		Large Funds		Index Funds
	pre-1999	post-1999	pre-1999	post-1999	post-1999
BvB gross alpha	–2.013	–0.833	4.092**	–1.057	1.357
	(–0.716)	(–0.919)	(3.167)	(–1.732)	(1.644)
BvB gross value added	0.949	0.290	–0.393	–0.109	–0.196
	(1.189)	(1.458)	(–0.405)	(–0.439)	(–0.535)
a	–9.403*	–4.646**	–3.499***	4.721***	1.472
	(–2.242)	(–2.391)	(–5.423)	(4.897)	(1.316)
b	13.469***	5.862***	6.378***	2.695***	–0.713
	(5.451)	(5.352)	(7.865)	(6.730)	(–0.600)
Skill ratio	2.825	5.599***	–1.747	2.348***	–0.778
	(1.266)	(5.153)	(–1.683)	(5.871)	(–1.077)
ln(TNA)	6.041***	2.821	9.219**	5.020***	3.470**
	(5.647)	(1.731)	(3.411)	(3.911)	(2.772)
Beta_mkt	5.004**	2.901**	4.298**	3.649***	2.335*
	(2.367)	(2.732)	(2.721)	(3.779)	(2.056)
Beta_smb	–3.532*	–2.187*	0.026	–2.181**	–2.141*
	(–2.293)	(–2.071)	(0.012)	(–2.463)	(–1.982)
Beta_hml	3.468*	–2.411**	5.605***	–0.146	–1.520**
	(2.158)	(–2.385)	(3.559)	(–0.186)	(–2.648)
Beta_umd	–19.817***	–11.375***	–12.780***	–7.339***	–9.756***
	(–10.659)	(–11.926)	(–8.548)	(–8.885)	(–12.828)
R^2	–14.715***	–10.347***	–0.702	–0.396	–5.172***
	(–9.299)	(–3.610)	(–1.277)	(–0.239)	(–2.898)
Sdmret	2.122	7.057***	–3.408	0.357	–1.173
	(0.778)	(3.540)	(–1.083)	(0.240)	(–0.613)
Fund age	0.786	4.771***	–7.623***	–1.248	1.379***
	(0.313)	(5.881)	(–9.058)	(–1.549)	(4.464)
Perf. sens.		5.416***		4.446***	0.403

	All Funds		Large Funds		Index Funds
	pre-1999	post-1999	pre-1999	post-1999	post-1999
		(14.028)		(6.844)	(0.406)
Turnover		3.618***		3.980***	13.628***
		(7.386)		(9.583)	(13.664)
ln(MgmtComp TNA)		–20.229***		–21.397***	–18.199***
		(–8.591)		(–19.679)	(–9.469)
Family100 dummy		23.501***		27.946***	16.598***
		(9.645)		(14.800)	(11.186)
Institutional dummy		–10.888***		–5.861***	–7.262***
		(–10.450)		(–5.968)	(–8.529)
Open dummy		–1.592***		0.652*	0.742***
		(–3.275)		(1.987)	(3.386)
Constant	136.498***	117.989***	104.035***	82.306***	40.781***
	(78.018)	(40.448)	(36.763)	(27.668)	(25.853)
N. of Obs.	8,564	47,481	1,710	9,491	6,656
Avg. Rsq	0.291	0.349	0.306	0.454	0.643
N. of Years	8	19	8	19	19

Table VI. Means of BVB gross and net value-added of the ex-ante and ex-post distributions (following Berk and van Binsbergen, 2015).

Panel A reports the average annual gross value added estimated according to Berk and van Binsbergen (2015), across expense quintiles. The calculation of the ex-ante and ex-post distributions are based on equations 8 and 9 in Berk and van Binsbergen (2015). The estimated value-added for a given fund that exists for T_i periods is calculated as:

$$\hat{S}_i = \sum_{t=1}^{T_i} \frac{V_{it}}{T_i}$$

Where V_{it} is the realized value-added in year t for fund i. The values reported in the *ex-post distribution* are the cross-sectional average of $\hat{S}_i$, weighted by the number of periods that a fund is in the sample. The values reported in the *ex-ante distribution* are the simple cross-sectional average of $\hat{S}_i$ with no weighting. Panel B reports similar results based on net value added. Panel C reports the sum of annual net value added across all funds and years. Numbers reported in Panels A through C are in millions USD in 2014dollars. Panel D reports statistics related to the misallocated capital measures from to Zhu (2018). Specifically, for *All* funds, we report the fraction of funds whose lifetime assets under management exceed their optimal size, the total sum of this amount of over-allocated capital across all funds, the extent of over-allocated capital in percent of TNA for the median fund, and, finally, we report underinvested fund capacity, which is the sum of all funds with negative misallocated capital.

