Fee dispersion and persistence in the mutual fund industry

Michael Cooper, Michael Halling and Michael Lemmon*

July 15, 2011

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

JEL Classifications: G10, G11, G23.

* All authors are with the David Eccles School of Business, University of Utah.

Preliminary and incomplete. We thank Darwin Choi, Rachel Hayes, Keith Jacob, Mark Seasholes and seminar participants at the University of Arizona and the University of Montana for their comments. We welcome additional comments, including references to related papers we inadvertently overlooked.

Fee dispersion and persistence in the mutual fund industry

Abstract

In this paper, we present striking new evidence of systematic differences in prices across *all* US equity funds. We examine the residuals from regressions of fees (annual expenses and 12b-1 fees) on important fund characteristics, essentially allowing us to compare the fees of "identical" funds. We find that the average spread in residual fees across all funds over the sample is approximately 2.3%. The dispersion in fees has not decreased over time, despite the fact that significant numbers of new funds have entered and the aggregate amount of assets under management has increased substantially. An investor purchasing identical lower fee funds would have outperformed an investor purchasing higher fee funds by approximately 32% over our sample. We test a number of hypotheses to explain our results including a random fee, competition, strategic fee setting, service, and captive investor hypothesis, and are able to explain only a small portion of the spread in residual fees. Surprisingly, a main determinant of fees is the initial fee set by a fund, which varies little over time. Overall, our evidence is largely inconsistent with a competitive market for mutual funds.

1. Introduction

A large literature exists that attempts to explain why similar products sell for different prices.[1] For example, Lach (2002) documents considerable price dispersion for similar refrigerators, chicken, coffee, and flour. He concludes that because stores change their pricing on a regular basis, consumers cannot learn which stores are the low cost sellers, and as a consequence, price dispersion persists.

In the mutual fund markets, Elton, Gruber, and Busse (2004) document price dispersion of more than 2% per year for essentially identical S&P500 Index funds.[2] They conclude that a 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&P500 Index fund market. Other papers, focusing on sub-categories of funds, also provide evidence of differential prices being charged for funds with similar characteristics.[3]

In contrast, other papers suggest that the mutual fund markets are more or less competitively priced. For example, Khorana, Servaes, and Tufano (2009) examine mutual fund fees in 18 countries and find that most of the cross-sectional dispersion in fees can be explained by economic variables, such as investment objective, sponsor, national characteristics, and levels of investor protection. More recently, Wahal and Wang (2010) provide evidence that incumbents with high overlap in their portfolio holdings with entrants subsequently engage in price competition by reducing their management fees. In addition, they also find evidence that incumbents with higher portfolio overlap with entrants have lower future fund inflows. They conclude that the mutual fund market has "evolved into one that displays the hallmark features of a competitive market." Overall, while the existing literature provides evidence of price dispersion in specific areas of the mutual fund market, there is little existing evidence on how widespread

[1] See Bakos (2001), Brown and Goolsbee (2002), Brynjolfsson and Smith (2000), Nakamura (1999), Pratt, *et al.* (1979), Scholten and Smith (2002), and Sorensen (2000).
[2] See also Hortacsu and Syverson (2004).
[3] See also Elton, Gruber, and Rentzler (1989) who find that public commodity funds exists that underperform the risk free rate and Christoffersen and Musto (2002) who find a wide dispersion in fees across similar money market funds.

the phenomenon is or on how it has changed over time given the dramatic growth in the mutual fund market.

In this paper, we present striking new evidence of systematic differences in prices across *all* US equity funds. We examine the residuals from regressions of fees (annual expenses and 12b-1 fees) on important fund characteristics, essentially allowing us to compare the fees of "identical" funds. We find that the average spread in residual fees (between the 1st and 99th percentile) across all funds over the sample is 2.34%. More interestingly, the dispersion in expense ratios has not decreased over time, despite the fact that significant numbers of new funds have entered and the aggregate amount of assets under management has increased substantially over time. Our results hold for both the largest total net asset (TNA) funds as well as smaller size funds; the average spread in residual fees is 2.88% for the smallest quintile of TNA funds and is 1.18% for the largest quintile of funds. In fact, for the largest quintile funds, representing 82% of the market value of our sample, the spread in residual fees has actually increased from 1990 to 2009, evidence that is difficult to reconcile with a competitive market for mutual funds.

We examine the implications of our findings for investors. Based on raw (residual) fees, an investor purchasing the lowest fee funds would have earned compounded abnormal returns 67% (32%) higher than an investor purchasing the most expensive funds. As a basis for comparison, the compounded differences in fees (residual fees) over the period were 90% (64%). Thus, while the difference in abnormal returns between high fee and low fee funds is less than the cumulative difference in fees, it appears that investors bear significant costs from investing in high fee mutual funds that are not recouped through higher performance of these funds[4].

We explore various explanations for the large price dispersion across similar funds. We first note that controlling for product characteristics that investors are likely to care about when purchasing a fund,

4 Ramadorai and Streatfield (2011) document that hedge funds with high initial management fees (i.e., non performance sensitive fees) have no better or slightly lower performance than hedge funds with low initial management fees. They conclude that high management fees are "money for nothing" in the hedge fund industry. Our results for the mutual fund industry are stronger in the sense that high-fee mutual funds significantly underperform on average.

such as search costs, service levels, fund size, retail versus institutional funds, fund age, fund flows, fund portfolio characteristics (i.e., lagged performance and style betas) explains about 44% of the dispersion in raw fees, leaving a sizable unexplained dispersion in fees.

We test various hypotheses to explain this dispersion. We first test a "random fee" hypothesis (Lach (2002) and others) which posits that mutual funds engage in frequent changes in their fees, resulting in fund consumers not being able to easily identify low fee funds, which in turn results in persistent price dispersion. We estimate transition probabilities among low and high fee funds and find evidence strongly inconsistent with a random fee explanation. For both raw and residual fees, low fee funds tend to stay low (70% – 87% transition probabilities) and high fee funds tend to stay high (70% - 85%). It is very rare that high fee funds become low (0.7% - 2.4%) or that low fee funds become high (0.6% - 2%)[5].

Given that there is scant evidence for fee switching over time, we examine the role of a fund's initial fee in explaining the cross-section of raw fees. As mentioned above, regression specifications that include standard determinants of fund fees obtain average adjusted R-squareds of approximately 44%. When we add a fund's initial fee to the regressions, the R-squareds of the regressions jump to 56% to 70%. Moreover, including the fund's initial fee substantially reduces the explanatory power of many of the other variables. To shed further light on this issue we study the determinants of initial fees. We find that initial fees are higher for funds that enter with smaller size, smaller fund families, fund families that charge higher average fees, and for families that have a higher dispersion of fees within the fund family. We also find that factor returns (i.e., the market risk premium, SMB, HML, and UMD) are consistently negatively related to first fees. A potential interpretation is that in periods when factors are doing poorly, actively managed funds can charge excess fees. Interactions between factor returns and betas are consistently positively related to first fees, implying that in situations when funds are founded in styles that have been successful in the past, higher fees are charged. Thus, first fees are strongly related to a fund's lifetime expense ratio.

[5] It seems that fees do not change much in the investment services industry, in general. Liang (2000), among others, documents that fees of hedge funds do not change much over time, as well.

We test four more hypotheses in an attempt to gain a better understanding of how funds set fees and how fees evolve over time. We first test the competition hypothesis. Wahal and Wang (2010) show that when existing funds face competition from new, similar funds, the existing funds lower their fees to better compete with the upstart funds. We regress fee changes on important fund characteristics and variables designed to capture the entry of new funds and the fees that those new funds charge relative to incumbent funds. We find statistically insignificant coefficients on our competition variables, suggesting that incumbent funds do not lower their fees when faced with new competing funds. In fact, after 1999, there is a positive (albeit statistically insignificant) effect on existing funds' fees when faced with a new competitor (i.e., as new funds' average fees go up by 1%, existing funds raise their fees by approximately 1 basis point).

Another hypothesis that we test is the strategic fee setting hypothesis. Using a sample of money market funds, Christoffersen and Musto (2002) show that performance sensitive investors withdraw assets from poorly performing funds, leaving only performance insensitive investors as holders of the funds' shares. Funds respond to the fact that the fund flows of the remaining investors are not sensitive to fund performance by raising fees. Gil-Bazo and Ruiz-Verdu (2008) provide some evidence of a similar effect in equity mutual funds. Specifically, they find evidence that funds with lower flow-performance sensitivity charge higher fees in the cross-section. We provide additional evidence on this issue by examining the determinants of fee changes and generally find minimal support for the strategic fee setting hypothesis. First, although lagged flows and returns have negative effects on future fee changes, the effects are economically small. We further examine the flows and changes in fees for funds segmented by good and bad prior performance. We find negative median flows for high-fee poor-performing funds, inconsistent with the strategic fee setting hypothesis that suggests little or no outflow for high-fee poor-performing funds. We also find median increases in fees of approximately zero for poor performing funds, a result that is also inconsistent with the strategic fee setting hypothesis.

Next, we test the fund family service hypothesis. Hortacsu and Syverson (2004), Collins (2005) and others have suggested that variation in services, such as financial advice or complementary investment instruments, may explain fee variation. Assuming that large fund families offer better service, we find that funds that are part of a family with more than 100 funds charge, on average, an extra 12 to 14 basis points in fees, but controlling for large fund families does not alter our finding of large spreads in residual fees. In addition, we find a wide dispersion in residual fees across large fund families, suggesting that even if larger fund companies do offer better service, there is significant heterogeneity in what they charge for such service. Our results do not appear to be explained by differences in the services offered by funds.

Finally, we test a "captive" investor hypothesis. We examine if funds that are likely to inhibit easy investor exit (thus creating captive investors) are the funds with high residual fees. We define captive investors two ways. First, we define funds that are "easy-in, hard-out" funds. These are funds that spend more than other funds on advertising and have higher back-end loads. Second, using a subsample of funds with high flow autocorrelations, we examine the fees of funds that are likely to be held within pension funds, thus not allowing for easy investor exit. We find evidence consistent with the captive investor hypothesis: high residual fee funds have higher back-end loads and spend more on advertising than do lower residual fee funds. Also, high residual fee funds have higher positive autocorrelation of flows than do lower residual fee funds. Overall, the evidence is consistent with fund managers recognizing that some investors may be less likely to switch out of certain funds and charging those investors higher fees.

Overall, our results raise two important questions. First, given that fees are important sources of underperformance, why do funds not manage their fees more actively?[6] Second, why do investors not learn to distinguish cheap from expensive funds? A potential explanation for these questions is that investors may have a difficult time learning about the quality of funds, which allows different funds to

[6] There are several papers that develop theoretical models of the mutual fund industry including endogenous fees. Usually, however, these papers do not focus on the cross-sectional and time-series properties of the equilibrium fees. Nanda, Narayanan and Warther (2000), for example, concentrate on the structure of mutual funds, i.e., on the combination of loads and fees. Das and Sundaram (2002) compare fulcrum fees (legally required for mutual funds in the US) to incentive fees. Pastor and Stambaugh (2010), as another example, use their model to study the aggregate size of the active management market.

charge different prices for delivering a similar product. Of course, if this is the case, one wonders why all funds do not then charge high fees. Overall, we conclude that the large dispersion in prices for similar funds is inconsistent with a competitive market for mutual funds.

The remainder of the paper is organized as follows. In Section I we describe the data used in our analysis and describe the characteristics of high and low fee funds. In Section II we present results that document price dispersion in the residual fee distribution of funds and perform tests to quantify the economic effects of mispricing for fund investors. In Section III we test various hypotheses to explain the apparent large mispricing of funds. Section IV concludes.

2. Data

2.1 Sample Construction

The sample selection follows Pastor and Stambaugh (2002). Accordingly, we select only domestic equity funds and exclude all funds not investing primarily in equities such as money market or bond funds. In addition, we exclude international funds, global funds, balanced funds, flexible funds, and funds of funds. The ICDI classification codes that were used by Pastor and Stambaugh (2002) are, however, no longer available. Thus, we follow Bessler et al. (2008) who use a combination of Lipper codes, Wiesenberger codes and Strategic Insight codes to identify domestic equity funds. Table A in the Appendix lists the specific codes that we use to identify the funds in our sample.[7]

In short, the above screens result in our sample focusing on active and passive US domestic equity funds. Our sample includes approximately 35% of all funds covered in the CRSP Mutual Fund Database (our sample consists of a total of 13817 funds while the CRSP Mutual Fund Database universe has approximately 40000 funds). As measured by total net assets, our sample covers approximately 32% of

[7] Frequently, mutual funds are sold in different share classes that usually differ by distribution-related costs (i.e., sales loads and 12b-1 fees). The CRSP mutual fund database treats these share classes of a fund as different entities and that's what we do as well in this paper. There does not seem to be an automatic way to identify different share classes of the same fund. According to Nanda, Wang and Zheng (2009) one would have to manually extract the share class information from mutual fund names.

the cumulative net assets represented in the database. The sample period spans 1963 to 2008 and the data frequency is yearly, as we focus on fund fees.[8]

2.2 Descriptive Statistics

Table 1 Panel A reports summary statistics of our fund sample. Details of the variable construction can be found in Table B in the Appendix. Throughout the paper 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 and flags for institutional funds 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 the pre-1999 sample the mean number of funds per year is 545, while it increases to 5562 in the post-1999 sample. Note that the mean size of a fund (*mtna*) also increases from 436 Million USD pre-1999 to 464 Million USD post-1999. Thus, 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 decrease in average expense ratios – but this is not the case. Average expense ratios (*exp_rati*) and initial expense ratios of entering funds (*first_exp_rati*) both increased, from 1.3% to 1.4% and 1.3% to 1.5%, respectively. It is also interesting to observe that average yearly changes of expense ratios (*Δ(exp_rati)*) are on average zero (with a tiny standard deviation of 18 bp pre-1999 and 9 bp post-1999).