Panel A. Annual BvB Gross Value Added

		Ex-post distribution			Ex-ante distribution		
		All	Large	Index	All	Large	Index
	All funds	11.57***	56.91***	10.05***	5.582***	37.16***	6.319***
Fee Quintile	1	30.57***	67.96***	44.93***	15.95***	30.73***	30.59***
	2	11.99***	102.0***	2.041**	5.736***	62.78***	2.254
	3	7.968***	50.14***	1.018	4.196***	37.74***	–2.602
	4	4.614***	33.78***	–1.701*	1.723**	35.55***	0.204
	5	2.542***	30.04***	1.544***	0.301	18.61***	0.593

Panel B. Annual BvB Net Value Added

		Ex-post distribution			Ex-ante distribution		
		All	Large	Index	All	Large	Index
	All funds	–1.929***	4.166***	2.873***	–2.777***	0.589	0.896
Fee Quintile	1	8.416***	19.21***	29.88***	2.225	–1.339	19.53**
	2	–3.657***	40.12***	–4.347***	–3.872***	23.67**	–3.292
	3	–5.419***	–1.504	–3.774***	–4.342***	–0.369	–5.950
	4	–5.206***	–17.33***	–5.494***	–4.532***	–1.512	–2.397*
	5	–3.868***	–20.28***	–3.803***	–3.365***	–17.86**	–3.863**

Panel C. Sum of Annual BvB Net Value Added

		All	Large	INDEX
	All funds	–124,650	53,606	20,966
Fee Quintile	1	145,409	10,436	41,782
	2	–72,716	162,516	–3,635
	3	–87,457	–25,068	–6,432
	4	–68,163	–52,909	–5,997
	5	–41,723	–41,369	–4,752

Panel D. Misallocated Capital

		Fraction of Overinvested	Sum of Overinvested	Median of Relative Misallocation	Underinvested Fund Capacities
	All funds	70%	1,432,609	61%	–2,240,648
Fee Quintile	1	74%	665,134	64%	–461,816
	2	71%	325,932	59%	–450,362
	3	65%	220,576	50%	–545,374
	4	68%	140,822	57%	–430,587
	5	72%	80,146	72%	–352,508

Figure 1. Fund Expense Dispersion

These figures show the dispersion of reported expense ratios (left column) and residual expense ratios (right column) across funds and over time. The graphs show the ranges between the 25th and 75th (darkest grey), 10th and 90th (medium dark grey), and 1st and 99th percentile (light grey) points of the distributions. Graphs in the top row, for the *All* Funds sample, also plot the aggregate TNA of all funds in the sample in billions of USD (solid line). In rows 2 (*Large* funds - top TNA quintile funds) and 3 (*Index* funds), the solid line represents the fraction of aggregate TNA represented by funds in those respective sub-samples. The residual expenses are defined as the regression residuals of the expense models specified in Table III. Our sample consists of equity mutual funds. The data covers the period from 1980 to 2017 and is a yearly panel. The left column y-axis is annual reported expenses in basis points (bps) and the right column y-axis is annual residual expenses in basis points (bps).