The average performance (*yalpha*) of our sample funds, as measured by the yearly 4-Factor alphas, is slightly negative, consistent with Carhart (1997) and others who show that funds do not earn positive abnormal returns net of fees. The average fund, over both time periods, has a market beta (*beta_mkt*) that is slightly less than 1, a small, negative exposure to HML (*beta_hml*), and small positive exposures to SMB (*beta_smb*) and UMD (*beta_umd*). After 1999, funds load more on the market, and less on HML,

[8] Quarterly data on fees and other fund characteristics is only available starting in 1999.

SMB, and UMD, consistent with an aggregate strategy shift of market indexing. The 4-Factor model works very well on average in explaining fund returns, yielding R-squareds (*r_squared*) of 84% and 87%, respectively.

Panel B (pre-1999 sample) and Panel C (post-1999 sample) of Table 1 report summary statistics by expense ratio deciles. Each year we split all funds into deciles by their expense ratios and then report contemporaneous means and standard deviations of fund characteristics. The last column looks at differences between the means of Decile 1 and 10.

Average expense ratios of Decile 10 exceed those of Decile 1 by 2.1%, in both the pre-1999 and post-1999 periods. In the pre-1999 sample, average expense ratio changes are most negative (-6 bp) in Decile 1 and most positive (15 bp) in Decile 10. These mean changes become even smaller in the post-1999 sample: funds in the bottom expense ratio decile decrease their fees on average by 1 bp in the same year, while funds in the top decile increase their fees on average by 3 bp in the same year.

All of the fund performance variables decrease monotonically by expense ratio deciles. The spread in yearly 4-Factor alphas, for example, equals 2.7% pre-1999 and 2% post-1999, which in both cases basically equals the spread in expense ratios. Thus, these simple descriptive statistics suggest that funds with higher expense ratios on average underperform their cheaper competitors by approximately their expense ratios (consistent with Berk and Green (2004)). Even given the limitations of this simple analysis it seems difficult to justify the expense ratios of the funds in the top deciles of the expense ratio distribution.

Of course, the most important limitation of this univariate analysis is that it completely ignores that expense ratios may reflect different fund strategies and characteristics. This is something that we will explore 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 returns similar to small cap, value, and momentum styles of investing (as judged by their loadings on the

SMB, HML, and UMD factors, respectively). Similarly, the average R-squared of the 4-Factor model decreases as we move from Decile 1 to Decile 10, suggesting that the managers of the higher fee funds may be following "unique" strategies, likely in an attempt to outperform. However, these managers also trade much more (the turnover (*turn_rat*) is much higher for the high fee funds relative to the low fee funds), which may contribute to their low return performance. 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 also implement strategies that go beyond the standard risk factors.

We also find that average funds in Decile 1 are much larger than average funds in Decile 10 suggesting that economies of scale play a role for expense ratios. The average fund in Decile 1 is approximately 1.5 Billion USD larger in both the pre-1999 and post-1999 periods than the average fund in expense ratio Decile 10. We also find that funds which are part of a larger fund family on average have lower fees.[9] This result is potentially consistent with an economies-of-scope argument. Moreover, we also find that institutional funds and ETFs have lower fees, as one would expect.

Finally, Panel D of Table 1 shows time-series means of cross-sectional correlations between fund characteristics. These correlations are consistent with our previous interpretations of patterns between expense ratio deciles and other fund characteristics. In general, none of these correlations seem to be high enough to cause worries about multi-collinearity problems in the subsequent multivariate analysis.

3. The Pricing of Mutual Funds

3.1 Residual Fee Estimation

Our goal is to compare prices (annual 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. Thus, we follow Lach (2002) and Sorensen (2000) to control for fund heterogeneity. As controls we use

[9] This pattern, however, is non-monotonic across raw fee deciles. In later cross-sectional tests we find that large families actually tend to charge greater fees.

the standard fund characteristics that have been shown to be important in determining fund fees (see for example Gil-Bazo and Ruiz-Verdu (2009) and Wahal and Wang (2010)).

We regress fund fees 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 fees. The residuals of these regressions can be interpreted as deviations of fund fees from expected fees given the set of characteristics used in the regression. Thus, using the residuals, we can compare prices across "identical" funds, under the assumption that we have controlled for the correct fund characteristics. Later in the paper we perform robustness tests on the characteristics used to estimate the residuals, and show that our results are qualitatively similar.

3.2 Results

In Table 2 we present the results of the yearly cross-sectional regressions used to estimate the residuals. The reported coefficients are time series averages of cross sectional regression betas obtained from the annual cross sectional regressions. Specification 1 serves as our basic model for the pre-1999 period and specification 4 serves as our basic model for the post-1999 period. Both models explain approximately 44% of the variation in fees. The signs of the coefficients are consistent with the literature (note that it is not the goal of this paper to interpret these relationships): e.g., across the two periods, better performing funds, less volatile funds, larger funds, older funds, lower turnover funds, institutional funds and ETFs, and funds with higher R-squareds from the Carhart four-factor model have lower fees. In the post-1999 period, we essentially see the same relationships, with the exception of some sign switching of the coefficients from the four factor model.

Our main point of interest, the spread in residual fees, is presented in Table 3 and in Figure 1. In the figure, we plot, across all funds in the sample each year, the residual fee spread between the 1% and 99% point of the distribution (note that the mean residual is, of course, zero) and the raw fee spreads. We do

this for the full sample, and for the bottom and top quintile of funds based on a yearly TNA sort.[10] Given that we have controlled for fund characteristics that investors care about in their selection of funds, the residual fee figures are striking. Essentially, these figures show that there exist huge dispersions in fees for basically identical funds. For the full sample, the fee dispersion is large and variable in the 1970-1990 period, with spreads ranging between 2 to 4%. After 1990, the spreads stabilize at approximately 2%. Overall, as reported in Table 3, Panel A, the mean spread for the basic fee model for the full sample (see the row labeled "yearly return") is 2.34%. The figures also plot the growth in TNA. We see a clear pattern of enormous growth in the fund industry, but no decrease in the residual fund spread, results apparently at odds with a competitively priced market. In fact for the larger funds ("top quintile (size) funds"), we actually see an increase in the residual fee spread; in pre-1990, the average spread is approximately 0.5% to 1%, and in the post-2000 period, it grows to an average of approximately 2%. We note that the largest quintile of funds represent 82% of the market value of our sample, suggesting market wide mispricing effects that are not confined solely to the smaller funds.

Our results are robust to different models for calculating the fee residuals. In Table 3, rows 2 through 4, we report mean spreads for models that include a persistence dummy, and two measures of the Carhart four-factor alpha. The results are strongly robust to variations in these fund performance characteristics,4(s)7.83545( )-7
f-0.95847(n)-0.958493( )-0.4785(e)-2.50657(r)-4.60238( )-0.47927()-5.15007(o)-0.954593( )-22.259gefn n ttilcd, tus iis e tos

(residual) fees and sells funds in the top fee decile. We rebalance these portfolios every year and compute the cumulative Carhart four factor model alphas over the sample period to equally-weighted portfolios.

The results are reported in Table 4 and Figure 4. Interestingly, in Panel B of Figure 4 and in Table 4, we observe that until the beginning of the 80ties, investors actually benefited from investing in higher residual fee funds, suggesting that managers of such funds were able to "earn their keep." However, over the entire sample, based on raw (residual) fees, an investor purchasing the lowest fee funds would have earned compounded abnormal returns 67% (32%) higher than an investor purchasing the most expensive funds. As a basis for comparison, the compounded differences in fees (residual fees) over the period were 90% (64%). Thus, while the difference in abnormal returns between high fee and low fee funds is less than the cumulative difference in fees, it appears that investors bear significant costs from investing in high fee mutual funds that are not recouped through higher performance of these funds.

Overall, our finding of large pricing differences for essentially identical products across all US equity funds is a new finding with wide-spread implications for both fund investors and for our understanding of how prices are set in the mutual fund industry. Next we explore various hypotheses, motivated from the price dispersion and mutual fund literature, in an attempt to gain a better understanding of the pricing mechanisms at work in the mutual fund industry.

4. Explaining the Dispersion in Residual Fees

4.1 The Random Fee Hypothesis

If buyers and sellers are identical and in a world with perfect information (i.e., search is costless), the unique Nash equilibrium is the perfectly competitive price. If we relax the assumption of perfect information (i.e., if search becomes costly), then the equilibrium price would be the monopoly price (see Diamond (1971)). Price dispersion in equilibrium, thus, requires some heterogeneity: either sellers (e.g., differences in their production costs) and/or buyers (e.g., differences in search costs or frequency of purchase) differ (see Lach (2002)).

In this context, it is interesting that we find large fee dispersion in mutual funds even after controlling for many fund characteristics; either we are missing important dimensions of fund heterogeneity, or the price dispersion is driven by heterogeneity among fund investors. Our empirical results, however, highlight another important characteristic of fee dispersion, namely its persistence over time.

From a theoretical point of view, persistent price dispersion requires additional assumptions. Moving away from a static equilibrium concept challenges the previously mentioned explanations of

fee funds become high (0.6% - 2%). These results are strongly inconsistent with a random fee explanation.

Note that entry of new funds might potentially bias these transition probabilities, as entering funds might observe the fee distribution of existing funds and chose their fee level accordingly. We want to make sure that entering funds are not causing or distorting the patterns that we observe in the data. Thus, we adjust the thresholds to construct the quintiles such that they only consider funds that were around in the prior year (i.e., "old" funds). Table 5, Panels C and D summarize these results. The results are qualitatively similar to the full sample. Thus, the results in Table 5 clearly document that on average funds do not migrate much between fee quintiles.

Given that we do not find evidence for fee switching over time, we examine the role of a fund's initial fee in explaining the cross-section of raw fees. We re-estimate the fee regressions from Table 2, adding a fund's initial fee (*First_exp_rati*) to the models. The results are presented in specifications 2, 3, 5 and 6. The coefficient on the first fee is positive and highly statistically significant in all models and is economically large, especially compared to the other explanatory variables. Adding the first fee to the models increases the adjusted R-squareds from approximately 44% to approximately 57% (pre-1999) and 70% (post-1999). Moreover, including the fund's initial fee reduces the explanatory power of many of the other variables.

Thus, the first fee explains up to 30% of the variance in fees, a new finding that seems quite at odds with competitive pricing. To better understand this result, we study the determinants of initial fees. In Table 6, we estimate cross-sectional regressions of a fund's total yearly fees from the year in which the fund was initiated on fund characteristics from that period. Each fund appears in the regression only during the year in which it was initiated. Fund family characteristics are potentially important drivers of an individual fund's initial fee. Thus, we include the mean and standard deviation of a fund's family as explanatory variables and estimate these regressions for almost all funds (if we assign a fund to be part of a fund family if the fund's management company has more than 10 funds, i.e., columns labeled "Number

of Funds in Family > 10", we basically include all funds) and for funds from larger fund families ("Number of Funds in Family > 100"). We also estimate separate models where we equal-weight or value-weight the fund characteristics within a family to arrive at an aggregate, across-fund characteristic value for each family.

The models can explain between 38 to 45% of the variation in first fees. Initial fees are statistically significantly higher for funds that enter with smaller size, smaller fund families, fund families that charge higher average fees, families that have a higher dispersion of fees within the fund family, families that have a higher average return, and for non-ETF and non-institutional funds[13]. We also find that factor returns (i.e., the market risk premium, SMB, HML, and UMD) are consistently negatively related to first fees. A potential interpretation is that in periods when factors are doing poorly, actively managed funds can charge excess fees. Interactions between factor returns and betas[14] are consistently positively related to first fees, implying that in situations where funds are founded in styles that have been successful in the past, higher fees are charged.

So far we have established that fees do not vary much, rejecting the random fee hypothesis, and we have shown that a fund's first fee is a strong indicator of the fund's lifetime expense ratio. The first fee, combined with other standard fund characteristics, can explain approximately 70% of the variation in fees. We next test three additional hypotheses to see if we can gain more of an understanding of how fees are set and how fee evolution, what little does occur, is determined.