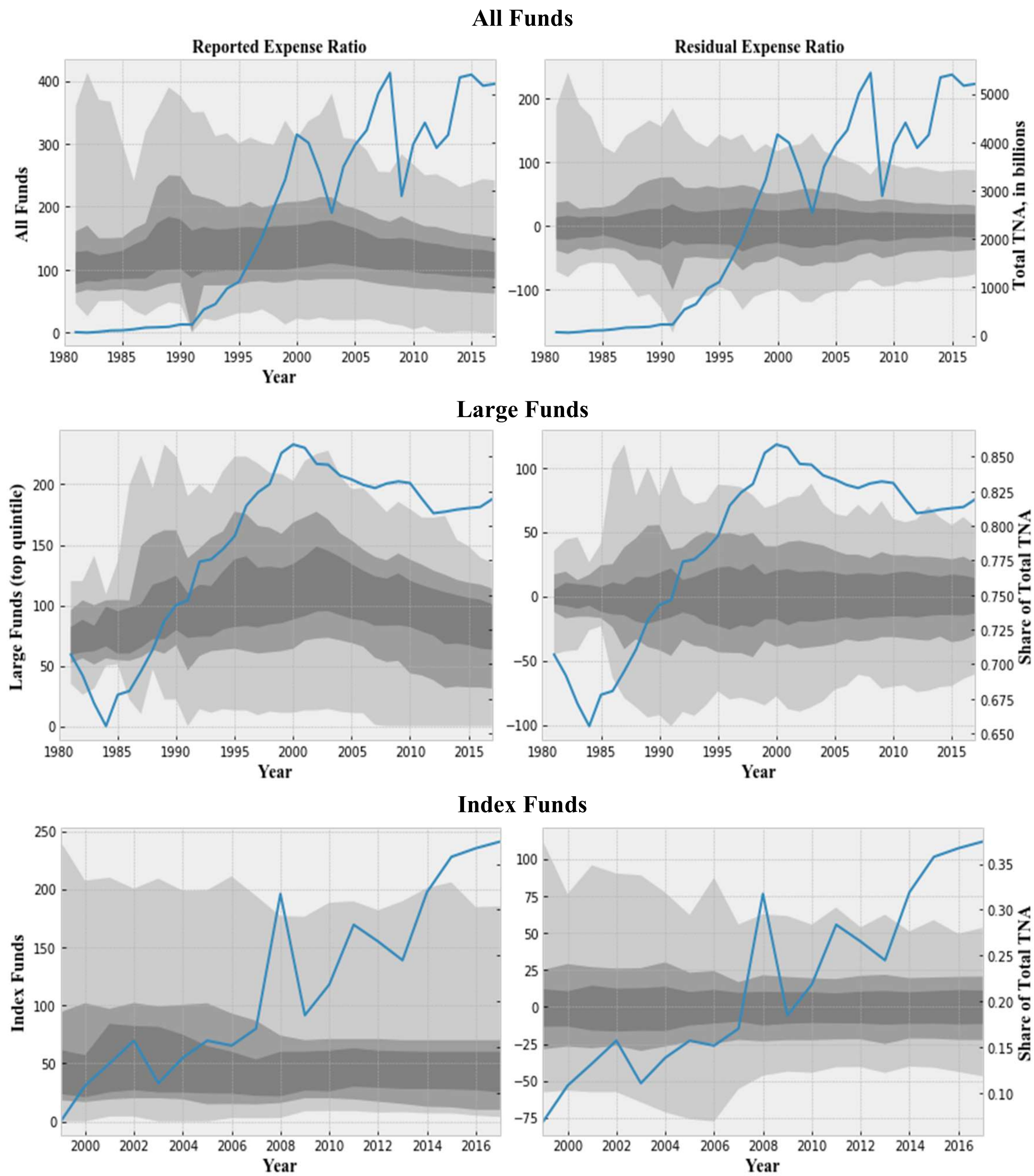
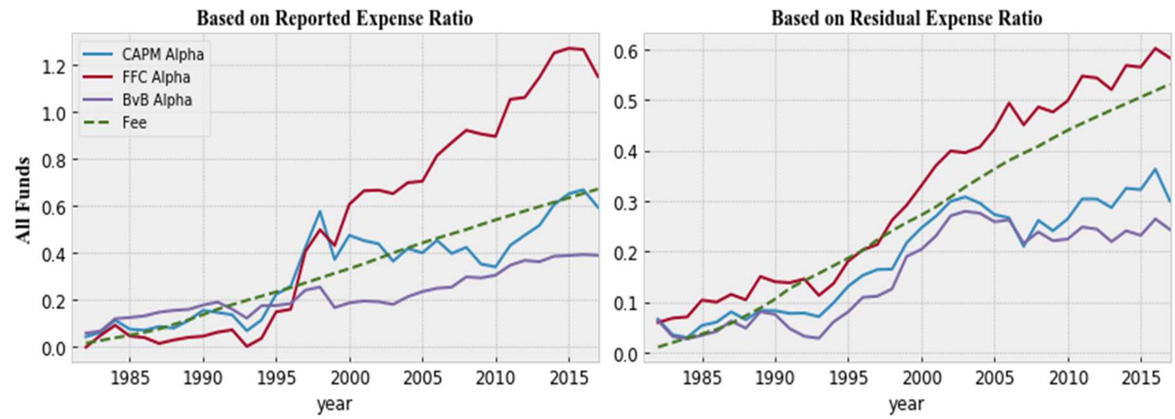