4.2 The Competition hypothesis

Wahal and Wang (2010) show that when existing funds face competition from new, similar funds (as defined by the overlap in quarterly holdings), that the existing funds lower their fees to better compete with the upstart funds, evidence consistent with a competitive market for mutual funds. To test for this

[13] Ramadorai and Streatfield (2011) perform a similar analysis for hedge fund fees. They also find that better performing fund families launch high fee funds. In contrast to our results for mutual funds, they document a positive relationship between fund family size and new hedge fund fees.
[14] Given that we don't observe fund performance before initiation these betas are forward-looking; i.e., they are estimated using future returns.

effect in our sample, we define two competition variables. The first, the number of entering competing funds (*comp_funds*) is estimated as follows. For an existing fund, we identify an entering, competing fund as a new fund that has a similar beta with the existing fund. To estimate betas for a new fund, we regress the time series of monthly returns for the fund against an intercept, MKT, SMB, HML and UMD using 3 years of data from year t to t+2. To estimate betas for the existing funds, we regress the time series of monthly returns for the fund against an intercept, MKT, SMB, HML and UMD using 3 years of data from year t to t-2. For both the existing and new funds, we require a minimum of 12 monthly returns to estimate the betas. To match funds, we discretize each beta into buckets (each bucket has a size of 0.2). For each existing fund, we then sum up the total number of new funds that have at least one similar beta. The second competition variable, the average fees of entering competing funds (*comp_fees*) is based on the same procedure as described in the case of *comp_funds* but instead of counting the number of entering competing funds we determine the average first fees of these funds. Under the assumption that the matching funds are close substitutes and that there is little or no asymmetric information on the part of investors concerning that substitutability, the competition hypothesis suggests that the coefficient from a regression of fee changes on *comp_funds* should be negative and that the coefficient on *comp_fees* should be positive.[15]

We regress fee changes on important fund characteristics and the two competition variables. The results are reported in Table 7. In the pre-1999 period, the fee change models have R-squareds ranging from 13% to 17% that decrease in the post-1999 period to 6% to 7%. In pre-1999, the coefficients on both the number of entering competing funds and the average fees of entering competing funds are negative, but statistically insignificant, showing that incumbent funds do not lower their fees when faced with new competing funds. In fact, after 1999, there is a positive (albeit insignificant) effect on existing funds' fees when faced with a new competitor (e.g., as new funds' average fees go up by 1%, existing funds raise

[15] The prediction for the coefficient on *comp_fees* is ambiguous. The ideal variable to capture in this situation would be a combination of *comp_funds* and *comp_fees* that measures that extent to which entering funds put price pressure on existing funds (i.e., relative to the existing funds' fee levels). The current results represent a first attempt to get at this notion but we are in the process of refining this analysis.

their fees by approximately 1 basis point). Overall, competition from new funds does not appear to materially affect the fees of existing funds.

4.3 The Strategic Fee Setting Hypothesis

We next test the strategic fee setting hypothesis (SFSH). Christoffersen and Musto (2002) and Gil-Bazo and Ruiz-Verdu (2008) show that performance sensitive investors withdraw assets from poorly performing funds leaving only performance insensitive investors as holders of the funds' shares. Funds respond to the fact that the fund flows of the remaining investors are not sensitive to fund performance by raising fees. We note that our fee regressions from Table 2, where we estimate residual fees, already control for lagged fund returns, thus to some extent controlling for a SFSH. We provide additional evidence on this issue in Table 7, where we regress fee changes on lagged flows and lagged returns. The coefficients on lagged returns are negative and highly statistically significant, consistent with the SFSH, however, the economic effects are small. For example, a 10% drop in returns for a fund results in an average increase in fees of approximately 1.6 basis points. The economic effect from decreasing flows is even smaller. In Table 8, we further examine flows and changes in fees for funds segmented by good and bad prior performance. We perform this analysis for first fees, average fees, and residual fees. The SFSH posits that flows should be low or close to zero for already-high fee funds that have performed poorly, under the assumption that those funds are primarily held by performance insensitive investors. We find negative mean and median flows for high-fee poor-performing funds, inconsistent with the strategic fee setting hypothesis that suggests little or no outflow for high-fee poor-performing funds. The SFSH also posits that fees should increase for poor performing funds as performance sensitive investors flee these funds, leaving behind primarily fee-insensitive investors. We find some increase in the means of poor performing high fee funds, but little increase in mean fees for poor performing low fee funds. Also, for median fees, the change in fees for poor performing funds, across high and low fee funds, is approximately zero, a result that is inconsistent with the SFSH.

4.4 The Service Hypothesis

Finally, we test a fund family service hypothesis. Hortacsu and Syverson (2004), Collins (2005) and others have suggested that variation in services, such as financial advice or complementary investment instruments, may explain fee variation[16]. As with the strategic fee setting, we note that our fee regressions from Table 2, where we estimate residual fees, already control for fund family characteristics that may proxy for service. For example, assuming that large fund families offer better service, we find in Table 2 that funds that are part of a family with more than 100 funds charge, on average, an extra 12 to 14 basis points in fees, but obviously, controlling for large fund families does not result in a small spread in residual fees. We provide additional evidence on service and family size in Figure 2. We report plots of residual fees, using the base specification 4 of the Table 2 regressions for funds that are part of a large fund family with greater than 100 funds ("funds within families") and for funds that are in families of less than 100 funds ("funds outside of families"). In the residual fee plots, for the funds outside of families, there is clearly a large spread (approximately 2%-2.5%). The more interesting finding is that for funds within large families, where presumably there is greater customer service, we still find large spreads in residual fees (approximately 2%). Assuming that the number of funds in a family is positively correlated with service, this clearly suggests that service does not explain our main finding of large spreads in residual fees for essentially identical funds.

4.5 The Captive Investor Hypothesis

Finally, we test a "captive" investor hypothesis. We examine if funds that are likely to inhibit easy investor exit (thus creating captive investors) are the funds with high residual fees. The intuition behind this hypothesis is that investors may get trapped within certain funds, and the managers of such funds recognize that these investors are unable or unwilling to exit these funds, and thus charge high fees. One potential trapping mechanism is that investors are lured into high-fee funds via low front-end loads and

[16] In an experimental setup, Choi, Laibson and Madrian (2010) reject the hypothesis that investors buy high-fee index funds because of nonportfolio services. Interestingly, they also report that their subjects did not minimize fees even when search costs were eliminated.

high marketing efforts. Once investors are shareholders in the high-fee funds, they are kept as shareholders by making exit costly, for example via high back-end loads. We deem these types of funds "easy-in, hard-out" funds. Another potential mechanism is that investors get stuck with high-fee funds in situations in which they are very restricted in their investment choices, e.g., in the case of funds residing within pension plans.

As a first test of this hypothesis, we look at the distribution of fee residuals across front-end loads, back-end loads and marketing expenses. The captive-investor hypothesis predicts that fund with low front-end loads, high marketing expenditures, and high back-end loads are likely to be high residual fee funds. Table 9 Panel A shows the average and standard deviations of front-end, back-end, and 12b-1 fees for Deciles 1, 3, 5, 7, and 10 of the residual fee distribution. The higher fee funds on average appear to be "easy-in." These funds spend more on advertising than the low fee funds (and the differences are statistically significant). These funds also have significantly lower front-end loads than the low fee funds. The average back-end loads are also consistent with the hypotheses: high residual fee funds have approximately an 80 basis point higher deferred sales load than the low fee funds.[17]

The evidence in Table 9 Panel A is also consistent with the idea that high fee funds may have an overconcentration of pension-plan funds. We define funds with high flow autocorrelation as "pension-plan" funds. We estimate each fund's flow autocorrelation using the entire time-series of monthly flows per fund. We find that high fee funds have much higher autocorrelated flows (in fact, more than double) than the low fee funds, evidence consistent with the captive-investor hypotheses.

We also estimate in Table 9 Panel A the percentage of easy-in, hard out, and pension plan funds in each residual fee decile. Non of the funds in the lowest three residual fee deciles and between 4 to 6% of the funds in the top five deciles are "easy-in, hard-out" funds. For the pension plan funds, the percentage of funds in decile 1 is 5%, increasing to 23% in decile 10. Thus, the pension plan funds make up a much

[17] These results are consistent with Barber, Odean and Zheng (2005) in the sense that investors seem to be sensitive to front-end loads. Our overall evidence of persistent price dispersion and the empirical pattern that funds with low front-end loads tend to be relatively more expensive are, however, in contrast to their results.

higher percentage of the high fee funds than would be expected by chance. Finally, in Table 9 Panel B, we estimate the residual fee distributions (using the residuals from the Table 2 regressions) for the two categories of captive investor funds and the difference between the full sample residuals and the two categories. The mean of the residual fee distribution for the "easy-in, hard-out" funds (pension funds) is 54bp (24bp) and the average spread in residuals between the 1% and 99% points of the distribution is 97bp (192bp). Thus, these two categories of captive investor funds do not, obviously, entirely explain our main findings (i.e., the spreads in residuals of the full sample minus these two categories of funds are still large). However, the results in this section show that these captive investor funds charge on average higher residual fees and thus help to explain which funds reside in the higher points of the residual fee distribution.

4.6 The Remainders

We add new variables to directly control for the random fee and competition hypotheses and re-estimate the residual fee spread from the regressions of Table 2 (recall that the strategic fee setting hypothesis and the service hypothesis variables are already controlled for in the residual fee spread estimation). For the random fee hypothesis, we add a random fee changes variable (*rand_feechgs*). For each fund and each year, we determine the fraction of positive and negative fee changes of all changes that we have observed for the fund since its first appearance in the CRSP Mutual Fund Database. Then we use the minimum value of the fractions of positive and negative changes as our variable, motivated by the idea that randomized pricing requires both increases and decreases of fees (and not just unidirectional changes). For the competition hypothesis, we calculate the total number of competing funds for each existing fund (*compAll_funds*). For a given existing fund, we identify competing funds as funds that have a similar beta to the existing fund. To estimate betas for an existing fund, we regress the time series of monthly returns for the fund against an intercept, MKT, SMB, HML and UMD using 3 years of data from year t to t-2. We require a minimum of 12 monthly returns to estimate the betas. To match funds, we discretize each beta into buckets (each bucket has a size of 0.2). For each fund, we then sum up the total

number of other funds that have at least one similar beta. The second competition variable, the average fees of competing funds (*compAll_fees*) is based on the same procedure as described in the case of *compAll_funds* but instead of counting the number of competing funds we determine the average fees of these funds. We present the results in the last row of each panel of Table 3. The spread in residual fees, after controlling for these new fund characteristics, is qualitatively similar to our previous results. For example, in Panel A, the base model generates an average spread in residuals between the 1% and 99% points of the distribution of 2.34%, 2.88%, and 1.18% for all funds, the smallest quintile, and the largest quintile, respectively, and the new model generates residuals of 2.28%, 2.64% and 1.09%, respectively. Thus, while there are slight drops in the residual fee spreads, the random fee setting and competition hypothesis do not come close to fully explaining the large differences in pricing across similar funds.

Finally, we examine the differences in pricing between institutional and retail funds. Christoffersen and Musto (2002) discuss how institutional funds charge, on average, lower fees relative to retail funds, and that the strategic fee setting hypothesis applies only to retail funds and not to institutional funds. Indeed, the literature (see Bris, Gulen, Kadiyala, Rau (2007) and others) has shown that institutional funds tend to have lower fees and are presumed to be held by more sophisticated investors relative to retail funds. Thus, if holders of institutional funds are more educated about funds and have a greater influence on prices, it is possible that our results do not hold for institutional funds. In figure 3, we plot raw fees and estimate residual fees separately for both retail and institutional funds. The raw and residual spreads are indeed higher for retail funds, but we still see evidence of relatively large spreads in residual fees for institutional funds (ranging from about 1% to 1.2%) with no clear trend of decreasing fee spreads in more recent years. Thus, our results also apply to institutional funds.

5. Conclusion

In this paper we examine how mutual funds price their services (management and 12b-1 fees). Surprisingly, after we control for a variety of fund characteristics we find that the unexplained portion of

fund fees exhibits considerable dispersion and that the dispersion in the prices charged by funds has not declined over time, despite significant entry and growth in assets under management.

Similar to others, we first show that fees are an important determinant of fund underperformance — that is, investors earn low returns on high fee funds, which indicates that investors are not rewarded through superior performance when purchasing "expensive" funds. We explore a number of hypotheses to explain the dispersion in fees and find that none adequately explain the data. Most importantly, there is very little evidence that funds change their fees over time. In fact the most important determinant of a fund's fee is the initial fee that it charges when it enters the market. There is little evidence that funds reduce their fees following entry by similar funds or that they raise their fees following large outflows as predicted by the strategic fee setting hypothesis. We also do not find evidence that higher fees are associated with proxies for higher service levels provided to investors.

Overall, our findings provide little evidence that competitive pricing exists in the market for mutual funds. However, our findings leave an important question unanswered - - namely, why investors are not able to distinguish cheap from expensive funds. We hypothesize that this is the case because performance of funds is noisy and not persistent on average. Thus, although fees are observable and persistent, after-fee performance – which is presumably the decision criterion of the average investor – is not, which makes it difficult for investors to distinguish good from bad investments. One possible explanation is that high fee funds "market" themselves to investors.

6. References

Bakos, Y., 2001, The Emerging Landscape of Retail E-commerce, *Journal of EconomicPerspectives,* 15, 69-80.

Barber, B. M., Odean, T., and L. Zheng, 2005, Out of Sight, Out of Mind: The Effects of Expenses on Mutual Fund Flows, *Journal of Business*, 78 (6), 2095-2119.

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

Bessler, W., D, Blake, P. Luckiff, and I. Tonks, 2008, Why is Persistent Mutual Fund Performance so Difficult to Achieve? The Impact of Management Turnover and Fund Flows, working paper.

Bris, A., H. Gulen, P. Kadiyala, and P. Raghavendra Rau, 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 A. Goolsbee, 2002, Does the Internet Make Markets More Competitive? Evidence from the Life Insurance Industry, *Journal of Political Economy,* 110, 481-507.

Brynjolfsson, E. and M. Smith, 2000, Frictionless Commerce? A Comparison of Internet and Conventional Retailers, *Management Science,* 46(4), 563-585.

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

Choi, J. J., Laibson, D., and B. C. Madrian, 2010, Why Does the Law of One Price Fail? An Experiment on Index Mutual Funds, *The Review of Financial Studies*, 23, 1405-1432.

Christoffersen, S. and D. Musto, 2002, Demand Curves and the Pricing of Money Mangement, *Review of Financial Studies, Vol. 15, No. 5.*

Collins, S., 2005, Are S&P 500 Index Mutual Funds Commodities?, *Investment Company Institute Perspective*, 11(3).

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

Diamond, P., 1971, A Model of Price Adjustments, *Journal of Economic Theory*, 3 (3), 156-168.

Elton, E., M. Gruber, and J. Busse, 2004, Are Investors Rational? Choices Among Index Funds, *Journal of Finance*, Vol. 59, No. 1.

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

French, K. 2008, Presidential Address: The Cost of Active Investing, *Journal of Finance*, Vol. 63, No. 4.

Gil-Bazo, J., and P. Ruiz-Verdu, 2008, When Cheaper is Better: Fee Determination in the Market for Equity Mutual Funds, *Journal of Economic Behavior and Organization*, pp. 871-885.

Gil-Bazo, J., and P. Ruiz-Verdu, 2009, The Relation Between Price and Performance in the Mutual Fund Industry, *Journal of Finance*, forthcoming.

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. *The Quarterly Journal of Economics*, 403-456.

Khorana, A., Servaes, H. and P. Tufano, 2009, Mutual Fund Fees around the World, *Review of Financial Studies*, 22 (3), 1279-1310.

Lach, S., 2002, Existence and Persistence of Price Dispersion: an Empirical Analysis, *The Review of Economics and Statistics*, 84 (3), 433-444.

Liang, B., 2000, Hedge Funds: The Living and the Dead, *Journal of Financial and Quantitative Analysis*, 35, 309-326.

Nakamura, L., 1999, The Measurement of Retail Output and the Retail Revolution, *Canadian Journal of Economics,* 32(2), 408-425.

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

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

Ṕastor, L. and R. Stambaugh, 2002, Mutual funds performance and seemingly unrelated assets, *Journal of Financial Economics,* 63, 315-349.

Ṕastor, L. and R. Stambaugh, 2010, On the Size of the Active Management Industry, Working Paper.

Pratt, J., D. Wise and R. Zeckhauser, 1979, Price Differences in Almost Competitive Markets, *Quarterly Journal of Economics,* 939, 189-211.

Ramadorai, T. and M. Streatfield, 2011, Money for Nothing? Understanding Variation in Reported Hedge Fund Fees, Working Paper.

Scholten, P. and S. Smith, 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(4), 833-862.

Varian, H., 1980, A Model of Sales, *American Economic Review*, 70 (4), 651-659.

Wahal, S., and A. Wang, 2010, Competition Among Mutual Funds, forthcoming in the *Journal of Financial Economics*.

Appendix

Table A. Sample selection

The sample selection follows Pastor and Stambaugh (2002). Accordingly, we select only domestic equity funds and exclude all funds not investing primarily in equities such as money market or bond funds. In addition, we exclude international funds, global funds, balanced funds, flexible funds, and funds of funds. The ICDI classification codes that were used by Pastor and Stambaugh (2002) are however no longer available. Thus, we follow Bessler et al. (2008) who use a combination of Lipper codes, Wiesenberger codes and Strategic Insight codes to identify domestic equity funds. Specifically we include funds in our sample with the following classification codes:

1. Lipper: CA, EI, EIEI, G, GI, I, LCCE, LCGE, LCVE, MC, MCCE, MCGE, MCVE, MLCE, MLGE, MLVE, SCCE, FS, H, NR, S, SESE, TK, TL, UT.
2. Wiesenberger: AGG, G, G-I, G-I-S, G-S, G-S-I, GCI, GRI, GRO, IG, I-G-S, I-S, I-S-G, IEQ, ING, LTG, MCG, S-G, S-GI, S-I-G, S-I, SCG, ENR, FIN, HLT, TCH, UTL.
3. Strategic Insight: AGG, GMC, GRI, GRO, ING, SCG, ENV, FIN, HLT, NTR, SEC, TEC, UTI.

This leaves us with around 35% of all funds covered in the CRSP Mutual Fund Database (our sample consists of a total of 13817 funds while the CRSP Mutual Fund Database universe has approximately 40000 funds). In total net assets our sample covers around 32% of the cumulative net assets represented in the database.

Table B. Variable construction and definitions

Variable Name	Variable Definition	Source
Total yearly expense ratio (*exp_rati*)	Ratio of total investment that shareholders pay for the fund's operating expenses, which include 12b-1 fees.	CRSP Mutual Fund Database
Change in yearly expense ratios (Δ(*exp_rati*))	Yearly change in total expense ratio.	Calculated
Expense ratio when fund started (*first_exp_rati*)	This is the total yearly expense ratio in the year when the fund was initiated.	Calculated
Average expense ratio of funds within a fund family (*avgFamFee*)	The value-weighted means of expense ratios of funds that belong to a fund family.	Calculated

Variable Name	Variable Definition	Source
Average performance of funds within a fund family (*avgFamPerf*)	The value-weighted average Carhart alpha of funds that belong to a fund family.	Calculated
Standard deviation of performances of funds within a fund family (*sdFamPerf*)	The value-weighted standard deviation of Carhart alphas of funds that belong to a fund family.	Calculated

Variable Name	Variable Definition	Source
Different definitions of funds being part of a fund family (*part_family1, part_family10, part_family100, part_family250*)	We define a fund family as a management company with more than 1 (10) [100] {250} funds associated with it. The standard case in our analysis is *part_family100*.	Calculated.
Flags identifying whether a fund is institutional (*inst_flag*), open to new investment (*open_flag*) or an ETF (*etf_flag*)	These are three flags that indicate whether a fund is an institutional fund, whether it is open to new investment and whether it is an ETF.	CRSP Mutual Fund Database
Number of competing funds (*compAll_funds*)	For a given existing fund, we identify competing funds as funds that have a similar beta with the existing fund. For the matching we discretize each beta into buckets (each bucket has a size of 0.2). We do this separately for each beta (MKT, SMB, HML and UMD) and then calculate – across betas – the total number of competing funds for an existing fund.	Calculated
Average fees of competing funds (*compAll_fees*)	We use the same procedure as described in the case of *compAll_funds* but instead of counting the number of competing funds in any given year we determine the average fees of these competing funds.	Calculated
Number of **entering** competing funds (*comp_funds*)	For a given existing fund, we identify an entering, competing fund as a new fund that has a similar beta with the existing fund. For the entering fund we use its beta over the next 3 years. For the matching we discretize each beta into buckets (each bucket has a size of 0.2). We do this separately for each beta (MKT, SMB, HML and UMD) and then calculate – across betas – the total number of competing funds for an existing fund.	Calculated
Average fees of **entering** competing funds (*comp_fees*)	We use the same procedure as described in the case of *comp_funds* but instead of counting the number of entering competing funds we determine the average first fees of these funds.	Calculated
Random fee changes (*rand_feechgs*)	For each fund and each year, we determine the fraction of positive and negative fee changes that we have observed for the fund since its first appearance in the CRSP Mutual Fund Database. Then we use the minimum value of the frequency of positive and negative changes as our variable. The idea is that randomized pricing requires increases and decreases of fees (and not just changes).	Calculated
12b-1 fees	Ratio of the total assets attributed to marketing and distribution costs. Available since 1992.	CRSP Mutual Fund Database
Front Load	Front loads for investments represent maximum sales charges. They often change with the level of investment. Thus, the database reports front load schedules. The front load value that we use in the paper is the equal weighted average of all front	CRSP Mutual Fund Database

Variable Name	Variable Definition	Source
	loads charged by a fund across different investment levels.	
Rear Load	The rear load is a fee charged by the fund when an investor withdraws funds. The rear load might vary by investment level and by the amount of time the funds were invested in the fund. The rear load value that we use in the paper is the equal weighted average across all reported rear load values across these dimensions.	CRSP Mutual Fund Database
Flow Autocorrelation	For each fund, we calculate monthly flows (see definition of yearly flows, *w_yflow*, for details) and then estimate the autocorrelation of these monthly flows.	Calculated
Easy-In Hard-Out Funds	We define funds to be "easy-in hard-out funds" when their rear loads are in the top decile and their 12b-1 fees are in the top quartile of all funds within a given year.[18]	Calculated
"Pension Plan" Funds	We define a mutual fund to be a "Pension Plan Fund" if its monthly autocorrelation is in the top decile of all funds. This classification does not vary over time.	Calculated

[18] An obvious, alternative definition also considers front loads. If a fund is "easy-in" it should have low front loads. The load data is, however, not very complete and if we require to know a fund's rear and front loads, we reduce the sample size dramatically.

Table 1. Summary Statistics

The table reports summary statistics (means and between standard deviations, i.e., means of yearly standard deviations) and correlation tables of our sample of domestic equity mutual funds (see Table A in the Appendix for a detailed description of the sample). The data covers the period of 1963 to 2008 and is a yearly panel. Variables are defined in Table B in the Appendix. The table focuses on our main variables; i.e., the ones used to explain fund expense ratios. Some information is only available after 1999 (e.g., information on fund families) and, thus, we split the sample into a pre-1999 and a post-1999 subset. Panel B and C summarize the sample by expense ratio deciles. The last column in each table reports the difference between Decile 1 and Decile 10. Stars indicate significance at the 1% (***), 5% (**) and 10% (*) level.

Panel A. Full Sample

	Pre-1999		Post-1999	
	Mean	**SD**	**Mean**	**SD**
Number of Funds per Year	545	586	5562	1318
exp_rati	0.0129	0.0061	0.0141	0.0065
Δ(exp_rati)	0.0000	0.0018	0.0000	0.0009
yret	0.1546	0.1149	0.0064	0.1324
yret_pos	0.1724	0.1016	0.0983	0.0602
yret_neg	-0.0178	0.0299	-0.0918	0.0979
yalpha	-0.0053	0.0792	-0.0053	0.0478
ycarh_alpha	0.0329	0.4022	-0.0133	0.0431
beta_mkt	0.9243	0.2156	0.9766	0.2612
beta_hml	-0.0591	0.3834	-0.0274	0.369
beta_smb	0.1619	0.3859	0.0665	0.2579
beta_umd	0.0557	0.2518	0.0241	0.1783
r_squared	0.8396	0.1495	0.8683	0.13
first_exp_rati	0.0128	0.0067	0.0145	0.0069
w_yflow	0.3191	1.8946	0.5046	1.6483
ln_mtna	4.0982	2.0987	3.3519	2.6044
mtna	436.4834	669.054	463.5951	1790.59
fund_age	9.2409	4.592	6.5168	5.8506
sd_mret	0.0444	0.0135	0.0466	0.0196
turn_rat				

Panel B. Summary Statistics by Expense Ratio Deciles – Pre-1999 Sample

	Decile 1		Decile 3		Decile 5		Decile 7		Decile 10		Decile 1-10
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Diff.
exp_rati	0.0052	0.0018	0.0089	0.001	0.0111	0.0014	0.0141	0.0022	0.0257	0.0059	-0.021***
Δ(exp_rati)	-0.0006	0.0052	-0.0002	0.0018	-0.0001	0.0021	-0.0003	0.0025	0.0015	0.0057	-0.002***
yret	0.1581	0.112	0.1613	0.1152	0.1518	0.1241	0.1503	0.1332	0.1287	0.1479	0.029***
yret_pos	0.1725	0.0965	0.1757	0.0965	0.169	0.107	0.1697	0.1139	0.1578	0.119	0.015***
yret_neg	-0.0143	0.0326	-0.0143	0.0386	-0.0171	0.0366	-0.0193	0.0406	-0.0289	0.0527	0.015***
yalpha	-0.0012	0.0539	-0.0018	0.0444	-0.0037	0.0655	-0.0061	0.0697	-0.0279	0.0963	0.027***
ycarh_alpha	-0.0021	0.0423	-0.0041	0.0547	-0.0055	0.0592	-0.0119	0.0629	-0.0251	0.0711	0.023***
beta_mkt	0.9327	0.2071	0.929	0.1699	0.9274	0.211	0.918	0.2143	0.9274	0.2723	0.005
beta_hml	-0.026	0.3245	-0.0461	0.3033	-0.05	0.3711	-0.0419	0.3583	-0.108	0.4576	0.082***
beta_smb	-0.0029	0.2884	0.0733	0.2886	0.1579	0.3586	0.2008	0.3685	0.3341	0.511	-0.337***
beta_umd	0.0116	0.1893	0.0395	0.1942	0.0473	0.23	0.0746	0.2324	0.0807	0.3337	-0.069***
r_squared	0.8896	0.1613	0.8698	0.1257	0.8446	0.1448	0.8224	0.1594	0.7632	0.1968	0.126***
first_exp_rati	0.0065	0.0047	0.0092	0.0046	0.0115	0.0046	0.0142	0.0053	0.0215	0.0075	-0.015***
w_yflow	0.1908	1.7286	0.2367	1.6338	0.2603	1.409	0.3424	1.3676	0.4338	2.0533	-0.243***
ln_mtna	5.7637	2.2359	4.9996	1.926	4.2314	1.9302	3.6769	1.8869	2.2615	1.7367	3.502***
mtna	1581	4294	581	1870	294	1088	215	671	48	105	1533.5***
mtna_share	0.723%	0.553%	0.237%	0.474%	0.119%	0.314%	0.067%	0.122%	0.025%	0.103%	0.70%***
fund_age	13.2252	9.1036	12.2635	8.9289	10.1815	7.4482	7.5967	5.9195	6.329	4.8624	6.896***
sd_mret	0.0405	0.0171	0.0415	0.0122	0.0436	0.0166	0.0449	0.0154	0.0511	0.0164	-0.011***