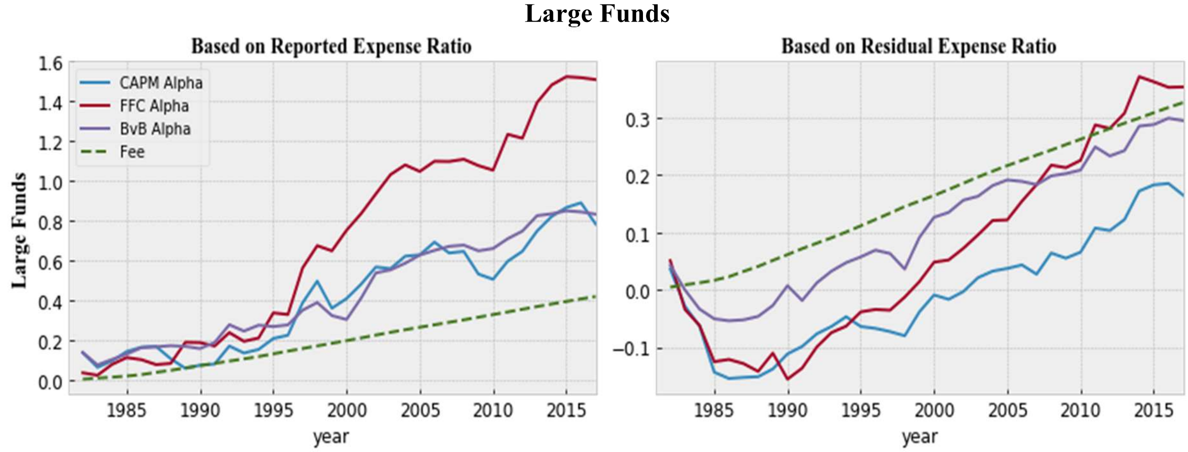
Figure 2. Evaluation of Trading Strategy

These figures show the cumulative alphas (CAPM net alpha in blue, Fama-French-Carhart net alpha in red, and BvB net alpha in purple) of a strategy that buys funds in the bottom quintile of reported expense ratios (residual expense ratios) and shorts funds in the top quintile of reported expense ratios (residual expense ratios). The figures also report the cumulative spread between average reported expense ratios (residual expense ratios) of funds in the top and the bottom quintile (the dashed line). The residual expenses are defined as the regression residuals of the expense models specified in Table III. The top panel (labeled *All* funds) include all the active equity funds in our sample, the second panel (labeled *Large* funds) includes the funds in the top quintile in terms of TNA, and the bottom panel includes all the *Index* funds. The data covers the period from 1980 to 2017 and is a yearly panel.

All Funds



Large Funds



Index Funds

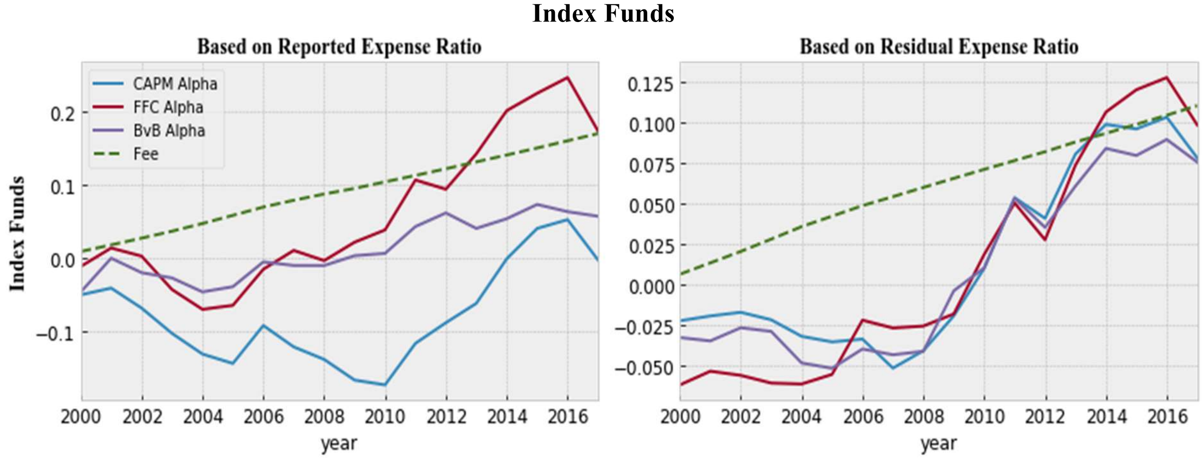


Figure 3. Cross-sectional Dispersion of BvB Net Value Added

These figures show the dispersion of the BvB net value added across funds and over time. The graphs show the ranges between the 25th and 75th (darkest grey), 10th and 90th (medium dark grey), and 5st and 95th percentile (light grey) points of the distributions. The top left graph shows the dispersion of *All* active equity funds. The top right graph shows the dispersion among the *Large* funds, and the bottom graph shows the dispersion among the *Index* funds. The data covers the period from 1980 to 2017 for all active funds and large funds and 1999 to 2017 for index funds. The left y-axis is in millions of dollars (inflation adjusted to 2014).