Panel C. Summary Statistics by Expense Ratio Deciles – Post-1999 sample

	Decile 1		Decile 3		Decile 5		Decile 7		Decile 10		Decile 1-10
	Mean	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Diff.**
exp_rati	0.0042	0.0021	0.0098	0.0009	0.0127	0.0009	0.0164	0.0016	0.0253	0.0049	-0.021***
Δ(exp_rati)	-0.0001	0.001	-0.0001	0.001	-0.0001	0.0011	-0.0001	0.0014	0.0003	0.0024	-0.0004***
Yret	0.0121	0.1522	0.0115	0.1789	0.0117	0.1892	0.0048	0.2019	-0.0061	0.2049	0.018***
yret_pos	0.0901	0.0837	0.0994	0.0864	0.1009	0.0963	0.0983	0.1143	0.1049	0.1193	-0.015***
yret_neg	-0.078	0.0981	-0.0879	0.1206	-0.0892	0.1259	-0.0935	0.1277	-0.111	0.1296	0.033***
Yalpha	0.0010	0.0326	-0.0015	0.0447	-0.0023	0.0507	-0.0065	0.0548	-0.0186	0.0741	0.020***
ycarh_alpha	-0.0049	0.0446	-0.0104	0.0564	-0.0118	0.0639	-0.0139	0.0651	-0.0233	0.0828	0.018***
beta_mkt	0.8789	0.2617	0.9735	0.2085	0.9836	0.2402	0.9886	0.2679	1.0389	0.3234	-0.160***
beta_hml	0.0351	0.2856	-0.0002	0.339	-0.0008	0.3601	-0.044	0.3902	-0.1738	0.4545	0.209***
beta_smb	0.0075	0.191	0.0272	0.2404	0.0751	0.2586	0.0875	0.2763	0.1755	0.3102	-0.168***
beta_umd	-0.0073	0.1459	0.0202	0.1701	0.0299	0.1787	0.0323	0.191	0.0426	0.2236	-0.050***
r_squared	0.8929	0.146	0.8803	0.1078	0.8683	0.1113	0.863	0.1303	0.8264	0.1503	0.066***
first_exp_rati	0.0047	0.003	0.0105	0.0031	0.0132	0.0034	0.0168	0.0037	0.0243	0.0059	-0.020***
w_yflow	0.563	2.1902	0.6233	2.5717	0.5171	1.9354	0.5043	1.9244	0.3448	1.5014	0.218***
ln_mtna	4.5465	2.9541	4.156	2.6665	3.4314	2.524	3.0345	2.3985	2.0143	1.9524	2.532***
Mtna	1523	4353	652	1977	273	656	223	641	44	98	1479***
mtna_share	0.060%	0.175%	0.025%	0.075%	0.010%	0.025%	0.009%	0.025%	0.002%	0.004%	0.06%***
fund_age	6.0691	6.5409	8.5349	8.9145	7.3196	7.6957	5.691	5.465	5.2391	4.2274	0.830***
sd_mret	0.0404	0.0174	0.0453	0.0178	0.0465	0.0217	0.0475	0.0206	0.0537	0.0241	-0.013***
turn_rat	0.4624	0.4833	0.7992	0.6617	0.9202	0.9211	1.023	1.181	1.3328	1.2522	-0.870***
ln_mgmt_mtna	10.7632	2.0674	9.8911	2.2031	9.4268	2.3247	9.6709	2.5008	8.6126	2.4426	2.151***
part_family	0.9974	0.0647	0.9833	0.116	0.9783	0.1421	0.9758	0.1606	0.9729	0.1429	0.025***
part_family10	0.9618	0.1956	0.9064	0.2757	0.8671	0.3149	0.9068	0.3032	0.8952	0.2818	0.067***
part_family100	0.6482	0.4376	0.5753	0.4716	0.5583	0.4781	0.6309	0.4653	0.5155	0.4651	0.133***
part_family250	0.2775	0.4345	0.2842	0.4307	0.2618	0.4232	0.2823	0.4275	0.1506	0.3545	0.127***
inst_flag	0.612	0.4777	0.4868	0.495	0.2413	0.4423	0.1325	0.3772	0.0183	0.1644	0.594***
open_flag	0.9748	0.1236	0.9792	0.1245	0.9682	0.1494	0.9834	0.1109	0.9657	0.1654	0.009***
etf_flag	0.0799	0.2143	0.0024	0.0606	0.0002	0.0219	0	0	0.0009	0.0566	0.079***

Panel D. Average (time-series) cross-sectional correlations

	Δ(exp_rati)	yret	yalpha	ycarh alpha	beta mkt	beta hml	beta smb	beta umd	r_squared	w_yflow	ln mtna	fund age	sd mret	turn rat	ln mgmt mtna
exp_rati	0.21	-0.07	-0.13	-0.07	0.01	-0.02	0.22	0.06	-0.23	0.07	-0.49	-0.18	0.19	0.25	-0.18
Δ(exp_rati)	1.00	-0.05	-0.07	-0.05	0.00	-0.02	-0.03	-0.02	0.01	-0.06	-0.02	0.02	0.01	0.00	-0.03
yret		1.00	0.40	0.60	0.01	0.08	0.07	0.07	0.02	0.19	0.07	-0.03	0.02	0.02	-0.01
alpha			1.00	0.49	-0.09	-0.16	0.06	-0.03	-0.09	0.16	0.11	-0.04	0.14	-0.04	0.05
ycarh_alpha				1.00	-0.08	-0.06	0.01	-0.06	-0.01	0.13	0.05	-0.04	-0.02	-0.03	-0.02
beta_mkt					1.00	-0.10	0.07	0.05	0.36	-0.01	0.08	0.00	0.52	0.07	-0.01
beta_hml						1.00	-0.09	-0.14	-0.10	0.00	-0.04	0.02	-0.42	-0.09	0.00
beta_smb							1.00	0.20	-0.14	0.10	-0.17	-0.13	0.50	0.21	-0.07
beta_umd								1.00	0.00	0.05	0.01	-0.02	0.21	0.24	-0.02
r_squared									1.00	-0.03	0.25	0.05	-0.17	-0.15	0.07
w_yflow										1.00	-0.05	-0.15	0.05	0.07	0.02
ln_mtna											1.00	0.33	-0.08	-0.14	0.18
fund_age												1.00	-0.06	-0.12	0.03
sd_mret													1.00	0.22	-0.05
turn_rat														1.00	-0.04
ln_mgmt_mtna															1.00

Table 2. Fund Fee Regressions

The table reports results of Fama-MacBeth regressions, in which we use the yearly expense ratio as dependent variables (see Table A in the Appendix for a detailed description of the sample). The data covers the period of 1963 to 2008 and is a yearly panel. Variables are defined in Table B in the Appendix. All variables are lagged by one year. Some information is only available after 1999 (e.g., information on fund families) and, thus, we split the sample into a pre-1999 and a post-1999 subset.

	Pre 1999						Post 1999					
	Spec. 1		Spec. 2		Spec. 3		Spec. 4		Spec. 5		Spec. 5	
	Coeff.	t-value	Coeff.	t-value	Coeff.	t-value	Coeff.	t-value	Coeff.	t-value	Coeff.	t-value
$Yret_{t-1}$	-0.0077	-4.61	-0.0063	-4.62			-0.0068	-5.2	-0.0045	-5.29		
Positive $Yret_{t-1}$					-0.0035	-2.92					-0.0043	-4.7
Negative $Yret_{t-1}$					-0.0169	-1.8					-0.0098	-1.89
beta_mkt $_{t-1}$	0.0016	1.72	0.0009	1	0.0013	1.47	-0.0006	-1.24	-0.001	-4.23	-0.0008	-2.4
beta_hml $_{t-1}$	0.0007	1.37	0.0002	0.5	0	0.1	0.0006	1.45	0.0001	0.42	0	0.21
beta_smb $_{t-1}$	0.0019	3.35	0.0008	1.79	0.0008	2.03	-0.0005	-2.06	-0.0005	-2.02	-0.0005	-1.66
beta_umd $_{t-1}$	0.0009	1.62	0.0007	1.6	0.0008	1.84	0.0001	0.1	-0.0005	-0.97	-0.0006	-1.21
r_squared $_{t-1}$	-0.0049	-5.26	-0.0031	-2.95	-0.003	-3.01	-0.0031	-3.68	-0.0003	-0.77	-0.0004	-0.99
First_exp_rati			0.3066	10.79	0.3044	10.82			0.5624	27.07	0.5619	27.16
flow $_{t-1}$	0	0.18	-0.0003	-1.76	-0.0003	-2.02	-0.0001	-2.75	0	-2.85	0	-2.97
ln_mtna $_{t-1}$	-0.0016	-18.64	-0.0014	-14.71	-0.0014	-14.68	-0.0007	-29.54	-0.0004	-16.77	-0.0004	-16.93
fund_age $_{t-1}$	0	-5.08	0.0001	1.63	0.0001	1.69	-0.0001	-13.31	0	-0.53	0	-0.51
sd_mret $_{t-1}$	0.0249	1.34	0.0291	1.75	0.0141	0.95	0.1129	6.71	0.0699	3.46	0.0668	2.94
turn_rat $_{t-1}$							0.0001	2.87	0	-0.02	0	0.09
ln_mgmt_mtna $_{t-1}$							-0.0004	-8.86	-0.0003	-8.55	-0.0003	-8.56
part_family100							0.0014	3.45	0.0012	5.5	0.0012	5.39
inst_flag							-0.0063	-42.4	-0.0025	-13.53	-0.0025	-13.57
open_flag							-0.0001	-0.21	0.0005	1.98	0.0005	2
etf_flag							-0.0034	-4.5	-0.0013	-4.39	-0.0013	-4.37
Constant	0.0203	21.59	0.0145	12.2	0.0144	12.36	0.0209	19.04	0.0089	18.11	0.0089	17.63
# of years	34		34		34		10		10		10	
# of obs.	12108		9818		9818		29052		26655		26655	
Avg. R-Squared	44.3%		56.3%		57.1		44.2%		69.5%		70.2%	

Table 3. Residual Fee Spreads

The table summarizes (time-series) mean spreads between percentiles of the residual fee distribution for our sample of mutual funds (see Table A in the Appendix for a detailed description of the sample). In Panel A, the residual fee is defined as the regression residual of Spec. 1 (pre-1999) and Spec. 4 (post-1999) as specified in Panel A of Table 2. To get the residual fee distribution used in Panel B and C, we re-estimate Spec. 1 and Spec. 4 of Panel A (Table 2) for the subsamples of smallest (bottom quintile) and largest (top quintile) funds per year. Coefficient estimates of these regressions are consistent with the ones reported in Table 2 and available from the authors upon request. For this table we extend the basic specification as described in the first column. The first row of each table contains the base specification that we use throughout the paper. The data covers the period of 1963 to 2008 and is a yearly panel. Variables are defined in Table B in the Appendix. The "*persistence dummy*" equals 1 if the fund performed in the top quintile with respect to 4-Factor alphas over the past two years.

Panel A. Full Sample

	Mean Spread (bp) 75%-25%	Mean Spread (bp) 90%-10%	Mean Spread (bp) 99%-1%
Yearly Return (**$Yret_{t-1}$**)	44	89	234
Yearly Return (**$Yret_{t-1}$**) + Persistence Dummy	44	88	234
4 Factor Alpha (**$Yalpha_{t-1}$**) + Persistence Dummy	44	88	231
Carhart Performance (**$Ycarh\_alpha_{t-1}$**) + Persistence Dummy	44	88	234
Yearly Return (**$Yret_{t-1}$**) + Competition (***$compAll\_funds_{t-1}$, $compAll\_fees_{t-1}$***) + Randomization (***$rand\_feechgs_{t-1}$***)	43	87	228

Panel B. Bottom Size Quintile of Funds

	Mean Spread (bp) 75%-25%	Mean Spread (bp) 90%-10%	Mean Spread (bp) 99%-1%
Yearly Return (**$Yret_{t-1}$**)	65	137	288
Yearly Return (**$Yret_{t-1}$**) + Persistence Dummy	65	136	284
4 Factor Alpha (**$Yalpha_{t-1}$**) + Persistence Dummy	62	131	279
Carhart Performance (**$Ycarh\_alpha_{t-1}$**) + Persistence Dummy	65	134	282
Yearly Return (**$Yret_{t-1}$**) + Competition (***$compAll\_funds_{t-1}$, $compAll\_fees_{t-1}$***) + Randomization (***$rand\_feechgs_{t-1}$***)	62	127	264

Panel C. Top Size Quintile of Funds

	Mean Spread (bp) 75%-25%	Mean Spread (bp) 90%-10%	Mean Spread (bp) 99%-1%
Yearly Return (**$Yret_{t-1}$**)	28	60	118
Yearly Return (**$Yret_{t-1}$**) + Persistence Dummy	28	60	117
4 Factor Alpha (**$Yalpha_{t-1}$**) + Persistence Dummy	28	60	118
Carhart Performance (**$Ycarh\_alpha_{t-1}$**) + Persistence Dummy	28	60	117
Yearly Return (**$Yret_{t-1}$**) + Competition (***$compAll\_funds_{t-1}$, $compAll\_fees_{t-1}$***) + Randomization (***$rand\_feechgs_{t-1}$***)	27	58	109

Table 4. Trading Strategy

The table summarizes the cumulative (i.e., compounded) Carhart alphas of a strategy that buys funds, which are in the bottom decile according to reported expense ratios (residual expense ratios), and shorts funds, which are in the top deciles according to reported expense ratios (residual expense ratios). Funds are equally-weighed within portfolios. The table also reports the compounded spread between the average reported expense ratio (residual expense ratio) of funds in the top and the bottom decile. The residual fee is defined as the regression residual of Spec. 1 (pre-1999) and Spec. 4 (post-1999) as specified in Panel A of Table 2. Our sample consists of domestic equity mutual funds (see Table A in the Appendix for a detailed description of the sample). The data covers the period of 1963 to 2008 and is a yearly panel. Variables are defined in Table B in the Appendix.

Reported Expense Ratio						Residual Expense Ratio					
Year	**α**	**Fee Spread**	**Year**	**α**	**Fee Spread**	**Year**	**α**	**Fee Spread**	**Year**	**α**	**Fee Spread**
1963	0.2%	1.5%	**1986**	30.0%	43.9%				**1986**	2.1%	28.1%
1964	3.7%	2.6%	**1987**	31.5%	46.1%				**1987**	5.9%	29.8%
1965	-0.6%	3.8%	**1988**	31.9%	48.6%				**1988**	8.2%	31.5%
1966	1.1%	5.1%	**1989**	32.3%	51.0%	**1966**	1.0%	0.9%	**1989**	8.6%	33.3%
1967	0.1%	6.7%	**1990**	32.8%	53.3%	**1967**	-10.7%	2.2%	**1990**	9.5%	35.2%
1968	-3.3%	7.8%	**1991**	34.0%	55.5%	**1968**	-8.9%	2.9%	**1991**	8.2%	37.0%
1969	6.6%	9.7%	**1992**	35.4%	57.7%	**1969**	-6.6%	3.9%	**1992**	5.9%	38.9%
1970	7.2%	11.5%	**1993**	41.5%	59.8%	**1970**	-5.6%	5.2%	**1993**	9.1%	40.5%
1971	9.4%	14.0%	**1994**	46.1%	61.8%	**1971**	-10.7%	6.9%	**1994**	11.7%	42.2%
1972	2.3%	15.6%	**1995**	48.6%	63.9%	**1972**	-9.9%	7.9%	**1995**	14.9%	43.8%
1973	1.5%	17.8%	**1996**	54.2%	65.8%	**1973**	-4.6%	9.6%	**1996**	16.3%	45.5%
1974	0.9%	19.9%	**1997**	52.4%	67.8%	**1974**	-4.4%	11.0%	**1997**	19.3%	47.1%
1975	0.5%	22.1%	**1998**	48.1%	69.7%	**1975**	-2.4%	12.6%	**1998**	20.0%	48.7%
1976	5.1%	24.4%	**1999**	46.7%	71.6%	**1976**	-1.8%	14.3%	**1999**	21.5%	50.0%
1977	9.9%	26.5%	**2000**	52.4%	73.6%	**1977**	-1.7%	15.6%	**2000**	23.6%	51.5%
1978	10.1%	28.5%	**2001**	56.3%	75.7%	**1978**	-4.2%	17.0%	**2001**	27.1%	53.0%
1979	11.1%	30.4%	**2002**	58.5%	78.0%	**1979**	-5.7%	18.4%	**2002**	28.2%	54.5%
1980	14.3%	32.5%	**2003**	61.9%	80.2%	**1980**	-2.6%	20.0%	**2003**	30.1%	56.2%
1981	28.2%	34.6%	**2004**	60.9%	82.3%	**1981**	1.7%	21.4%	**2004**	31.5%	57.8%
1982	31.3%	36.6%	**2005**	63.5%	84.4%	**1982**	0.6%	22.8%	**2005**	32.9%	59.3%
1983	33.2%	38.5%	**2006**	64.5%	86.4%	**1983**	2.3%	24.1%	**2006**	33.5%	60.8%
1984	31.1%	40.2%	**2007**	67.1%	88.5%	**1984**	4.2%	25.3%	**2007**	32.2%	62.3%
1985	29.1%	42.1%	**2008**	67.1%	90.5%	**1985**	3.4%	26.8%	**2008**	32.2%	63.8%

Table 5. Transition Probabilities between Fee Quintiles

The table summarizes 1-year and 2-year transition probabilities between fee quintiles for our sample of mutual funds (see Table A in the Appendix for a detailed description of the sample). In Panel A and C we look at expense ratios. In Panel B and D, the residual fee is defined as the regression residual of Spec. 1 (pre-1999) and Spec. 4 (post-1999) as specified in Panel A of Table 2. In Panels A and B, we determine the quintile thresholds using all funds within one year. In Panels C and D, we determine the quintile thresholds using only "old" funds (i.e., funds that were in the sample in the previous year) within one year. The data covers the period of 1963 to 2008 and is a yearly panel. Variables are defined in Table B in the Appendix.

Panel A. Expense Ratios (All Funds)

	Low Fees (t-12)	**2 (t-12)**	**3 (t-12)**	**4 (t-12)**	**High Fees (t-12)**
Low Fees	87.0%	11.0%	0.7%	0.7%	0.6%
2	8.4%	72.7%	16.5%	1.7%	0.7%
3	1.1%	11.3%	68.7%	17.4%	1.5%
4	0.5%	1.6%	11.8%	69.8%	16.3%
High Fees	0.7%	0.5%	1.6%	12.4%	84.7%
	Low Fees (t-24)	**2 (t-24)**	**3 (t-24)**	**4 (t-24)**	**High Fees (t-24)**
Low Fees	80.9%	16.1%	1.4%	0.8%	0.7%
2	10.7%	61.4%	24.1%	2.7%	1.0%
3	1.9%	14.1%	57.3%	24.0%	2.7%
4	0.7%	2.5%	14.0%	59.5%	23.3%
High Fees	0.8%	0.7%	1.8%	15.6%	81.0%

Panel B. Residual Fees (All Funds)

	Low Fees (t-12)	**2 (t-12)**	**3 (t-12)**	**4 (t-12)**	**High Fees (t-12)**
Low Fees	70.1%	19.4%	5.9%	2.5%	2.0%
2	20.6%	46.1%	22.5%	7.7%	3.0%
3	6.6%	21.8%	43.2%	21.9%	6.5%
4	3.4%	7.9%	21.3%	48.8%	18.5%
High Fees	2.4%	4.1%	6.1%	17.7%	69.7%
	Low Fees (t-24)	**2 (t-24)**	**3 (t-24)**	**4 (t-24)**	**High Fees (t-24)**
Low Fees	62.2%	22.5%	7.5%	4.5%	3.2%
2	23.9%	38.5%	23.8%	9.1%	4.6%
3	9.1%	22.9%	36.7%	23.2%	8.1%
4	4.7%	11.0%	21.5%	41.5%	21.4%
High Fees	3.2%	4.7%	8.7%	20.9%	62.5%

Panel C. Expense Ratios (Old Funds Only)

	Low Fees (t-12)	2 (t-12)	3 (t-12)	4 (t-12)	High Fees (t-12)
Low Fees	86.8%	11.3%	0.7%	0.4%	0.8%
2	8.3%	71.6%	17.5%	2.1%	0.5%
3	1.0%	11.1%	67.5%	18.4%	1.9%
4	0.5%	1.3%	12.5%	69.9%	15.8%
High Fees	0.7%	0.4%	1.5%	10.4%	87.0%
	Low Fees (t-24)	**2 (t-24)**	**3 (t-24)**	**4 (t-24)**	**High Fees (t-24)**
Low Fees	80.3%	16.5%	1.5%	0.7%	0.9%
2	10.8%	60.1%	25.0%	3.0%	1.1%
3	2.0%	13.3%	56.9%	25.2%	2.7%
4	0.6%	2.4%	13.9%	59.9%	23.2%
High Fees	0.7%	0.5%	1.7%	12.9%	84.1%

Panel D. Residual Fees (Old Funds Only)

	Low Fees (t-12)	2 (t-12)	3 (t-12)	4 (t-12)	High Fees (t-12)
Low Fees	70.1%	19.2%	6.2%	2.3%	2.1%
2	21.5%	44.9%	22.7%	7.7%	3.1%
3	7.5%	22.0%	41.7%	22.2%	6.7%
4	3.8%	7.5%	20.9%	47.8%	20.1%
High Fees	2.7%	4.1%	5.9%	17.9%	69.5%
	Low Fees (t-24)	**2 (t-24)**	**3 (t-24)**	**4 (t-24)**	**High Fees (t-24)**
Low Fees	63.0%	21.4%	8.1%	4.2%	3.4%
2	25.0%	37.9%	22.6%	9.7%	4.9%
3	10.1%	22.3%	36.2%	23.0%	8.4%
4	4.8%	11.1%	21.0%	40.4%	22.8%
High Fees	3.5%	4.9%	8.6%	20.4%	62.6%

Table 6. Determinants of First Fees

The table summarizes results from cross-sectional regressions (per fund there is one observation) that use the first reported expense ratio of a fund as their dependent variable. Our sample consists of domestic equity mutual funds (see Table A in the Appendix for a detailed description of the sample). The data covers the period of 1963 to 2008 and is a yearly panel. Variables are defined in Table B in the Appendix.

	Number of Funds in Family > 10				Number of Funds in Family > 100			
	Value-Weighted Family Char.		Equal-Weighted Family Char.		Value-Weighted Family Char.		Equal-Weighted Family Char.	
Variables	**Coeff.**	**t-value**	**Coeff.**	**t-value**	**Coeff.**	**t-value**	**Coeff.**	**t-value**
$\mathbf{ln\_mtna_t}$	-0.0004	-12.51	-0.0003	-11.04	-0.0003	-8.23	-0.0003	-6.94
$\mathbf{ln\_mgmt\_mtna_t}$	-0.0002	-5.72	-0.0003	-6.74	-0.0004	-5.52	-0.0005	-5.01
$\mathbf{avgFamFee_t}$	0.3531	14.56	0.4922	18.33	0.1314	3.49	0.3016	6.28
$\mathbf{sdFamFee_t}$	0.1425	2.57	0.0871	1.42	0.3259	3.93	0.3256	3.07
$\mathbf{avgFamPerf_t}$	0.0031	1.75	0.0048	2.32	0.0092	2.9	0.013	3.43
$\mathbf{sdFamPerf_t}$	0.0075	3.31	0.0026	1.24	-0.0008	-0.22	0.0011	0.34
$\mathbf{beta\_mkt_{t+3}}$	0.0024	9.32	0.0026	10.28	0.0035	10.75	0.0038	11.69
$\mathbf{yret_{mkt,t}}$	-0.0136	-5.13	-0.0111	-4.19	-0.0196	-3.88	-0.0102	-2.01
$\mathbf{beta\_mkt_{t+3}}$ x $\mathbf{yret_{mkt,t}}$	0.0033	2.1	0.0026	1.68	-0.0003	-0.18	-0.0005	-0.29
$\mathbf{beta\_smb_{t+3}}$	0.0008	3.38	0.0009	3.97	0.0005	1.45	0.0006	1.75
$\mathbf{yret_{smb,t}}$	-0.0072	-4.69	-0.0057	-3.72	-0.0119	-4.37	-0.0081	-2.95
$\mathbf{beta\_smb_{t+3}}$ x $\mathbf{yret_{smb,t}}$	0.0074	3.91	0.0062	3.33	0.0095	3.56	0.0087	3.27
$\mathbf{beta\_hml_{t+3}}$	-0.001	-6.04	-0.001	-6.01	-0.0011	-4.86	-0.0011	-5.26
$\mathbf{yret_{hml,t}}$	-0.0051	-2.96	-0.0045	-2.6	-0.0119	-3.41	-0.0047	-1.28
$\mathbf{beta\_hml_{t+3}}$ x $\mathbf{yret_{hml,t}}$	0.0044	4.29	0.0036	3.63	0.0031	2.33	0.0034	2.59
$\mathbf{beta\_umd_{t+3}}$	-0.0003	-0.97	-0.0004	-1.64	-0.0003	-0.92	-0.0004	-1.14
$\mathbf{yret_{umd,t}}$	-0.0029	-2.54	-0.0029	-2.55	-0.0038	-2.29	-0.0029	-1.72
$\mathbf{beta\_umd_{t+3}}$ x $\mathbf{yret_{umd,t}}$	0.0043	2.86	0.0049	3.28	0.0052	2.52	0.0048	2.36
inst_flag	-0.0054	-33.89	-0.0052	-33.14	-0.005	-26.27	-0.0049	-25.87
etf_flag	-0.0015	-3.29	-0.0003	-0.7	-0.0037	-5	-0.0016	-2.21
Constant	0.0140	18.24	0.0115	14.46	0.0190	12.30	0.0136	7.89
Year FE	Yes		Yes		Yes		Yes	
# of funds.	4935		4948		3288		3297	
R-Squared	42.8%		45.4%		38.7%		39.7%	

Table 7. Determinants of Fee Changes

The table reports results of Fama-MacBeth regressions, in which we use the yearly changes in expense ratios as dependent variables (see Table A in the Appendix for a detailed description of the sample). The data covers the period of 1963 to 2008 and is a yearly panel. Variables are defined in Table B in the Appendix. Δ indicates that we calculate simple changes of our base variables. Variables with subscripts *t-1* are lagged by one year. Some information is only available after 1999 (e.g., information on fund families) and, thus, we split the sample into a pre-1999 and a post-1999 subset.

	Pre 1999						Post 1999					
	Spec. 1		Spec. 2		Spec. 3		Spec. 4		Spec. 5		Spec. 6	
	Coeff.	**t-value**	**Coeff.**	**t-value**	**Coeff.**	**t-value**	**Coeff.**	**t-value**	**Coeff.**	**t-value**	**Coeff.**	**t-value**
$Yret_{t-1}$	-0.0009	-1.14	-0.0014	-2.22			-0.0016	-9.17	-0.0016	-8.13		
Positive $Yret_{t-1}$					-0.0006	-0.62					-0.0014	-7.79
Negative $Yret_{t-1}$					-0.0008	-0.11					-0.0042	-2.30
Δbeta_mkt	0.0010	0.81	-0.0004	-0.31	-0.0004	-0.36	-0.0002	-0.61	-0.0003	-0.73	-0.0003	-0.76
Δbeta_hml	-0.0004	-0.50	0.0003	0.51	0.0003	0.50	-0.0001	-0.31	0.0001	0.48	0.0001	0.53
Δbeta_smb	0.0007	0.82	0.0003	0.69	0.0003	0.61	0.0006	1.76	0.0005	1.52	0.0004	1.35
Δbeta_umd	0.0010	1.03	0.0003	0.37	0.0001	0.13	0.0004	0.93	0.0004	0.97	0.0004	0.96
Δr_squared	-0.0011	-1.34	-0.0011	-1.27	-0.0010	-1.13	-0.0001	-0.13	-0.0002	-0.39	-0.0002	-0.36
First_exp_rati			-0.0106	-0.41	-0.0133	-0.53			-0.0107	-2.53	-0.0109	-2.52
Comp_funds	0.0000	-1.20	0.0000	-1.19	0.0000	-0.93	0.0000	0.46	0.0000	0.81	0.0000	0.86
Comp_fees	-0.0153	-0.73	-0.0102	-0.42	-0.0097	-0.45	-0.0121	-0.29	0.0116	0.24	0.0056	0.12
$flow_{t-1}$	-0.0003	-1.45	-0.0002	-1.40	-0.0002	-1.36	0.0000	-2.39	0.0000	-2.25	0.0000	-2.45
Δln_mtna	-0.0008	-3.02	-0.0007	-3.84	-0.0007	-3.60	-0.0001	-1.41	-0.0001	-1.53	0.0000	-1.54
fund_age	-0.0000	-0.37	0.0000	-0.71	0.0000	-0.70	0.0000	3.23	0.0000	1.08	0.0000	1.06
Δsd_mret	-0.0205	-1.05	-0.0130	-0.73	-0.0102	-0.58	0.0045	0.73	0.0054	0.78	0.0062	0.92
Δturn_rat							0.0000	1.33	0.0001	1.48	0.0001	1.48
Δln_mgmt_mtna							-0.0001	-2.04	-0.0001	-1.70	-0.0001	-1.66
part_family100							0.0000	0.38	0.0000	0.50	0.0000	0.56
inst_flag							0.0001	1.46	0.0000	0.11	0.0000	0.10
open_flag							0.0002	2.70	0.0002	2.71	0.0002	2.72
etf_flag							0.0000	0.53	0.0000	-0.44	0.0000	-0.51
Constant	0.0007	1.47	0.0008	1.56	0.0007	1.45	0.0001	0.11	-0.0001	-0.13	-0.0001	-0.07
# of years	33		33		33		10		10		10	
# of obs.	11780		9625		9625		28869		26502		26502	
Avg. R-Squared	13.4%		15.7%		17.0%		5.9%		7.0%		7.2%	

Table 8. Performance-Flow Sensitivity

The table shows the means, standard deviations and medians of flows and expense ratio changes of groups of funds. All statistics are calculated from pooled data. We group funds by performance: low (high) performance funds are funds in the bottom (top) decile of past year's returns. We also split funds by different measures of fees into "cheap" (below median) and "expensive" (above median) funds. Performance deciles and expense medians are calculated separately for each year.

			First Expense Ratio		Expense Ratio		Residual Fee	
			Below Median	Above Median	Below Median	Above Median	Below Median	Above Median
Yearly $Return_{t-1}$	Low	Mean Flows	3.6%	-1.3%	4.7%	-2.1%	7.5%	-7.3%
		SD of Flows	113.1%	115.6%	128.4%	106.0%	130.1%	94.7%
		Median Flows	-11.1%	-15.4%	-11.2%	-15.3%	-11.1%	-16.8%
		Mean Fee Changes	3 bp	3 bp	-1 bp	5 bp	0 bp	7 bp
		SD of Fee Changes	20 bp	28 bp	13 bp	30 bp	17 bp	32 bp
		Median Fee Changes	0 bp	0 bp	0 bp	0 bp	0 bp	1 bp
		Funds	1434	2692	1507	2619	2153	1973
	High	Mean	79.4%	65.3%	72.4%	69.6%	86.2%	56.9%
		SD	258.5%	188.9%	235.1%	203.7%	272.4%	154.6%
		Median	14.0%	15.4%	15.3%	14.3%	15.9%	14.2%
		Mean Fee Changes	-2 bp	-6 bp	-3 bp	-5 bp	-5 bp	-4 bp

Table 9. Captive-Investor Hypothesis

The table shows descriptive statistics related to the captive-investor hypothesis. Panel A reports means and standard deviations of variables related to the captive-investor hypothesis by residual fee deciles. The last column reports the difference between Decile 1 and Decile 10. Stars indicate significance at the 1% (***), 5% (**) and 10% (*) level. Panel B reports mean residual fees and fee spreads for groups of funds. Variables are defined in Table B in the Appendix.

Panel A. Distribution of Fee Residuals

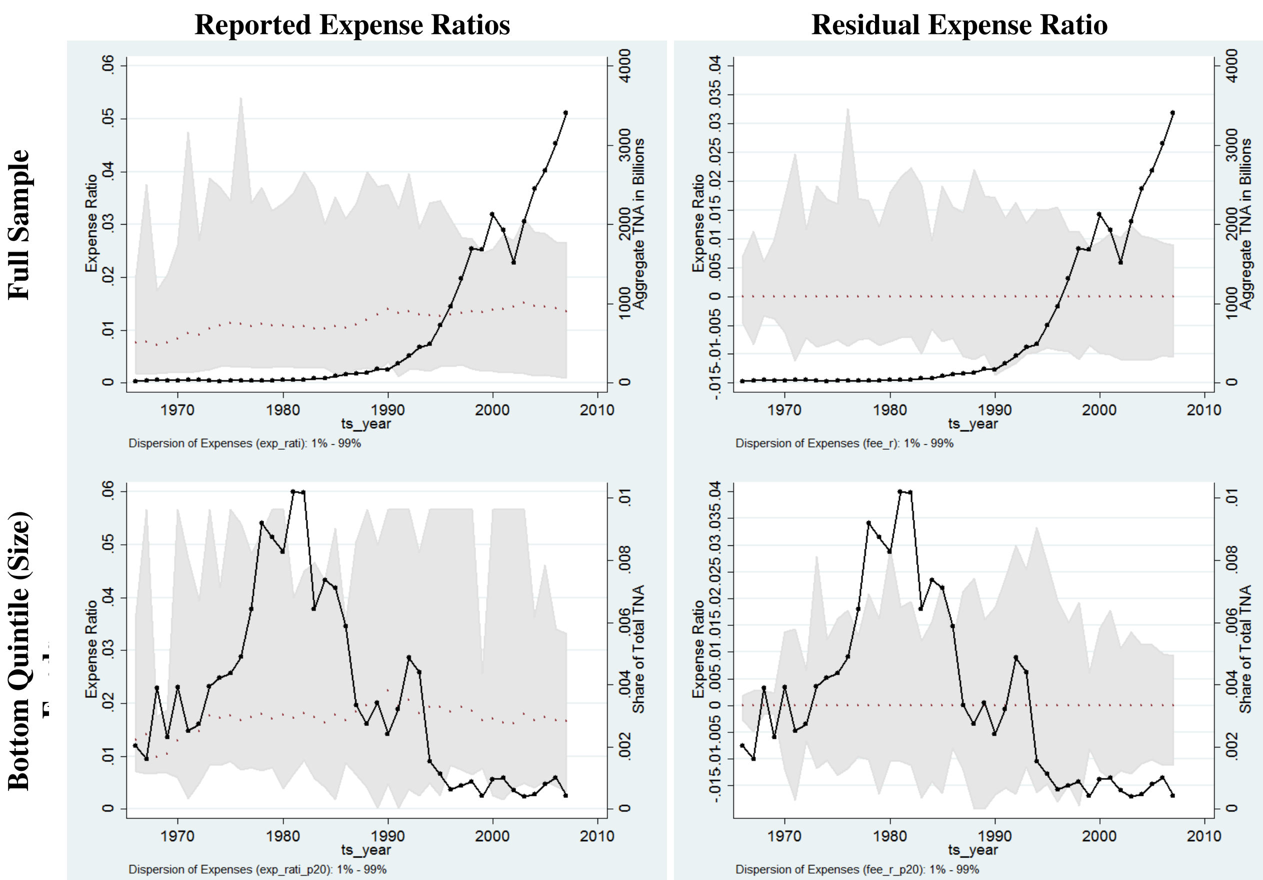
	Decile 1		Decile 3		Decile 5		Decile 7		Decile 10		Decile 1-10
	Mean	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Diff.**
12b-1 fees	0.0011	0.0021	0.0016	0.0024	0.0024	0.0032	0.0046	0.0042	0.0087	0.003	0.0075***
Front-End Load	0.0235	0.0223	0.0271	0.0211	0.0287	0.0206	0.028	0.0236	0.0145	0.0271	-0.0090***
Back-end Load	0.0023	0.0036	0.0025	0.0045	0.0034	0.0057	0.0056	0.0076	0.01	0.0087	0.0077***
Flow AutoCorrel.	0.163	0.292	0.163	0.298	0.161	0.309	0.204	0.327	0.380	0.376	0.218***
***Easy-In Hard-Out* Funds**	0.0%	0.0%	0.0%	0.0%	0.0%	5.1%	0.6%	6.4%	4.6%	20.4%	-4.6%***
"*Pension Plan*" Funds	4.8%	20.5%	4.5%	22.5%	6.5%	26.4%	8.5%	29.7%	23.2%	41.5%	-18.5%***

Panel B. Mean Fee Residual and Mean Spreads

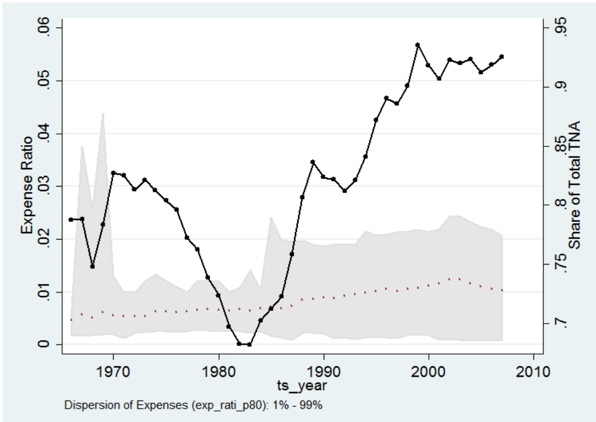
	Mean Residual Fee (bp)	Mean Spread (bp) 75%-25%	Mean Spread (bp) 90%-10%	Mean Spread (bp) 99%-1%
Full Sample	0	44	89	234
Full Sample minus *Easy-In Hard-Out* Funds	17	51	100	205
Full Sample minus "*Pension Plan*" Funds	-1	42	87	234
***Easy-In Hard-Out* Funds**	54	30	60	97
"*Pension Plan*" Funds	24	53	98	192

Figure 1. Fund Fee Dispersion

The figure shows the fee dispersion (99%-percentile vs. 1%-percentile) of expense ratios (left column) and residual expense ratios (right column) across funds and over time. In addition, all graphs include the mean expense ratios and the mean residual expense ratios (these are, by definition, lines in the origin). The means are represented by the dotted lines. Graphs in the top row also plot the aggregate TNA of all funds in the graph in Billions of USD (connected, dark line). In rows 2 and 3 we include a line (connected, dark line) that represents the fraction of aggregate TNA represented by funds in the bottom size quintile (row 2) and the top size quintile (row 3) of our sample. For the "Full Sample", the residual fee is defined as the regression residual of Spec. 1 (pre-1999) and Spec. 4 (post-1999) as specified in Panel A of Table 2. To get the residual fee distribution used in rows 2 and 3, we re-estimate Spec. 1 and Spec. 4 of Panel A (Table 2) for the subsamples of smallest (bottom quintile) and largest (top quintile) funds per year. Coefficient estimates of these regressions are consistent with the ones reported in Table 2 and available from the authors upon request. Our sample consists of domestic equity mutual funds (see Table A in the Appendix for a detailed description of the sample). The data covers the period of 1963 to 2008 and is a yearly panel. Variables are defined in Table B in the Appendix.



Top Quintile (Size) Funds



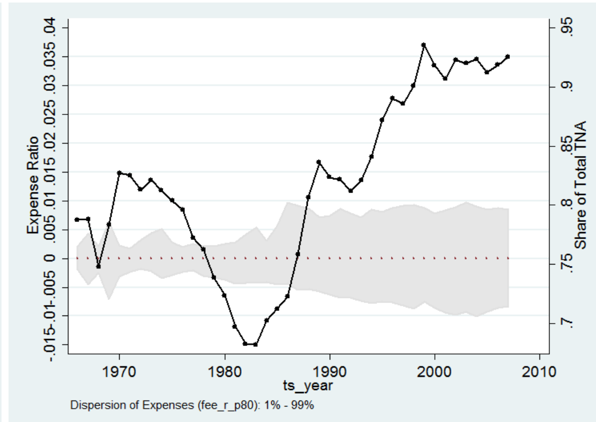


Figure 2. Fund Fee Dispersion within and outside of Fund Families

The figure shows the fee dispersion (99%-percentile vs. 1%-percentile) of expense ratios (left column) and residual expense ratios (right column) across funds that are part of a fund family (top row) and funds that are not (bottom row). For these plots management companies with more than 100 funds are defined as fund families. In addition, all graphs include the mean expense ratios and the mean residual expense ratios (these are, by definition, lines in the origin). The means are represented by the dotted lines. The residual fee is defined as the regression residual of Spec. 4 (post-1999) as specified in Panel A of Table 2. Our sample consists of domestic equity mutual funds (see Table A in the Appendix for a detailed description of the sample). The data in this analysis covers the period of 1999 to 2008 and is a yearly panel. Variables are defined in Table B in the Appendix.

Reported Expense Ratios | **Residual Expense Ratio**

Within Family

Outside Family

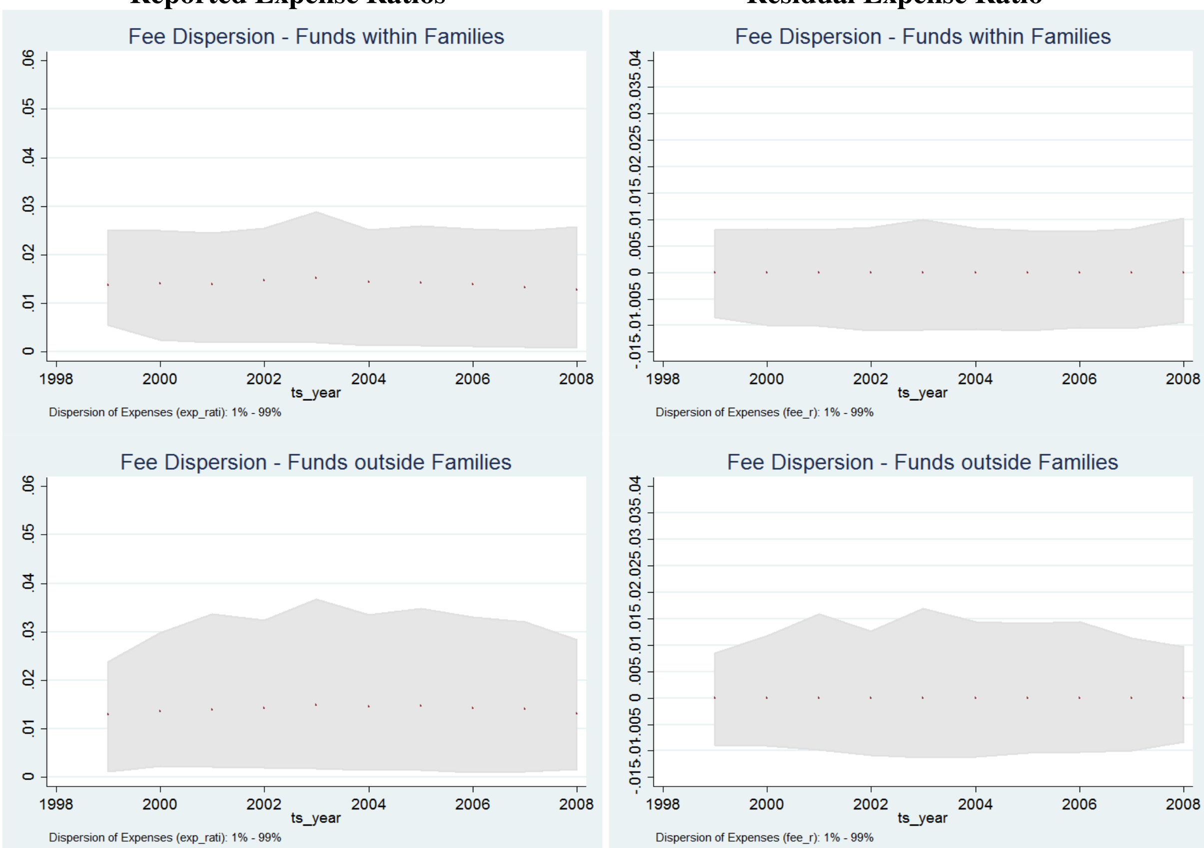


Figure 3. Fund Fee Dispersion of Institutional and Retail Funds

The figure shows the fee dispersion (99%-percentile vs. 1%-percentile) of expense ratios (left column) and residual expense ratios (right column) across funds that are institutional funds (top row) and funds that are retail funds (bottom row). In addition, all graphs include the mean expense ratios and the mean residual expense ratios (these are, by definition, lines in the origin). The means are represented by the dotted lines. The residual fee is defined as the regression residual of Spec. 4 (post-1999) as specified in Panel A of Table 2. Our sample consists of domestic equity mutual funds (see Table A in the Appendix for a detailed description of the sample). The data in this analysis covers the period of 1999 to 2008 and is a yearly panel. Variables are defined in Table B in the Appendix.

Reported Expense Ratios | **Residual Expense Ratio**

Institutional Funds

Retail Funds

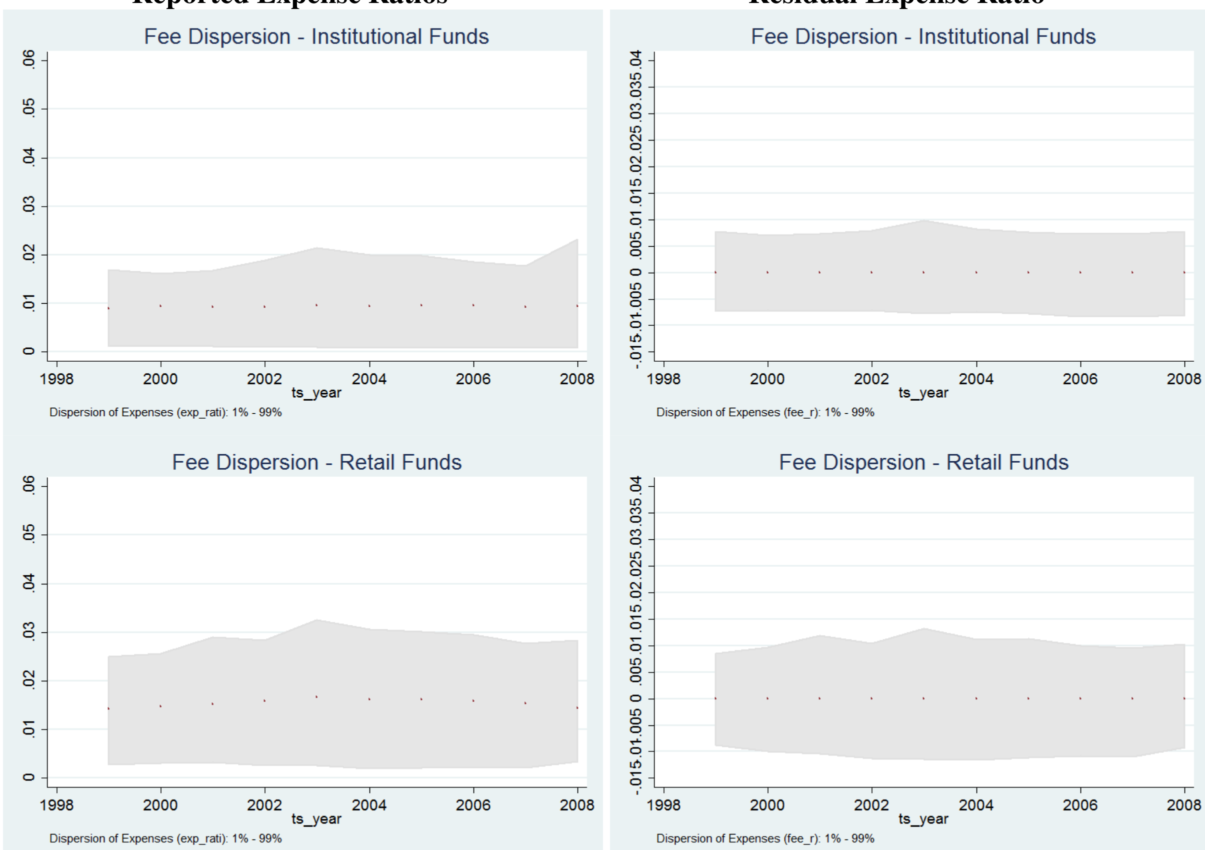
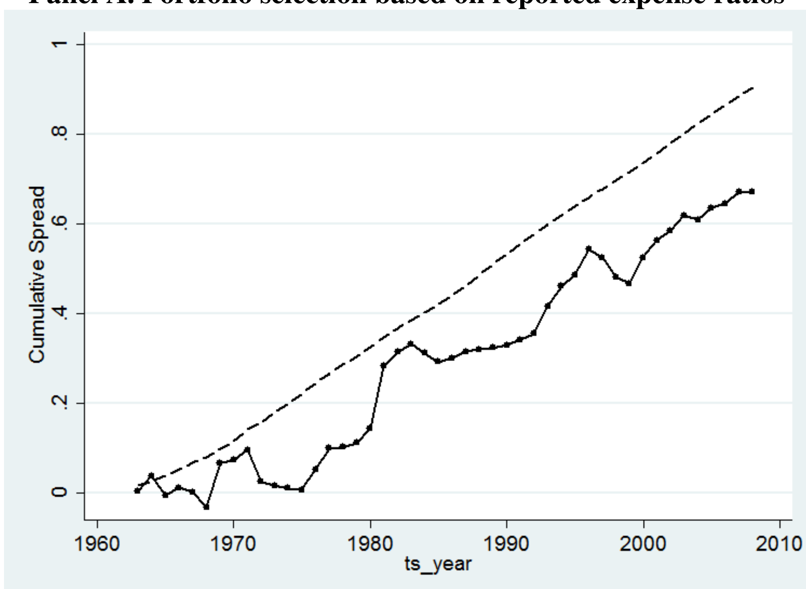


Figure 4. Evaluation of Trading Strategy

The figure summarizes the cumulative Carhart alpha (the solid line with markers) of a strategy that buys funds, which are in the bottom decile according to reported expense ratios (residual expense ratios), and shorts funds, which are in the top deciles according to reported expense ratios (residual expense ratios). The table also reports the cumulative spread between the average reported expense ratio (residual expense ratio) of funds in the top and the bottom decile (the dashed line). The residual fee is defined as the regression residual of Spec. 1 (pre-1999) and Spec. 4 (post-1999) as specified in Panel A of Table 2. Our sample consists of domestic equity mutual funds (see Table A in the Appendix for a detailed description of the sample). The data covers the period of 1963 to 2008 and is a yearly panel. Variables are defined in Table B in the Appendix.

Panel A. Portfolio selection based on reported expense ratios



Panel B. Portfolio selection based on residual fees